# NORMATIVE AUTHORITY AND THE FOUNDATIONS OF ETHICS

by

Matthew E. Silverstein

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Philosophy)
in the University of Michigan
2008

Doctoral Committee:

    Professor Peter A. Railton, Chair
    Professor Stephen L. Darwall
    Professor Allan F. Gibbard
    Associate Professor Mika T. Lavaque-Manty

For Bertie

## ACKNOWLEDGEMENTS

# CONTENTS

# INTRODUCTION

> But why, we ask…. Because … And when we have gleaned that reason we will want
> to know, But why *that* reason? Because … And when we have that further reason,
> But why again—? Because … Why?… Because … Why?… Until … that incessant
> question Whywhywhy has become like a siren wailing in our heads and a further
> question begins to loom: when—where—how do we stop asking why?
> —Graham Swift, "About the Question Why," *Waterland*

My wife and I are having a "discussion." We're going to invite our dear friend Helen to a small dinner party, but we're debating whether we should include her husband. You see, we enjoy her company but not his, and the evening will be much more pleasant for both us and our other guests if Helen comes alone. I contend that we should ask her to leave her lesser half at home. My wife disagrees.

What are we disagreeing *about*, though? There are at least a couple of possibilities. We might, for example, be differing over what the rules of etiquette require. My wife is convinced that these rules dictate that invitations should always include spouses, even when those spouses are irritating beyond belief. I then observe that on more than one occasion I've opened the mailbox to find an invitation addressed only to her. She has attended countless bridal showers over the past several years, and I have yet to be invited to one! She insists that bridal showers fall into a special category. I reply by suggesting that our little get-together also falls into a special category: parties that exclude annoying husbands. Fed up with my lack of social grace, my wife dryly observes that such a category would exclude me as well.

Now, if we are simply disagreeing about what the rules of etiquette require, we can resolve our dispute by consulting an authority on those rules. Sure enough, the foremost authority on all matters of etiquette—Miss Manners—rules in favor of my wife: it would indeed be rude not to invite the annoying spouse. And so the matter is settled.

Or perhaps not. I might concede that etiquette requires us to include our friend's husband yet still wonder whether we would be justified in leaving his name off of the invitation envelope. That is, I might accept Miss Manners' verdict as valid from the standpoint of etiquette but challenge the authority of the standpoint itself. Yes, the rules of etiquette dictate that we should invite him. But why—I inquire of my wife—should we abide by those rules? Why are the norms of etiquette the ones we should follow rather than some other norms? Notice that she cannot look to Miss Manners for answers to these questions. Etiquette's demand that we follow the rules of etiquette is beside the point, for I am challenging the reasonableness of etiquette itself. Since I wonder whether I should even be concerned with what is rude, it will not help to point out that it would be rude to be unconcerned. And so, in order to convince me that we should invite Helen's troublesome husband, my wife must appeal to some other set of norms— the norms of morality, for example.

As it happens, that is precisely what she does. She notes that if we do not observe the rules of etiquette, poor Helen will be offended and hurt, and we're morally obligated not to hurt our friends. That's why we should adhere to the rules of etiquette, my wife explains: because the norms of morality require it! I then respond with another challenge: Why should we adhere to the norms of morality? I admit that morality require us to invite Helen's husband, but why should we follow those norms rather than some others? At this point my wife appeals to the norms self-interest. No one wants to spend time with reprobates. And so if we refuse to conform our behavior to the norms of morality, then sooner or later we won't have any friends at all and we'll be lonely and miserable. But why, I can't help asking, should we care about the norms of self-interest? Sure, it would be better for us in the long run to abide by the norms of morality, but why should we care about what is better for us in the long run? "We just should," my wife sighs as she turns away and wonders again why she married a philosopher.

### The Regress of Justification

This connubial fable reveals a troubling aspect of the ways in which we justify our actions. When I ask why we should heed the norms of morality, I am asking for a reason to be moral. But a reason is a consideration that justifies; and to justify something is, as David Velleman

puts it, "to show that it is just, in the old-fashioned sense meaning 'correct.' Something is subject to justification only if it is subject to a *jus*, or norm of correctness."[1]

Yet norms are everywhere: we are awash in an ocean of norms and rules. There are the rules of etiquette, the norms of morality, the norms of good taste, the rules of law, the norms of prudence, the norms of instrumental reason, and so forth. Most of us accept these norms most of the time. There are also norms we accept only in certain circumstances, such as the norms of business ethics, the norms of psychotherapy, the rules of basketball, and the norms of academia. And of course there are plenty of norms that most of us reject: the norms of white supremacy, the norms of rugged individualism, the norms of anti-Semitism.

This plethora of norms gives rise to a problem. An action that is justified according to some standards will invariably be forbidden by others. For example, whereas spitting in my face is probably condemned by the norms of morality and etiquette, it is heartily recommended by the norms of anti-Semitism. Often there is conflict even between norms we accept, as when what I ought to do prudentially speaking is not what I ought to do morally speaking. However, we do not think that all norms are created equal. We believe that we should heed the norms of morality but reject the norms of white supremacy. We hold the former to be *authoritative* or justified in a way that the latter are not. The problem arises when we try to spell out exactly *how* these norms are justified, for to be justified is just to be correct according to some norm or standard. Now, perhaps certain norms are *self*-justifying. The norms of morality, for example, require that we follow the norms of morality. But, as Stephen Darwall has observed, "this response hardly satisfies," since the justification it offers is "wholly internal to the system of reasons being questioned."[2] Darwall adds:

> It is no more satisfying a response to the philosophical request for justification than would 'because etiquette requires it' be a satisfying response to the question of why we must abide by its demands.[3]

And as I suggested above in my description of my marital dispute, that is indeed an unsatisfying response. So, if a particular norm is justified, it must be in virtue of some *other* norm. But that further norm will support the norm with which we started only if it is itself justified. And so we must ask whether this further norm is justified, which calls for us to look

---

1    Velleman, introduction to *The Possibility of Practical Reason*, 15.

2    Darwall, "Autonomist Internalism and the Justification of Morals," 257.

3    Darwall, "Autonomist Internalism and the Justification of Morals," 257.

to yet another norm. We now face a seemingly infinite *regress of justification*. The norms of morality are justified only if the norms that justify them are themselves justified, and they are justified only if the norms that support them are justified. And so on and so on.[4]

If there is no end to this justificatory regress, then justification will only ever be relative or qualified: justification according to these norms or those. But could this be the sort of justification we are after when we ask what we ought to do, or the sort of justification we claim when we defend our actions by citing our reasons? William Frankena thought so. In "On Saying the Ethical Thing," Frankena wrote:

> The reasons … claimed by normative judgments may be of different sorts: aesthetic, legal, moral, prudential, perhaps even religious.… Each type of reason may be conceived of as relative to a 'point of view,' and each type of normative judgment claims that something is justified … (or the opposite) from some such point of view.[5]

According to Frankena, relative justification is the only kind of justification there is. But as Darwall has argued in response, deliberation seems to involve the search for a more robust and genuine sort of justification. For one thing, relative or qualified justification is not worth much: nearly every set of norms is supported by at least one other set of norms. Even the norms of anti-Semitism, for instance, are sanctioned by the norms of National Socialism. Moreover, when I deliberate about what I ought to do, I often already know what the norms of morality and prudence and etiquette require. That is, I already know what I morally ought to do and what I prudentially ought to do (and what Miss Manners thinks I ought to do). What I want to know in such situations is what I *really* ought to do—what I ought to do in some *unqualified* sense. I am after not moral or prudential justification but rather unqualified justification—justification *simpliciter*.[6] As Darwall notes, to regard a norm as authoritative is

---

4   My conception of the problem of the regress of justification has been shaped most significantly by the work of David Velleman ("The Possibility of Practical Reason," and the introduction to *The Possibility of Practical Reasoning*), Peter Railton ("Some Questions about the Justification of Morality" and "On the Hypothetical and Non-hypothetical"), and Stephen Darwall ("Autonomist Internalism and the Justification of Morals" and "Internalism and Agency"). Other work that has influenced my thinking on this subject includes Korsgaard, *The Sources of Normativity*; Gibbard, "Morality as Consistency in Living"; Radzik, "A Normative Regress Problem" and "Justification and the Authority of Norms"; and Dreier, "Humean Doubts about the Practical Justification of Morality."

5   Frankena, "On Saying the Ethical Thing," 120. See Foot, "Morality as a System of Hypothetical Imperatives," 169–70, for a similar view.

6   The phrases "unqualified justification" and "justification *simpliciter*" both come from Darwall. See Darwall, "Autonomist Internalism and the Justification of Morals," 258, and "Internalism and Agency," 156. Compare Dreier, "Humean Doubts about the Practical Justification of Morality," 83–4; and contrast Copp, "The Ring

really just to treat is as justified *simpliciter.* "To people who accept etiquette's norms, reasons of etiquette present themselves as [unqualified] reasons, relevant to what a person should do ([unqualifiedly])."[7]

That deliberation presupposes the possibility of justification *simpliciter* is one thing. That there are norms that really are justified *simpliciter* is something else. How are such norms possible? How could there be norms or reasons whose authority is unqualified? Since it is the regress of justification that gives rise to these questions, an answer can be found only by halting that regress. But how are we going to do *that?* Stopping the justificatory regress would seem to require norms that are justified not in virtue of any other norms but rather entirely on their own. It would, in other words, seem to require norms that are intrinsically justified. Above, however, I defined justification in a way that makes the notion of an intrinsically justified norm incoherent. I stated that to justify something is to show that it is correct according to some norm or standard. This rules out the possibility of norms that are justified independently of any other norm.

David Copp expresses a similar concern. He acknowledges that when we deliberate, "we want to make the decision that is best *period*."[8] A morally virtuous person does what she thinks is morally best because she believes that "to do what is morally best *is* to do what is best, period."[9] But Copp also contends that deliberative conclusions about what is "best *period*" are always false. He holds that there is no such thing as justification *simpliciter*—that "there are only the various reasons of the various special kinds."[10]

Copp's view amounts to an error theory about "the unity of practical reason."[11] His defense of this theory revolves around his claim that the idea of a norm that is somehow justified above

---

of Gyges," 86; and Radzik, "Justification and the Authority of Norms," 455–9. It is important not to conflate unqualified reasons with what are commonly called all-things-considered reasons. An unqualified reason to perform one action can still be overridden or outweighed by a stronger, competing unqualified reason to do something else.

7   Darwall, "Learning from Frankena," 699.

8   Copp, "The Ring of Gyges," 100.

9   Copp, "The Ring of Gyges," 100.

10   Copp, "The Ring of Gyges," 103.

11   Copp, "The Ring of Gyges," 103.

all others (or is "normatively supreme") is incoherent.[12] Here is a condensed version of Copp's *reductio:*

> The claim that a standpoint has the property of supremacy is the claim that it is the *normatively most important* standpoint…. Hence, the claim that the candidate *S* has the property of supremacy is the claim that it is normatively more important than any other standpoint, as assessed from a relevant authoritative standpoint. That is, if *S* is normatively most important, then there is some authoritative standard *R* that yields the verdict that *S* is normatively the most important standpoint…. Now, either standard *R* is identical to *S*, or it is not. We cannot suppose that *R* is identical to *S*. For a standard cannot be normatively the most important in virtue of its meeting criteria that it *itself* specifies…. Therefore, if *S* is normatively the most important standpoint, then it meets the criteria specified by some authoritative standpoint *R* that is distinct from *S*. This standard *R* must be normatively the most important standard…. Otherwise, there would be some standpoint superior to *R*, and *its* assessment of the relative importance of *S* and the special standpoints would be the definitive one. But then, if *R* is normatively more important than *S*, it follows that *S* is not in fact the normatively most important standpoint …, which contradicts the assumption with which we began…. It appears, then, that the idea of [justification *simpliciter*] is incoherent.[13]

I have quoted Copp's argument at length because I think his reasoning is cogent. If finding an end to the regress of justification required there to be a standard that is "normatively more important" than any other standard, such a search would indeed be futile. To put the same point in terms I have been employing, if the justificatory regress could be brought to a halt only by norms that are intrinsically justified—that is, only by norms that are justified "all on their own"—then there would be no such thing as what we *really* ought to do. There would be only justification according to this norm or that, only qualified or relative reasons. We would not have any unqualified reasons to do or believe anything. There would, I think, be no genuine normativity or authority at all.

Fortunately, putting a halt to the regress of justification does not require there to be norms that are intrinsically justified in this strict sense. I reject Copp's conclusion because I reject the premise stated in the argument's very first sentence: "The claim that a standpoint has the property of supremacy is the claim that it is the *normatively most important* standpoint." While I accept that an end to the regress depends on the possibility of norms that are "important" or

---

12    Copp, "The Ring of Gyges," 103.

13    Copp, "The Ring of Gyges," 101–3, emphasis in original.

*privileged* in some way, I do not believe, as Copp does, that such norms must be *normatively* privileged. Copp assumes that the norms that stop the regress of justification can do so only if they are privileged in the same way that any norm is privileged when supported by another norm, only more so (or supremely so). And that does seem incoherent. How could there be a norm that is justified just like any other norm except that its justification does not rest on further norms? Resting on further norms is how other norms are justified! But this is not the only way in which a norm can be privileged. To halt the justificatory regress, what we need are not norms or reasons that are intrinsically justified, but rather norms or reasons that somehow resist the demand for justification altogether.[14]

This is precisely what my wife claims to have identified when she declares, "We just should." When she offers this reply to my question about why we should act in our own self-interest, she is not suggesting that the norms of prudence are justified by some other norms. Rather she is claiming that they do not need to be justified, and that there is therefore something wrong with my "Why?" question. I am demanding further normative justification where it is neither possible nor called for. When the regress of justification leads us to a norm that is immune from further "Why?" questions, justification has come to end. To borrow a metaphor from Derek Parfit, we have reached normative bedrock.[15]

This is what justification *simpliciter* requires. If there are any unqualified 'oughts' and genuinely authoritative reasons, they are rooted in norms that are privileged because they do not require justification—because it is somehow inappropriate to question their justification. Philosophers on the lookout for this sort of normative bedrock are engaged in the search for what Christine Korsgaard calls "the source of normativity."[16] This dissertation constitutes my contribution to that search. I shall be concerned with questions about how we can bring the regress of justification to a halt. How are bedrock norms possible? How, as Graham Swift asks, can we stop asking why? What would make the demand for further justification somehow

---

14  Here I differ from Radzik, who claims that "the real challenge to formulating a foundational theory of normative authority is to find some explanation of what it is that makes a norm justified in itself, justified without appeal to any other norm" ("Incorrigible Norms," 637).

15  Parfit, "Normativity, Naturalism, and Noncognitivism," 24.

16  Korsgaard, *The Sources of Normativity*, 18. For perhaps the earliest explicit statement of this approach to the problem (and maybe of the problem itself), see Price, *Review of the Principal Questions in Morals*, as reprinted in Raphael, *British Moralists*, § 672.

inappropriate? In the process of answering these general questions about what the sources of normativity could be, I shall develop and defend an account of what those sources actually are. But I am getting ahead of myself. In the next section I shall delimit more completely the scope of my project.

## The Foundations of Ethics

Strictly speaking, the quest for the sources of normativity takes us well beyond the domain of what has traditionally been called "ethics." Normativity is not confined to the moral or even the practical realm: we can have reasons to believe and to feel as well as to act. That said, I shall be concerned here almost exclusively with practical normativity. My reasons for limiting my investigation in this way are partly autobiographical. I came to the problem of the regress of justification by way of challenges to the justification of morality and questions about the relationship between morality and rationality. I have flirted with various forms of moral skepticism over the years; and at some point I realized that the aspect of morality with which I was struggling—namely, its normativity—was also a feature of various notions I had been taking for granted, such as rationality and well-being. Hence it is the grounds of *ethics*, broadly construed to include the entire practical realm, that I am most keen to uncover and investigate. Another reason for my focus on practical normativity is that there is considerably less controversy about theoretical reason and reasons to believe than there is about practical reason and reasons for action. There are, for example, a number of philosophers who accept an account of the sources of epistemic normativity but deny that a similar account is plausible in the practical realm.[17] One aspect of my project is to challenge this skepticism about the unity of practical reason.

My aim, then, is to explore the sources of practical normativity or, as I shall often call them, the foundations of ethics. In the chapters that follow, I shall sometimes refer explicitly to "*practical* normativity" and "*practical* justification," but more frequently I shall write simply in terms of "normativity," "justification," and "normative authority." It is therefore important to remember that I do not mean to draw conclusions about normativity and justification more broadly. Of course I hope and believe that my arguments about reasons for action can be applied to reasons for belief and other attitudes, but I do not mean to argue for any

17    See note 59 in chapter 4 below.

such conclusions here. As I have already said, it is the foundations of *ethics*—the search for normative bedrock in the practical realm—that shall occupy me in this dissertation.

As my choice of terms reveals, this project presupposes what we might call *normative foundationalism*.[18] I believe that if there are any genuinely authoritative norms and reasons, it is only because those reasons are supported by foundational norms that resist the call for justification. I have already spelled out some of my reasons for believing this. A theory that relied only on qualified or relative justification would not leave room for the sort of justification our practical deliberations take for granted.[19] But of course the best defense of my foundationalist assumption would be actually to discover the foundations in question. And that is precisely what I shall endeavor to do. I hope to find practical norms and reasons for action that are immune to the dreaded "Why?" question and that can therefore serve as starting points for practical reasoning.

This *normative* foundationalism must be not be confused with any sort of *epistemological* foundationalism. The foundations I shall be exploring are the *metaphysical* foundations of normativity.[20] One can accept the possibility and even the existence of such foundations without subscribing to a foundationalist epistemology of normativity—that is, to a foundationalist theory about how we justify our normative and metanormative beliefs. One can, in other words, believe in normative bedrock without holding that our reasons for this belief are rooted in some sort of Cartesian epistemological bedrock. There is no tension here. It may very well be the case that an anti-foundationalist epistemology based on the idea of "reflective equilibrium" is what best supports my conclusions about the foundations of ethics.[21] That is, we might adopt a coherentist theory about the justification of our beliefs and then conclude that a belief in the foundations of ethics is justified.[22]

---

18    Compare Radzik, "Incorrigible Norms."

19    See Radzik, "A Coherentist Theory of Normative Authority," for a different approach.

20    Thanks to Jyl Gentzler for pressing me to clarify these distinctions.

21    For the classic discussion of the method of reflective equilibrium in ethics, see Rawls, *A Theory of Justice*, 18, 40–6. See also Daniels, "Wide Reflective Equilibrium and Theory Acceptance in Ethics"; and Brandt, "The Science of Man and Wide Reflective Equilibrium."

22    Even a foundationalist account of the metaphysical sources of *epistemic normativity* need not be in tension with an anti-foundationalist theory of epistemic justification. Thus, one could accept the widely held view that truth is the standard of correctness for belief while also holding a coherentist theory of justification (so

One final clarification is in order. Up to now I have been using the terms "ethics" and "practical reason" interchangeably. I have been assuming that the foundations of practical reason are also the foundations of ethics broadly construed. But this is a controversial assumption. There are philosophers who think that questions about what I *ought* to do and what I have *reason* to do do not exhaust the subject matter of ethics. Such philosophers usually believe that *values*—that is, questions about what is *good* or *valuable* in various respects—are independent of and perhaps even prior to questions about 'oughts' and reasons.[23] Despite this controversy, I shall continue to treat *ought* and *reason* as the fundamental normative notions. I shall, in other words, continue to assume that all of normativity—including the normativity of value, goodness, obligation, and duty—can be reduced to the normativity of 'oughts' and reasons (and hence justification).[24] Following Wilfrid Sellars, I regard normative concepts as those that are fundamentally "fraught with ought"—or as those that concern "the logical space of reasons."[25] Of course there are also disagreements over whether the notion of *ought* is prior to that of *reason* or vice versa.[26] I regard these ideas as roughly interchangeable.[27] I shall, however,

---

long as one believes that the coherence of one's beliefs is a good indication of their truth). Thanks to Nishi Shah for helpful discussion of this point.

23 The most familiar defender of such a view is probably G. E. Moore. See especially Moore, *Principia Ethica*, chs. 1 and 5. For other accounts that accord an independent place in ethics for value, see Griffin, *Value Judgement*; Crisp, *Reasons and the Good*, 61–7; FitzPatrick, "Reasons, Value, and Particular Agents"; and Väyrynen, "Resisting the Buck-Passing Account of Value." Note that strictly speaking there are at least two possible positions here. The first holds that value involves a type of normativity that is distinct from the normativity of 'oughts' and reasons. The second (and more radical) holds that value is not normative at all but that there is more to ethics than normativity. Then, of course, there are the philosophers who want to leave all this talk about value and normativity behind in favor of a virtue-centered approach.

24 For philosophers who share this assumption, see Sidgwick, *The Methods of Ethics*, 25, 32–4; Ewing, *The Definition of Good*, ch. 5; Frankena, "Obligation and Value in the Ethics of G. E. Moore"; Richards, *A Theory of Reasons for Action*; Anderson, "Value in Ethics and Economics," ch. 1; Gibbard, "Knowing What to Do, Seeing What to Do," 212; Raz, "Explaining Normativity"; Scanlon, *What We Owe to Each Other*, ch. 2; and Parfit, "Normativity."

25 Sellars, "Empiricism and the Philosophy of Mind," 169. Richard Rorty contrasts the logical space of reasons with the nonnormative logical space of causes (*Philosophy and the Mirror of Nature*, 157).

26 Joseph Raz, for instance, states that "the normativity of all that is normative consists in the way it is, or provides, or is otherwise related to reasons.... So ultimately the explanation of normativity is the explanation of what it is to be a reason, and of related puzzles about reasons" ("Explaining Normativity," 67). John Broome disagrees: "Reasons are undoubtedly important, but normativity has other important features, and our preoccupation with reasons distracts us from them" ("Reasons," 28). See also Broome, "Normative Requirements."

27 See note 63 in chapter 1 below for a brief explanation why.

write about normativity primarily in terms of *reasons*, both because that is what so many of the philosophers I shall be discussing do and because it captures my aim of discovering the foundations of practical reason. This will occasionally require me to reformulate someone else's arguments or theses in terms of reasons. I do not believe that doing so will obscure any important philosophical distinctions.

Note that readers who disagree with me about the primacy of reasons and 'oughts' need not cast the dissertation aside at this point. Even though they may reject my suggestion that the arguments and conclusions contained herein pertain to practical normativity or ethics as a whole, they can still accept them as relevant to an important and oft-discussed domain of ethics: the domain of reasons for action. I shall not be too disappointed if my theory turns out to be a theory "only" of the foundations of practical reason.

### Looking Ahead

Bernard Williams' papers on reasons for action were my introduction to the issues I have been discussing here. Both for that reason and because the crucial metanormative ideas at work in those papers have so seldom been fully recognized and appreciated, I devote chapter 1 to the debates surrounding Williams' view. I argue that buried within the seemingly endless exchanges about reasons and motives are two different approaches to the foundations of ethics and the problem of the regress of justification. The first—*internalism*—is a reductionist approach that regards justification as a function of explanation. For some consideration to be a reason for you to act, it must be the case that you could act for that reason (in which case the reason would explain your so acting). According to internalism, then, normative authority requires the capacity for normative governance. The second approach—*externalism*—rejects the putative connection between explanation and justification, or between authority and governance. According to externalists, normativity is primitive and *sui generis* and thus cannot be reduced to facts about explanation or motivation.

Both externalists and internalists accept that there are foundational norms and reasons, and both claim to have discovered them (or at least some of them). In other words, externalism and internalism both purport to be accounts of the foundations of ethics. In chapters 2 and 3 I evaluate these competing approaches in comparison to one another. Chapter 2 focuses on what I take to be the flaws of externalism. I conclude that the notion of external authority is

too obscure to serve as the centerpiece of an account of the foundations of ethics, and that a concept of authority entirely cut off from our motives and commitments cannot capture the essential *practicality* of reasons for action. I also observe that externalism's explanatory miserliness contrasts sharply with internalism's explanatory generosity. In chapter 3, however, I find the traditional internalist views wanting in other crucial respects. I argue that the virtues of internalism's reductive approach are outweighed by its inability to capture the *authority* of reasons for action. Internalism's emphasis on the practicality of reasons does not leave room for them to play their justificatory role.

Finally, in chapter 4, I present and defend my own approach to the foundations of ethics, which I call *constitutivism*. This approach takes off from David Velleman's claim that agency has a constitutive aim. Velleman himself has been cautious about deriving normative conclusions from his philosophy of action. I argue that this caution is unwarranted, and that we can derive a constitutive norm for action from action's constitutive aim. The theory of normativity that results is technically an internalist one, but one that has heeded the lessons of chapter 3 about the centrality of reasons' justificatory role. I conclude that my constitutivist approach captures both the practicality and the authority of normativity and therefore provides the most compelling solution to the problem of the regress of justification and the foundations of ethics.

The great moral philosopher (and externalist) H. A. Prichard once claimed that "in philosophy the truth can only be reached via the ruins of the false."[28] I agree. That is why I present and defend my own theory only after an extended critique of other views. The process whereby I arrived at my constitutivism was one of learning from the missteps and drawbacks of earlier attempts to unearth the foundations of ethics. Naturally, I regard my theory as a step in the right direction; it represents the approach that seems right to me at the moment. However, I do not presume to have put the issue to rest. There are undoubtedly objections yet to be considered, and some of them may prove decisive. If so, then I hope that the ruins of my own view will prove enticing to future explorers bent on discovering the foundations of ethics, and that my failures may lead them a little bit closer to the truth.

---

28    Prichard, "What is the Basis of Moral Obligation?" 1.

# CHAPTER 1

# THE DEBATE

> I am sure that I do not understand the idea of a reason for acting, and I wonder whether anyone else does either.
>
> —Philippa Foot, *Virtues and Vices*

Bernard Williams' 1980 paper "Internal and External Reasons" launched the contemporary debate about normativity and motivation. Of course Williams was not the first twentieth-century philosopher to propose an internal connection between reasons or obligations on one hand and motives on the other. Thirty years earlier, W. D. Falk had contrasted what he called the "motivation sense" or "internalist interpretation" of the word "ought" with the "externalist view" of that term.[1] In fact, Williams was not even the first to cast the issue in terms of reasons; Falk had broken that ground as well.[2] Nonetheless, it was Williams' arguments on behalf of this purported internal connection that placed the issue center stage, where it remains to this day.

Hardly a month goes by without the appearance of a new paper defending or criticizing some aspect of Williams' internalism. Unfortunately, this ongoing discussion has not sufficiently refined our understanding either of Williams' position or of the arguments he musters in its defense. It is not even clear whether the many philosophers who call themselves internalists and who take themselves to be supporting Williams' claims share a common view. Admittedly, Williams deserves at least some blame for this unhappy state of affairs. None of his four discussions of internalism is marked by a particularly clear presentation, and at least some

---

1   Falk, "'Ought' and Motivation," 33, 34, 40. Internalists before Falk had already posited such a connection, but he was the first to give it such explicit treatment.

2   See Falk, "Action-Guiding Reasons." Here Falk notes: "It would be odd to say that one gave someone a reason all right, but nothing that could possibly carry influence with him however well he considered it" (83).

of what he writes on the subject is actually misleading.[3] But more recent work on both sides of the issue has hardly helped matters. Nevertheless, philosophers continue to venture into this quagmire, perhaps with the conviction that hidden within are issues of great philosophical interest and importance.

I share this conviction. As I suggested in the introduction, I see in Williams' internalism hints or symptoms of a compelling but controversial account of normative authority. This account underlies many of the arguments and counterarguments surrounding internalism and externalism, yet it is seldom made explicit, let alone discussed on its own. My goal in this chapter is to correct that oversight. Starting with Williams' own statements on the subject, I shall isolate the different strands that make up his view. I shall then try to identify where in those strands the real normative and metanormative disagreements lie.

William Frankena praised Falk for raising the debate between internalists and externalists "to another level, a less merely logical and larger scale level"—for placing the issues "on a macroscopic rather than microscopic plane."[4] In revealing the competing accounts of normative authority that are driving so much of this debate, I hope to raise the discussion to an even higher, more macroscopic level.

## Williams' Internalism

In "Internal Reasons and the Obscurity of Blame," Williams states his view plainly: "there are only internal reasons for action."[5] According to Williams, this amounts to the claim that statements of the form "$A$ has a reason to $\phi$" (where $A$ is an agent and $\phi$ is an action) are true only if "$A$ could reach the conclusion that he should $\phi$ (or a conclusion to $\phi$) by a sound deliberative route from the motivations that he has in his actual motivational set—that is, the set of his desires, evaluations, attitudes, projects, and so on."[6]

---

3      The other three discussions can be found in "Internal Reasons and the Obscurity of Blame," "Replies," and "Postscript."

4      Frankena, "Obligation and Motivation in Recent Moral Philosophy," 71.

5      Williams, "Internal Reasons and the Obscurity of Blame," 35. In "Internal and External Reasons," Williams states his conclusion somewhat differently: "external reason statements, when definitely isolated as such, are false, or incoherent, or really something else misleadingly expressed" (111). This amounts, however, to the same claim.

6      Williams, "Internal Reasons and the Obscurity of Blame," 35.

Williams formulates his internalism in terms of a necessary condition for one to be in the state of having a reason for action.[7] This tells us nothing about the role that the actual reason must play, however. To see this difference, consider the following example. Suppose that my students' anxiety over their grades is a reason for me to mark their essays this evening. Williams' formulation can tell us what is required for me to have such a reason: in order for me to have a reason to mark my students' essays this evening, it must be the case that I could reach the decision to mark their essays by a sound deliberative route from the motivations I have in my actual motivational set. Notice that this statement of the necessary condition makes no mention of the reason in question—namely, my students' anxiety. This is an unfortunate omission. A properly specified necessary condition should require more than that I be able to arrive at the decision to mark essays via *some* sound deliberative route. If my students' anxiety is the consideration that is supposed to be my reason, then this very consideration had better be at least a stop along the deliberative route to my decision. And so, properly understood, Williams' view involves more than a necessary condition for $A$ to have a reason to $\phi$. It identifies a necessary condition for $R$ (where $R$ is some consideration) to be that reason! Reformulated in this way, internalism places the condition not on the agent's state of *having* a reason, but on the consideration's state of *being* a reason.[8]

---

7    As many philosophers have noted, this is only one of a plethora of internalisms in ethics. Stephen Darwall calls Williams' view an example of *existence internalism*, since it places a necessary condition on the existence of reasons for action. Darwall contrasts this sort of theory with *judgment internalism*, which involves a necessary condition for *judging* that one has a reason. I shall use the unmodified word "internalism" to refer to the former view. For more on Darwall's distinction, see Darwall, *Impartial Reason*, 54–5, as well as "Reasons, Motives, and the Demands of Morality." For an extensive but somewhat different taxonomy of the various internalisms in ethics, see Parfit, "Reasons and Motivation," 99–109.

8    Even though Williams never quite states his view this way, it is perfectly in keeping with the spirit of his internalism. For one thing, the revised version entails the original one. If some consideration $R$ satisfies the requirement of the reformulated view with respect to some agent $A$ and some action $\phi$, then $A$ will satisfy the requirement of the original view with respect to $\phi$. Put another way, if $A$ does *not* satisfy the requirement of the original view with respect to $\phi$—if, in other words, $A$ does not have a reason to $\phi$—then there will be no consideration $R$ that satisfies the requirements of the revised view with respect to $A$ and $\phi$. Moreover, in "Internal Reasons and the Obscurity of Blame," Williams defends his view by appealing to the following assumption: "if it is true that $A$ has a reason to $\phi$, then it must be possible that he should $\phi$ *for that reason*" (39, emphasis added). This assumption (to which I shall return below) clearly suggests my revised version of Williams' internalism, since it specifies that the sound deliberative route which must be open to the agent should include consideration of the reason in question. This is, after all, what $\phi$-ing *for that reason* involves.

Restating Williams' internalism in light of this point is somewhat tricky. A first stab might look something like this: $R$ is a reason for $A$ to $\phi$ only if $A$ could reach the conclusion that he should $\phi$ (or a conclusion to $\phi$) by a sound deliberative route from $R$ and the motivations that he has in his actual motivational set. The idea we want to capture is that it must be possible for *A's consideration of R* to lead from his motivational set—via sound deliberation—to his decision to $\phi$. In other words, the agent must be able to arrive at the conclusion to $\phi$ via a consideration of the reason in question.

Before taking another stab at restating Williams' view, we can stop for another small modification. A second disadvantage of Williams' formulation of internalism is that it amounts to a necessary condition only of decisive or all-things-considered reasons. It does not leave room for *pro tanto* reasons—that is, for reasons that might be outweighed by countervailing considerations. To see this, return to the above example about marking essays. According to Williams, my students' anxiety can be a reason for me to mark their essays only if my consideration of their anxiety could—by sound deliberation—lead me from my actual motivational set to a conclusion to mark their essays. As it happens I am a committed teacher: I care about my students, and I want to make their lives as anxiety-free as possible. Consequently, if I deliberate soundly on their anxiety about their grades, I will be motivated to spend the evening marking, at least to a certain extent. But suppose that despite my admirable commitment to my students, my overall motivational set is heavily skewed in some other direction. The Vienna Philharmonic is playing Mahler tonight at Hill Auditorium, and my thirst for Mahler is so great that no amount of sound deliberation on my students' troubles could lead me to conclude to spend the evening in my office grading essays. Mahler is what I care about the most, and thus sound deliberation will always yield a decision to go to the concert. Strictly speaking, then, given Williams' formulation of the necessary condition, the fact that my students are anxious *fails* to count as a reason for me to mark essays.

But this is not the conclusion Williams wants. After all, there *are* elements of my motivational set that are activated by my knowledge of my students' anxiety. This knowledge *does* motivate me. However, this motivation is overwhelmed or defeated by other, stronger motives. Williams wants to allow that an agent "can have reason to do a lot of things which he has other and stronger reasons not to do."[9] He states explicitly that "*A* has a reason to $\phi$'

---

9    Williams, "Internal and External Reasons," 104.

does not mean 'the action which *A* has overall, all-in, reason to do is *φ*-ing.'"[10] The problem is that Williams' formulation of internalism does not actually allow for this. He demands too much of reasons; he requires that they be able to have decisive or overriding influence even when they are not decisive or overriding reasons. Another reformulation of Williams' view is therefore required. *Pro tanto* reasons are defeasible, and so they need only be capable of generating defeasible motivation. Incorporating this insight into our reformulation, we arrive at something like: *R* is a reason for *A* to *φ* only if *A* could reach *defeasible motivation* to *φ* by a sound deliberative route from *R* and the motivations that he has in his actual motivational set. Alternatively: *R* is a reason for *A* to *φ* only if sound deliberation on *R* and his actual motivational set could defeasibly motivate *A* to *φ*.[11]

We now have a refined statement of Williams' internalism—that is, of his contention that there are only internal reasons for action. Most of the philosophers who have waded into the debates about this claim—including Williams himself—conclude that internalism entails subjectivism about reasons for action. But the view that there are no objective or universal reasons for action does not follow obviously from internalism, at least as it is stated above. To see how this subjectivism is supposed to arise, we need to examine the assumptions and arguments whereby Williams arrives at his view. It is to such an examination that I shall now turn.

**Williams' Argument**

Williams' theory of reasons is supported by two broad assumptions. The first can be found in Williams' claim that reasons must satisfy an explanatory requirement: "if there are reasons for

---

10    Williams, "Internal and External Reasons," 104. But see Williams, "Postscript," 91 for a contrary suggestion.

11    Compare Sobel, "Explanation, Internalism, and Reasons for Action," 220–2. There are further ambiguities in Williams' view. For instance, it is not immediately clear what is involved in deliberating "from a motive." In "The Real Problem with Internalism about Reasons," Talbot Brewer identifies two possible interpretations. According to the first, which Brewer calls the "inferential" interpretation, the recognition of the motive ("I want to *φ*") must play a role in the deliberation. According to the second, which he calls the "causal" interpretation, the existence of some motive must make it causally possible for *A* to deliberate soundly and arrive at the conclusion to *φ* (446–7.) Brewer claims that Williams must be an inferential internalist, since only that sort of internalism is supported by the assumptions underlying Williams' argument. For a similar, "inferential" interpretations of Williams, see Schueler, *Desire*, 72; and Pettit and Smith, "Backgrounding Desire," 278–81.

For reasons that should become clear presently, the inferential interpretation cannot be correct. See note 25 below, as well as Arkonovich, "Defending Desire," 510 n. 16.

action, it must be that people sometimes act for those reasons, and if they do, their reasons must figure in some correct explanation of their action."[12] Put more schematically: *R* is a reason for *A* to *φ* only if *A* could *φ* for reason *R*—that is, only if *R* could be *A*'s reason for *φ*-ing. Call this the *explanatory constraint*.[13]

What does this constraint really require? According to Williams, only something that motivates an agent can be his reason for acting and thus explain his action. And motivation, for Williams, is a causal process.[14] However, action for a reason requires more than action merely caused by a reason.[15] *R* causing *A* to *φ* is not sufficient for *R* to be *A*'s reason for *φ*-ing. *R* must cause or influence *A* to *φ* *in the right sort of way*. The problem here is that almost *any* consideration can causally influence an agent to perform almost any action. For example, suppose some nefarious psychologists condition me to ask for change whenever I am near a soda machine. If this conditioning program is successful, then whenever I see a soda machine I

---

12    Williams, "Internal and External Reasons," 102. In "Internal Reasons and the Obscurity of Blame," Williams cites this as one of the "fundamental motivations" of his internalism (38).

13    See Dancy, *Practical Reality*, 101. David Copp refers to this as "the explanation principle," and David Sobel calls it "the explanation condition." See Copp, "Against Internalism about Reasons," 457; and Sobel, "Explanation, Internalism, and Reasons for Action," 218.

14    See Williams, "Internal and External Reasons," where he claims that "nothing can explain an agent's (intentional) actions except something that motivates him so to act" (107). Since Williams thinks (as almost all philosophers do) that motivation is causal, here he commits himself to a casual theory of action for a reason (and hence of action explanation). One could certainly accept Williams' assumption about reasons and explanation (namely, that any normative reason for *A* to *φ* must be capable of explaining *A*'s *φ*-ing) but reject his causal account of action. I have worked these two claims together into a single assumption—the explanatory constraint—for the simple reason that few of Williams' critics challenge or reject his causal theory of action for a reason. The philosopher who comes closest is Jonathan Dancy in *Practical Reality*. Dancy embraces the explanatory constraint, and he even accepts the idea that action explanations appeal to motives (or to what he calls *motivating reasons*). However, on Dancy's view, motivational explanations are not causal explanations. He defends a noncausal theory of motivation! In *Reason and Value*, E. J. Bond proposes a similar view of action explanations that appeal to motivation (9–26).

    Interestingly, in "Are External Reasons Impossible?" Rachel Cohon misleadingly suggests that she rejects Williams' claim that only something that motivates can explain an agent's actions (549). However, her suggestion is based on a misreading of that claim. In reality, Cohon's objection is to Williams' broadly Humean theory of motivation, which I discuss below. In "Williams' Argument against External Reasons," Elijah Millgram insists that this claim—that only what motivates an agent can explain his actions—does not follow from the idea that reasons must be able to contribute to explanations of the actions for which they are reasons. Millgram is quite right on this point, but he is mistaken to think that Williams' internalism depends in any way on such an entailment.

    Even if there are few defenders of noncausal theories of action in the internalism debate, there are plenty to be found in and around the philosophy of action literature. See, for instance, Wilson, *The Intentionality of Human Action*, and Schueler, *Reasons and Purposes*.

15    This is one respect in which Williams' view is Davidsonian in spirit.

will ask for change without even thinking about it.[16] But when this happens, the consideration that I am in the vicinity of a soda machine does not become my *reason* for seeking change, even though it is causing my action. It is not my reason because I am not acting for a reason at all: the consideration that I am near a soda machine is not influencing me *as a reason*. The difference here is between merely causal influence and *rational* influence. If Williams' view required only that *R* be able to *cause A* to *ϕ* in order for *R* to be a reason for *A* to *ϕ*, then his internalism would rule out almost nothing.[17]

This is where Williams' emphasis on "sound deliberation" enters the picture. If *R* is to exert rational rather than merely causal influence on *A*, then *A* must be open or susceptible to that kind of influence. In other words, *A* must be an agent—someone in the market for reasons for action. The phrase "sound deliberative route" is simply Williams' way of describing the pathways whereby agents are influenced by reasons: a sound deliberative route is whatever path causal influence must travel in order to count as rational influence—as influence *as a reason*.[18] And to say that someone is deliberating soundly is just to say that he is susceptible to that influence.

Williams offers a fairly expansive account of what constitutes sound deliberation. He allows for far more than simple causal means-end reasoning. Sound deliberation can include reasoning about different ways of attaining or subserving one's ends, thinking about how the satisfaction of different ends can be combined or ordered, finding a specific form for an as yet unspecified project, seeking constitutive solutions to one's projects, and considering how much weight to attach to one's various commitments. Sound deliberation can also involve exercising

---

16  If it is not credible that even clever psychologists could accomplish this sort of conditioning, imagine some nefarious neurosurgeons instead. Compare Sobel, "Explanation, Internalism, and Reasons for Action," 222.

17  For a misreading of Williams' argument along these lines, see Cohon, "Internalism about Reasons for Action," 269–71.

18  Actually, as Williams at one point observes, even causal influence *through sound deliberation* is not *sufficient* for rational influence: "In particular, we need to exclude boring counter-examples in which the motivation to *ϕ* is merely a causal consequence of the activity of deliberating" ("Replies," 220 n. 4). Here Williams is referring to the problem of deviant causation—a general problem for causal theories of action. "Boring" counterexamples of causal deviance can be found throughout the philosophy of action literature. See, for instance, Davidson's case of the nervous climber in "Freedom to Act," 79; Goldman's story of the befouled soup in *A Theory of Human Action*, 54; and Chisholm's account of the greedy nephew in "The Descriptive Element in the Concept of Action," 616.

  For an extended and wonderfully detailed response to the problem of deviant causation, see Bishop, *Natural Agency*.

one's imagination to think about alternative courses of action and outcomes. Although Williams insists that there are many different forms of sound or rational deliberation, he is not interested in providing a fully determinate account.[19] He holds that it is not always clear what our reasons are, and so he sees it as an advantage of his view that it preserves this indeterminacy.

Recall that according to the explanatory constraint, reasons for action must be able to explain the actions for which they are reasons. But a reason can only explain an agent's action if the agent can act for that reason, and a consideration counts as an agent's reason for acting only if that consideration rationally influences the agent to act. This therefore amounts to the claim that reasons must be able to rationally influence the agents for whom they are reasons: $R$ is a reason for $A$ to $\phi$ only if $R$ could rationally influence $A$ to $\phi$.

This statement of the explanatory constraint is too strong, though, since it seems to imply that we cannot have reasons of which we are unaware. If I am not aware of $R$, then $R$ cannot rationally influence me, and thus—according to the constraint as I have just stated it—$R$ cannot be a reason for me. But Williams clearly rejects this conclusion. He believes that considerations of which we are unaware *can* be reasons for us. To borrow Williams' example, the fact that the stuff in my glass is petrol rather than gin is certainly a reason for me not to drink it, even if I do not know it is petrol.[20] The explanatory constraint is therefore best stated as follows: $R$ is a reason for $A$ to $\phi$ only if $R$ could rationally influence $A$ to $\phi$ were $A$ aware of $R$.

The second assumption underlying Williams' internalism is a broadly Humean theory of motivation. This theory places an upper limit on what sound or rational deliberation can accomplish. According to Williams, a consideration can never rationally influence an agent unless that consideration is appropriately related to the agent's actual motivational set. Specifically, $R$ could rationally influence $A$ to $\phi$ only by showing $A$ that $\phi$-ing would, in some way, subserve one of his ends, goals, or projects.[21] What makes this theory of motivation

19    For Williams' comments on sound deliberation, see "Internal and External Reasons," 104, and "Internal Reasons and the Obscurity of Blame," 38.

20    Williams, "Internal and External Reasons," 102.

21    In Christopher Cowley's words, a consideration can rationally influence an agent to act only if it can "find purchase" in or "speak to" an element of the agent's actual motivational set. See Cowley, "A New Defence of Williams's Reasons-Internalism," 349.

broadly *Humean* is its emphasis on the agent's *actual* motivational set, or on motives the agent *already has*.[22] It is *A*'s *current* motives that determine what can rationally influence *A* in the near future.[23] Motivation, on this view, cannot rationally arise in a vacuum; no amount of rational or sound deliberation can create new motivation out of whole cloth. Instead, rational motivation only ever issues from prior motivation. Put another way, action for a reason requires motivational fuel. Without it, rational deliberation can never get off the ground.[24]

What makes this theory of motivation only *broadly* Humean is its relatively expansive account of what counts as a motive (or as an element of the agent's actual motivational set). In addition to desires, an agent's motivational set includes "dispositions of evaluation, patterns of emotional reaction, personal loyalties, and various projects, as they may abstractly be called, embodying commitments of the agent."[25] Moreover, this list is not meant to be exhaustive. The class of motives should be broadly construed to include all of the agent's conative or motivational states or dispositions.[26]

22    As several commentators have noted, it is not clear whether Hume qualifies as a Humean. (See, for instance, Millgram, "Was Hume and Humean?") Hobbes, however, was certainly a Humean when it comes to motivation. See Darwall, *The British Moralists and the Internal 'Ought'*, 60.

23    As we have seen, strictly speaking *A* can become motivated to do almost anything, regardless of what his current motives are; we need only dispatch more of those nefarious psychologists or neurosurgeons.
           It is important not to overstate the limits placed on motivation by a broadly Humean theory. As Peter Railton has reminded me, a Humean can allow that our underlying motivational makeup might evolve over time in nondeliberative ways, perhaps due to physical or mental maturation, the influence of our peers, and so forth. And this opens the door to future rational influence that is unrelated to our *current* motives. What the Humean cannot allow is the possibility that *deliberation* can bring about such nondeliberative motivation—except, of course, through "deviant" channels (see note 18 above). Compare McDowell, "Might There Be External Reasons?" 97–108; and Williams, "Replies," 186–94.

24    In *Moral Realism*, Russ Shafer-Landau puts the point in a slightly misleading way. According to Shafer-Landau, Williams "announces the impossibility of a certain kind of motivation," which Shafer-Landau calls "motivation *ex nihilo*" (175). But Williams' position is not that motivation *ex nihilo* is impossible. It is that such motivation cannot be brought about through rational deliberation. Shafer-Landau acknowledges this point in a footnote (175 n. 12).

25    Williams, "Internal and External Reasons," 105. Williams sometimes seems to believe that prior motives must play a role not only in rational motivation but also in rational *deliberation*. This would make the broadly Humean theory implausible. Even when motivational fuel is powering our deliberations, it need not do so in the foreground. That is, we need not be actively deliberating on our own goals, inclinations, or projects. Our attention may instead be fixed entirely on the *objects* of our desires and plans, with the desires themselves relegated to the background. This does not make them any less necessary for rational motivation, though. For a nice account of this distinction, see Pettit and Smith, "Backgrounding Desire."

26    This way of stating the Humean theory of motivation nicely sidesteps the unproductive debate about whether and to what extent desires are required for motivation. For the highlights of that debate, see Nagel,

Let us now consider the two assumptions of Williams' argument together. According to the explanatory constraint, a reason must be able to rationally influence the agent for whom it is a reason. But according to the broadly Humean theory of motivation, whether a consideration can rationally influence an agent is determined by that agent's actual motivational set—that is, by his desires, goals, commitments, projects, values, and so forth. A consideration rationally influences an agent by engaging one of his motives. Does this add up to subjectivism about reasons? At least one sort of subjectivism seems unavoidable. From Williams' two assumptions it does indeed follow that all reasons for action are, at least to a certain extent, a function (although not necessarily a simple or straightforward one) of the (potential) actions and reactions of agents. And so we can confidently conclude, borrowing a useful technical term from Peter Railton, that on Williams' view reasons for action are all "subject-ive"—that is, they are all "essentially connected with the existence or experience of subjects."[27] There are plenty of philosophers who would recoil at even this mild subject-ivism. (We shall meet several in the next chapter.) But we cannot pause to soothe their anxiety here, for a far more serious and perhaps pernicious sort of subjectivism threatens. If an agent's reasons for action in a given situation are ultimately a function of his motives, then agents in similar circumstances but with different motives may have different reasons. And since motives are notoriously different from one agent to the next, Williams' view seems to entail that there are no reasons shared by everyone. What one has most reason to do in a given situation need not be what anyone else has most reason to do in that same situation.[28] An agent's reasons are ultimately contingent upon his ends; and thus reason is, to this extent, "the slave of the passions."[29]

*The Possibility of Altruism*, 29–32; Darwall, *Impartial Reason*, 39–40; Smith, *The Moral Problem*, 92–129; and especially Wallace, "How to Argue about Practical Reason."

27    Railton, "Subject-ive and Objective," 55. Compare McDowell, "Aesthetic Value, Objectivity, and the Fabric of the World," 113–4.

28    Compare Velleman, "The Possibility of Practical Reason," 170–1; Wallace, "How to Argue about Practical Reason," 15–16; and Millgram, "Williams' Argument against External Reasons," 198. Millgram suggests that Williams makes two distinct arguments in "Internal and External Reasons." According to Millgram, the first argument is about only motivation, whereas the second incorporates the idea of sound or rational deliberation. As I hope is clear from my discussion of Williams' papers, I think that his claims about motivation and rational deliberation are best seen as two stages of a single argument.

29    Hume, *A Treatise of Human Nature*, book II, part iii, section iii, 415.

## Williams' Subjectivism

To see how Williams' view is supposed to yield this pernicious subjectivism, consider his example of Owen Wingrave (from the short story by Henry James). Owen comes from a family with a strong tradition of military service and honor. However, he "has no motivation to join the army at all, and all his desires lead in another direction: he hates everything about military life and what it means."[30] Owen's family urges him to enlist nonetheless. They insist that their tradition of military service is a reason for Owen to join the army. Williams assesses the family's claim as follows. According to the explanatory constraint, were their tradition a reason for Owen to enlist, then it would be possible for this consideration to rationally influence him to enlist. That is, it would be possible for sound deliberation on his family's tradition to generate at least defeasible motivation for Owen to join the army. But, according to the broadly Humean theory of motivation, rational deliberation could motivate Owen to join the army only if Owen already had at least some motive that would be subserved by that course of action. And Williams specifies that this is precisely what Owen lacks: "*All* his desires lead in another direction."[31] Therefore, Williams concludes, Owen's family must be wrong: considerations of family tradition are not a reason for Owen to enlist. The family's claim to the contrary turns out to be an external reason statement; and according to Williams, all such statements are false.

It is worth pointing out here that Williams' subjectivism is not as radical a position as some philosophers have thought. Owen fails to have *any* reason to join the military only on the implausibly strong assumption that *none* of the elements in his subjective motivational set—none of his desires, values, commitments, concerns, or projects—would be subserved in any way by enlisting. The stipulation here is not merely that Owen does not care about honor and family traditions as such. It is, rather, that Owen does not care about *anything* that might be promoted or achieved by joining the army—that, for instance, he is not at all interested in pleasing his family, attaining higher social standing, earning money, or impressing the ladies with his uniform. And that is a very strong assumption. If, for example, he were committed to pleasing his family, then their desire for Owen to enlist would be something that could rationally influence Owen. In this case, the sound deliberative route would run via Owen's

---

30    Williams, "Internal and External Reasons," 106.

31    Williams, "Internal and External Reasons," 106, emphasis added.

interest in pleasing his family. Thus, the seemingly harsh conclusion that Owen has no reason at all to join the military follows only on this implausibly strong assumption about Owen's motives.

It is reasonable to wonder how often we will be able to say with any certainty that an agent has *no* motives that could be subserved by a particular course of action. Williams makes a similar point about someone who is not at all motivated to pursue his own needs:

> If an agent really is uninterested in pursing what he needs; … and he could not reach any such motive from motives he has by the kind of deliberative processes we have discussed; then I think we do have to say that in the internal sense he indeed has no reason to pursue those things. In saying this, however, we have to bear in mind how strong these assumptions are, and how seldom we are likely to think that we know them to be true. When we say that a person has reason to take medicine which he needs, although he consistently and persuasively denies any interest in preserving his health, we may well still be speaking in the internal sense, with the thought that really at some level he *must* want to be well.[32]

Williams is not backing away from his subjectivism here. He is merely pointing out how rarely an agent will have absolutely no reason at all to pursue what may seem to us to be a worthwhile course of action.

Nonetheless, subjectivism does indeed seem to follow from the explanatory constraint and Williams' broadly Humean theory of motivation. After all, it is not easy to find considerations that motivate everyone. And if reasons must bear an internal connection to motives, then reasons will be as subjective as the motives on which they depend. But there is a third assumption driving this inference—an assumption I did not mention above and without which subjectivism (at least of the pernicious variety) does not actually follow. This is revealed in Williams' talk of a *subjective* motivational set. Here Williams simply presupposes that motivation is always subjective. But if there are motives that all agents necessarily share, then internalism will allow for reasons that all agents necessarily share. Such reasons would still be subject-ive, since they would still be essentially connected with the existence or experience of subjects. They would not, however, be subjective in the stronger, more serious sense.

Williams acknowledges as much in his discussion of Kant, whom he accepts as "a limiting case of internalism."[33] If, for example, concern for universal law as such is somehow built into

---

32    Williams, "Internal and External Reasons," 106.

33    Williams, "Replies," 220 n. 3.

the nature of rational agency, then considerations that trigger this concern will be reasons for all rational agents. And so, even if internalism entails the weaker subject-ivism, one can accept Williams' internalism while consistently denying the more pernicious subjectivism about reasons. It is only internalism combined with subjectivism about motives that entails subjectivism about reasons.[34]

But if internalism by itself does not entail subjectivism, wherein lies its philosophical interest? If even dyed-in-the-wool Kantians can qualify as internalists, does internalism have anything interesting to say about reasons for action? At least one half of Williams' position does seem empty of metanormative content. On its own, Williams' broadly Humean theory of motivation says nothing at all about reasons. It belongs more to the philosophy of action than to ethics or the philosophy of normativity. The explanatory constraint is another matter, however. Recall that according to the explanatory constraint, $R$ is a reason for $A$ to $\phi$ only if $A$ could $\phi$ for reason $R$—that is, only if $R$ could be $A$'s reason for $\phi$-ing. The basic idea behind this constraint is that the explanatory role of reasons is somehow crucial to their normativity: considerations that could not explain $A$'s $\phi$-ing cannot be (normative) reasons for A to $\phi$. Here we find the real metanormative interest of internalism: a purported connection between normativity and explanation, or—in the practical case—between normativity and motivation. Thus, insofar as it is a normative or metanormative theory about reasons for action, internalism *just is* the explanatory constraint.[35] If there is an element of Williams' position that deserves to be center stage in discussions of reasons and normativity, it is the idea that we must have the capacity to act for the normative reasons we have.

Many philosophers (myself among them) have found the idea of a connection between normativity and explanation extremely compelling. Some even consider it obvious that such a connection exists. But the nature of this supposed connection remains deeply mysterious. After all, it is not immediately clear what counting in favor of—what justifying or recommending— has to do with explaining. While it may seem clear to internalists that reasons must meet the explanatory constraint, they have traditionally done a poor job explaining why. In fact, most

---

34    For more on the ways in which internalism can be combined with different theories of motivation and rational deliberation to yield different theories of reasons, see FitzPatrick, "Reasons, Value, and Particular Agents," 285–91.

35    Shafer-Landau is therefore correct to insist that the explanatory requirement cannot be used to defend internalism. It *is* internalism! See Shafer-Landau, *Moral Realism*, 179.

consider it too obvious even to bother attempting such an explanation. Through the rest of this chapter and much of the next, my goal will be to remedy this oversight. But first we must take a brief detour.

## Clarifications

Most philosophical debates are muddled in various ways, but the one surrounding internalism and the explanatory constraint seems more turbid than most. How did it get this way? Part of the problem stems from Williams' failure to state clearly whether his internalism is supposed to be a normative theory regarding what reasons we have or a metanormative theory about the very nature of reasons for action. It could be either. To see this, consider the explanatory constraint again: $R$ is a reason for $A$ to $\phi$ only if $R$ could rationally influence $A$ to $\phi$ were $A$ aware of $R$. One way to interpret this constraint is simply as a first-order normative claim about which considerations are *not* reasons for action. This sort of interpretation sees internalism as a substantive ethical theory, or theory of practical reasons, on a par with, say, egoism, consequentialism, or any of the several theories Derek Parfit discusses in *Reasons and Persons*.[36] Consider the following trio:

> *Egoism:* $R$ is a reason for $A$ to $\phi$ if and only if $R$ indicates that $\phi$ would contribute to $A$'s well-being.

> *Consequentialism:* $R$ is a reason for $A$ to $\phi$ if and only if $R$ indicates that—compared to the alternatives open to $A$—$\phi$ would yield the best consequences.

> *Internalism:* $R$ is a reason for $A$ to $\phi$ only if $R$ could rationally influence $A$ to $\phi$ were $A$ aware of $R$.

These are competing views of what we have reason to do. Each purports to be the expression of some basic ethical intuition about what sorts of reasons we have. The intuition captured by internalism is that I cannot have a reason to $\phi$ unless I have motivational dispositions that could be recruited on behalf of $\phi$-ing. Of course unlike egoism and consequentialism, internalism (at least as Williams delineates it) involves only a necessary condition, but we can

---

36    See especially pts. 1 and 2. Connie Rosati treats internalism as a class of substantive theories about what is good for a person. See Rosati, "Internalism and the Good for a Person."

nevertheless regard it as a stop on the path to a more complete view, perhaps along the lines of Parfit's "deliberative theory" or some other full information account of practical reason.[37]

Understood in this way, internalism is a ground-floor normative view. One could be an internalist in this sense without thinking that the explanatory constraint has anything to tell us about the nature of normativity and reasons for action. Often, however, Williams appears to be interested in tunneling beneath the ground floor and down to the foundations. His objection to externalism is not simply that there are no external reasons, but also that the very idea of an external reason borders on incoherence.[38] Conclusions like this make Williams' internalism seem less like a normative theory about what matters and more like a metanormative theory about what mattering consists in. On this second, metanormative interpretation, the explanatory constraint represents Williams' view of what reasons we have only because it also—and more fundamentally—represents his view of *what it is to be a reason*.

As a purely exegetical matter, the second of these two interpretations must be the correct the one. This is made clear by Williams' repeated insistence that external reasons statements are not merely false but also incoherent, as well as by his suggestion that he does not understand what normative force could be if it is not just some type of motivational force. Despite these clear indications, though, an unfortunately large number of philosophers have failed to recognize Williams' metanormative ambitions, and they have assumed that he is making a merely normative claim. Consequently, it is now commonplace to treat Williams' view as though it were simply the ground-floor normative theory I mentioned above.[39] Hence the muddled nature of the literature.

---

37  For Parfit's deliberative theory, see *Reason and Persons*, 117–8. Williams explicitly endorses the view that the explanatory constraint is both a necessary and a sufficient condition, although he never tries to defend the sufficiency claim. See "Internal Reasons and the Obscurity of Blame," 35.

38  See Williams, "Internal and External Reasons," 109–11; and "Internal Reasons and the Obscurity of Blame," 39–40.

39  See, for instance, the debate in *Philosophical Quarterly* about "internalism" and the conditional fallacy, sparked by Robert Johnson in "Internal Reasons and the Conditional Fallacy." Responses to Johnson include Brady, "How to Understand Internalism"; van Roojen, "Motivational Internalism"; and Gert, "Avoiding the Conditional Fallacy." See also Johnson, "Internal Reasons: Reply to Brady, van Roojen, and Gert."
    A remarkably similar exchange was initiated by David Sobel in "Explanation, Internalism, and Reasons for Action." For criticism of Sobel's argument, see Robertson, "Internalism, (Super)fragile Reasons, and the Conditional Fallacy." For Sobel's reply, see "Reply to Robertson." See also Enoch, "Why Idealize?"

I do not mean to suggest that this purely normative version of internalism is somehow inconsistent or incoherent, although I do regard it as unfortunate that the label "internalism" has come to refer to this view. The philosophers most responsible for introducing the label—Falk, Frankena, and Williams—clearly treated internalism as a metanormative account of the nature of morality, or the nature of normativity in general. Moreover, internalism as a merely normative theory seems radically unmotivated. It is not at all clear why one would think that considerations must satisfy a "motivational recruitment requirement" in order to count as normative reasons unless one also thinks that having motivational or explanatory potential is at least part of what it is to be normative.[40] Finally, it is only the metanormative version of internalism that promises to address the problem that interests me here, namely, the regress of justification and the foundations of ethics. For these reasons, I shall hereafter use the term "internalism" to refer only to those views that advance the explanatory constraint as at least a partial account of what being a reason consists in.[41]

Even with this distinction under our belts, however, further clarification is required. Another factor contributing to the muddled nature of the internalism literature is the fact that many philosophers fail to recognize that the explanatory constraint is separable from the broadly Humean theory of motivation. As a consequence, such philosophers often describe themselves as externalists even though they fully accept the connection between normativity and explanation at the heart of internalism.[42] But even among those who recognize that the explanatory constraint and the Humean theory of motivation are discrete views, there is disagreement over how to classify theorists who accept the former but not the latter. Williams himself tends to call them externalists. Throughout most of his work on internalism, Williams treats externalism as the view that reasons can motivate a rational agent and hence satisfy the

---

40    The phrase "motivational recruitment requirement" was suggested to me by Peter Railton. Throughout this section I am drawing extensively on his helpful comments.

       Here it is interesting to compare Prichard's claim that if we believe that to be right an action must be advantageous "we cannot stop short of maintaining that it is precisely advantageousness and nothing else which renders an action right" ("Duty and Interest," 27–9).

41    One unfortunate result of this usage is that I shall end up applying the label "externalist" to philosophers who are, in various other senses, internalists. This has the potential to create confusion, which I shall endeavor to avoid by flagging such cases in footnotes.

42    For some examples of critics who fail to separate Williams' internalism from his Humeanism, see Cohon, "Internalism about Reasons for Action" 268; and Gert, "Skepticism about Practical Reasons Internalism," 63–6.

explanatory constraint without finding purchase in motives the agent already has.[43] It never seems to have occurred to him that the most serious externalists might simply reject the explanatory constraint outright.

Interestingly, though, nearly all of Williams' critics—even those who are quite hostile to his internalism—claim to *accept* the explanatory constraint. This is partly due to Christine Korsgaard's formulation of internalism. Korsgaard defends what she calls the *internalism requirement*: reasons for action "must be a capable of motivating rational persons."[44] This has been treated by many philosophers as a canonical statement of internalism. Notice, however, that the term "rational" is used here to modify "persons." On Korsgaard's view, it is not merely the *influence* reasons exert that must be rational in nature; it is also the *agents* themselves. This makes her version of the explanatory constraint seem weaker than the one Williams proposes. In Korsgaard's hands, it becomes as much a constraint on who counts as rational as it is a constraint on what counts as a reason.[45] Internalism, she claims, does not require that reasons *always* be capable of motivating us. "All it requires is that rational considerations succeed in motivating us *insofar as we are rational*."[46]

Some such proviso is clearly necessary. Without it, the explanatory constraint would state that $R$ is a reason for $A$ to $\phi$ only if $R$ could influence $A$ to $\phi$ were $A$ aware of $R$. But this bare "could" is underspecified: it indicates neither what sort of influence reasons must be able to have nor the circumstances in which they must be able to have it. This leaves us with a constraint that is either much too weak or much too strong. On the one hand, the underspecified constraint seems to rule out next to nothing, for—as we saw above—pretty much any consideration is capable of exerting motivational influence, at least with the help of those nefarious psychologists. To strengthen the requirement and thereby disqualify considerations that can move us only under such extreme circumstance, we need to clothe that bare "could." But we must be careful not to overdress it. That is, we must not over-strengthen the explanatory constraint. We need to leave room for the possibility that when a consideration is incapable of influencing me, it is my fault rather than the consideration's. Otherwise we lose

---

43    This is the view Williams' criticizes in "Internal and External Reasons," 106–11.

44    Korsgaard, "Skepticism about Practical Reason," 317,

45    Compare Velleman, "The Possibility of Practical Reason," 172.

46    Korsgaard, "Skepticism about Practical Reason," 321, emphasis added.

the normativity of reasons. As Korsgaard puts it, "the fact that people are motivated differently by the reasons they have does not show that they have different reasons. It may show that some have virtues that others lack."[47] The failure of some consideration to influence me cannot by itself establish that the consideration is *incapable* of influencing me (and hence not a reason for me). If it did, then I would never fail to be influenced by my reasons, since a consideration's lack of influence would actually be a sign that the consideration is not a reason.

It may therefore appear as though Korsgaard has identified a crucial element missing from Williams' internalism: a rationality proviso for the explanatory constraint. Despite this appearance, however, Williams' version of the explanatory constraint is no less specified than the one Korsgaard advances. Recall that on Williams' view reasons must be able to exert *rational* influence, and that for a consideration to rationally influence an agent, the agent must first be susceptible to that sort of influence. This is why Williams requires that reasons be able to influence agents *only insofar as those agents are deliberating soundly*. Like Korsgaard, Williams leaves room—in situations where a consideration fails to motivate an agent—for it to be the agent's fault, rather than the consideration's. These are the cases in which the agent is *not* deliberating soundly. In such cases, the agent closes off certain avenues for the conveyance of rational influence and thereby makes himself immune to that influence. And so Williams' version of the explanatory constraint, rendered in terms of rational influence, also includes a rationality proviso and thus allows for the possibility of irrational action.

### The Rationality Proviso

In her first paper on internalism, Korsgaard observes that the strength of the explanatory constraint—and in particular whether it entails subjectivism about reasons—depends at least in part on the strength of the attached rationality proviso. What considerations count as internal reasons will depend on what is required for an agent to be rational. And this latter question, she maintains, must be resolved on its own. Moreover, such resolution will involve "a more metaphysical argument about just what reason does, what its scope is, and what sorts of operation, procedure, and judgment are rational."[48]

---

47    Korsgaard, "Skepticism about Practical Reason," 325.

48    Korsgaard, "The Possibility of Practical Reason," 323.

A number of philosophers have read this paper as providing a recipe for nonsubjective internalism: start with the explanatory requirement; add a robust account of rationality; and *voilà!* Most of the philosophers who follow this recipe pay only lip service to the explanatory constraint, however; and thus they are not really internalists at all. This is most evident when rationality is simply defined in terms of reasons.[49] The view that rationality is just a matter of responding appropriately to one's normative reasons, or that the rational thing to do is just what one ought to do, makes the explanatory constraint trivially true. Reasons will always have the capacity to influence rational agents if rational agents are just defined as ones that are influenced by their reasons. On this understanding of rationality, the explanatory constraint is actually no constraint at all. It is an empty formalism that tells us nothing at all about the nature of reasons or even about what considerations can be reasons. After all, on the view in question, in order to know whether a particular consideration can influence a rational agent, we first need to know whether that consideration is a reason. We need, in other words, an independently specified account of reasons for action. But this is precisely what internalism is supposed to provide. It should therefore be clear that defining rationality in terms of reasons drains the explanatory constraint (and hence internalism) of any content. Philosophers who adopt this approach are externalists, irrespective of whether they claim to accept the explanatory requirement.[50]

Of course, rationality need not be defined in this way. Some philosophers who accept the explanatory constraint adopt instead a substantive account of rationality according to which agents are rational only insofar as they possess certain motives (and lack others) or pursue certain courses of action (and avoid others).[51] David Velleman provides a sketch of how such a view might look:

> Being a rational agent entails having various motives, including a preference for one's own greater good and an acceptance of moral principles. A rational agent is

---

49  See, for instance, Parfit, "Reasons and Motivation," 109–20. Parfit does not claim to be an internalist, but he does suggest that he can accept the explanatory constraint, given his understanding of rationality.

50  Stephen Darwall makes a similar point when he notes that this approach to the explanatory constraint makes *reasons* the "independent variable" and *rationality* "the dependent variable" ("Reasons, Motives, and the Demands of Morality," 310). Darwall suggests that only views that make rationality the independent variable are really consistent with the spirit of internalism. See also Darwall, *The Second-Person Standpoint*, 292. Compare Korsgaard, "The Normativity of Instrumental Reason," 243.

51  See, for instance, Parfit, "Reasons and Motivation," 112–7; and Gert, *Morality*, 29–68.

influenced by a reason for doing something when, for example, he considers some respect in which doing it is morally required; and this consideration can influence him because an inclination to abide by moral requirements is partly constitutive of his rationality. If an agent lacks this inclination, its absence won't prevent him from having moral reasons for acting: moral requirements will still count as reasons for him to act. Rather, lacking an inclination to abide by moral requirements will render the agent irrational, by making him insensitive to this particular kind of reason.[52]

On the surface at least, this sort of approach seems to qualify as an instance of internalism. It appears to be consistent with the explanatory constraint's condition that reasons be capable of exerting rational influence. On this view, those who are not moved by moral considerations are simply not susceptible to the influence of moral reasons, since they lack the inclination to abide by moral requirements. They are, in other words, irrational, whereas the explanatory constraint requires only that reasons be able to motivate *rational* agents!

The literature is divided over whether philosophers who incorporate a substantive account of rationality into the explanatory constraint should count as internalists or externalists.[53] Such philosophers accept the explanatory constraint without rendering it trivial, and this may suggest that they should qualify as internalists. I shall nonetheless classify them as externalists. As Williams points out, it is the substantive account of rationality that is pulling all of the weight for these views. An explanatory constraint that incorporates such an account of rationality can allow *any* consideration *R* to be a reason, as long as rationality is delimited so that rational agents always respond to *R*-like considerations. If, for example, we simply define rationality such that one must be motivated by moral considerations in order to count as a rational agent, then moral considerations will always qualify as internal reasons. But the notion of rationality at work in Williams' internalism is not substantive in this way. Therefore, even though this sort of view accepts a nontrivial version of the explanatory constraint, it is more in the spirit of externalism than internalism.

Further clarification is required, though. A number of philosophers who claim to be internalists incorporate a *procedural* account of rationality into their formulations of the explanatory constraint. Stephen Darwall is one. In *Impartial Reason*, Darwall defends the

---

52    Velleman, "The Possibility of Practical Reason," 173. Of course, Velleman does not defend such a view. Compare Wallace, "How to Argue about Practical Reason," 17–18.

53    Korsgaard defends a substantive view in "Skepticism about Practical Reasons," where she describes herself as an internalist. Velleman, on the other hand, describes her view in that paper as an example of externalism. But see note 65 below.

seemingly internalist view that *R* is a reason for *A* to $\phi$ if and only if *R* could rationally motivate *A* to $\phi$ "under conditions of *rational consideration*."[54] But crucially, for Darwall rational deliberation must be dispassionate and impartial. Consequently, considerations that fail to move "passionate and partial" agents do not thereby fail to count as genuine reasons for action. It is the agents who fall short: they fail to count as rational. Since there are nonsubjective considerations that will always influence impartial and dispassionate agents, Darwall's internalism allows such considerations to qualify as reasons.[55] Michael Smith's take on the explanatory constraint incorporates a similarly procedural notion of rationality. According to Smith, internalism amounts to the claim that whether *A* has a reason to $\phi$ depends on whether *A* would be motivated to $\phi$ were he fully rational. However, "fully rational agents *converge* in their desires about what is to be done in the various circumstances they might face."[56] For Smith, this is part of what we mean when we describe an agent as rational.

Smith and Darwall both claim to be internalists, and their proceduralist approach is certainly closer to the spirit of internalism. Nevertheless, I shall not treat them as internalists. Even their more modest, procedural notion of rationality is not really akin to the one at work in Williams' view. Recall that the motivation for the explanatory constraint is the idea that we must be able to act for our reasons: if *R* is a reason for me to $\phi$, it must be the case that *R* could be my reason for $\phi$-ing. And so strictly speaking the explanatory constraint requires agents to be rational only in the sense that they must be susceptible to rational influence in general. They must, in other words, be in the market for reasons for action, and rationality is the price of admission. This is just the basic or minimal sort of rationality that is constitutive of agency.[57] Darwall and Smith add to this minimal notion a more robust, procedural view of what makes for sound deliberation. While this might add up to a perfectly coherent theory of reasons, it stretches the explanatory constraint beyond its humble, internalist origins.

---

54    Darwall, *Impartial Reason*, 81, emphasis in original.

55    Actually, since Darwall's internalism involves a necessary *and sufficient* condition for reasons, considerations that move impartial and dispassionate agents *always* count as reasons for action.

56    Smith, "Internal Reasons," 33–4.

57    Minimal rationality is akin to the ante that must be paid before agents can participate in (or be proper subjects of) discussions of what one has reason to do. This is what Korsgaard calls "the descriptive sense" of rationality ("Internalism and the Sources of Normativity," 62). The absence of this sort of rationality is not irrationality, but rather nonrationality.

Here it may seem that I am unreasonably restrictive in my use of the term "internalism." After all, Darwall and Smith are a far cry from the obviously externalist theorists who work substantive ends into the nature of rational agency or instead simply define rationality in terms of reasons. Darwall describes his view as one according to which "the normativity of reasons ... arises from an ideal of rational consideration and not through the externalist idea that there are simply some facts that have the intrinsic property of being action-guiding and some that do not."[58] At one point Darwall suggests that it is precisely the emphasis on a procedural or formal ideal of reasoning that marks internalism as a distinctive view. But this cannot be correct, especially given his own account of what internalism is supposed to be. According to Darwall, metanormative internalism seeks "an understanding of the unqualifiedly normative."[59] It hopes to capture "the nature of normativity itself."[60] But Darwall explains the notion of an unqualifiedly normative reason in terms of another robustly normative notion, namely, rationality. He describes rationality as an independent system of *norms*, and he takes his theory of rationality to establish an *ideal* of practical deliberation.[61] These are straightforwardly normative terms, as Darwall acknowledges. Describing his own view, he states that "the normative or justificatory weight of reasons, then, is held to derive from a *normative* ideal of rational consideration."[62] Instead of relying on certain facts that have the intrinsic property of being action-guiding (as most externalists do), Darwall appeals to standards of rationality that have the intrinsic property of being normatively deliberation-guiding. The correctness of dispassionate and impartial reasoning is, therefore, just as external as the standards of correctness postulated by the externalist view he rejects. The same goes for Smith's theory of rationality. Although both he and Darwall account for reasons in terms of an independently specified and procedural notion of rationality, they ultimately leave the "the nature of normativity itself" unexplained.[63]

---

58    Darwall, *Impartial Reason*, 81.

59    Darwall, "Internalism and Agency," 157.

60    Darwall, *The British Moralists and the Internal 'Ought'*, 10.

61    See Darwall, *Impartial Reason*, 201–11.

62    Darwall, "Reply to Terzis," 115, emphasis added.

63    Jean Hampton also takes philosophers like Darwall and Smith to be externalists. She makes the further claim that a normative ideal of rational deliberation entails that there are *reasons*—reasons of rationality, as it were—to "deliberate in certain ways" (*The Authority of Reason*, 78). If, for example, rational deliberation

This is most apparent when we consider their views in light of the problem of the foundations of ethics. If internalism is going to put a stop to the regress of justification, it must do more than explain the normativity of reasons in terms of a normative ideal of rationality. After all, any such ideal itself calls for justification. Confronted with Darwall's conclusion that ideal practical reasoning is impartial, we might reasonably ask why *that* is the model we should adopt as a guide for our own deliberations. Darwall proposes a system of norms for practical reasoning, but there are competing systems as well. Why should we deliberate according to one rather than the other? What reason do we have to accept norms that prescribe impartial deliberation over those that prescribe a more egocentric approach? The explanatory constraint is defended by internalists as at least a partial answer to such questions. Internalism as I understand it is the view that a philosophical account of the foundations of ethics must ultimately involve an appeal to facts about what can explain or guide our actions and deliberations. In *Impartial Reason*, however, Darwall acknowledges that there might be agents who cannot be rationally influenced to deliberate impartially. Given their current motives and deliberative practices, no amount of deliberation could guide such agents to adopt the impartial standpoint. But this in no way undercuts the normativity of the deliberative ideal, according to Darwall. He insists that the norms of rationality are "normative for *all* agents with rational

---

is always impartial (as Darwall suggests), then all agents have an external reason to deliberate impartially, according to Hampton. Such reason are external because "they direct the agent whether she likes it or not, and have a compelling rightness about them, whether she is willing to recognize that fact or not" (79–80).

In his review of Hampton's book, Darwall rejects the idea that a normative deliberative ideal commits one to external *reasons* for so deliberating. According to Darwall, "what a deliberative ideal prescribes is *how* to deliberate, which way of deliberating is ideal or correct. It does not tell (or give agents reasons) *to* deliberate" (review of *The Authority of Reason*, 584). John Broome has more recently made a similar point. Broome urges that reasons do not follow from rational requirements. From the fact that I am rationally required to take the means to my ends, it does not follow that I have a reason to take those means, since I might instead have a reason to change my ends. (See Broome, "Normative Requirements.")

I am inclined to agree with Hampton here, since I think that the "reasons of rationality" that follow from a deliberative ideal can be couched in conditional terms. I would argue that from the fact that rational requirements do not yield detachable reasons, it does not follow that they do not yield any reasons at all. There may be conditional reasons that are not detachable, and this might simply be the form that reasons of rationality take. (Here I side with Roger Crisp against Broome. See Crisp, *Reasons and the Good*, 56–61.) But even if Darwall and Broome are correct, my more basic point stands. A deliberative ideal is fully normative, and it can guide our practical deliberations in precisely the same way that reasons guide our actions. Darwall sometimes seems not to recognize this point. In his attempt to distinguish rational requirements (which prescribe *how* to deliberate) from reasons (which dictate *whether* to deliberate), he contends that "whether deliberation makes sense in any given situation is a practical matter" (review of *The Authority of Reason*, 584). But surely how to deliberate is also a practical matter! Any proposal for how to deliberate stands in just as much need of justification as a proposal regarding whether to deliberate at all.

capacity."[64] Darwall's normative ideal of deliberation therefore fails to meet the explanatory constraint, and this is why I classify him and philosophers who adopt a similar approach (such as Smith and Korsgaard, at least in her first paper on the subject) as externalists.[65]

Of course at this point we might reasonably wonder: if these three philosophers are not internalists, then who is? What must a theory of reasons or normativity look like in order to count as internalist? According to Williams, it must state not merely that certain norms are ideal, but that these norms "are necessarily part of rational agency."[66] In work published since the appearance of *Impartial Reason*, Darwall has moved in precisely this direction. He acknowledges that his earlier view leaves crucial questions of justification unanswered.

> If facts inherit their status as reasons, their justificatory weight, from their capacity to motivate on rational reflection, how does a sort of reflective consideration itself acquire normative status as rational? What does its normativity rest on?[67]

In response to these pressing questions, Darwall suggests that the normativity of his deliberative ideal is rooted in the nature of autonomy—that is, in the thought that "rational conduct is autonomous conduct."[68] This approach, inspired by Kant, revolves around the idea that impartial reflection is essential to genuinely autonomous action. Agents who are guided by deliberation that is not impartial are thereby less than fully autonomous. Now, if we assume that autonomous action is simply action for a reason, we can see how this further claim is more in the spirit of internalism and the explanatory constraint. On Darwall's revised view,

---

64    Darwall, "Impartial Reason," 116.

65    In more recent work, Korsgaard disavows her earlier definition of internalism. In "The Normativity of Instrumental Reason," she states the explanatory constraint in terms of what can motivate us insofar as "we are susceptible to the influence of reason" (226). And in "Internalism and the Sources of Normativity," Korsgaard rejects the suggestion that internalism involves a "presupposition of rationality in the normative sense" (63). Instead the rationality proviso is more a thesis in "moral psychology" or the philosophy of action (63). She also explicitly acknowledges that her formulation of the internalism requirement in "Skepticism and Practical Reason" is misleading: "the way I wrote that paper makes it possible to confuse my position with the one I just rejected—that rational principles are somehow just out there and you are rational if you respond to them correctly" (64).

66    Williams, "Postscript," 92. Here "rational agency"—agency that involves a general capacity to respond to reasons—must be distinguished from the fully normative notion of rationality *in* agency—that is, the notion of an agent who is also (procedurally or substantively) rational. Thanks to David Velleman for discussion of this point.

67    Darwall, "Impartial Reason," 117.

68    Darwall, "Impartial Reason," 118.

the normativity of rationality is grounded in claims about what is required for agents to act for reasons, and this is precisely the idea at the center of Williams' internalism.

In this section I have tried both to clarify and to justify my narrow use of the label "internalism." The clarification is more important for my larger project. Even if I have not shown that only philosophers who are internalists in my narrow sense really deserve to be called (metanormative) internalists, I have at least identified a particular strain of internalism, a strain that addresses the foundations of ethics by grounding the normativity of reasons in facts about agency and action for a reason. This is the sort of internalism that interests me, and it will be my focus in what follows.

### Behind the Explanatory Constraint

As I mentioned above, for many internalists the connection between normativity and explanation—or between reasons for action and acting for a reason—is readily apparent. But for most externalists, it is equally apparent that there is no such connection. This naturally raises the question of what can be said in defense of the explanatory constraint. Why think that reasons must be able to explain? Normative reasons are justifying reasons, and at least on the surface it is not clear why justification must involve the capacity for explanation or influence. Given that justification and explanation seem to be two entirely different things, why do internalists insist that they are so closely linked?[69] As I mentioned above, even Williams' critics often accept this assumption; Jonathan Dancy actually treats it as a "maxim."[70] But does it deserve this status?

Surprisingly few philosophers have discussed these questions, and even fewer have done so at any length. This is due at least in part to the aforementioned muddle that is the internalism literature, and in particular to the confusion surrounding the various understandings of the explanatory constraint. Having identified and clarified the version of internalism relevant to the foundations of ethics, we can now safely turn to an investigation of the explanatory constraint's mysterious philosophical allure.

---

69    Sobel points out that many philosophers appeal to the explanatory requirement to defend internalism. But, as Sobel observes, properly understood, the explanatory requirement just is internalism! Shafer-Landau notices that internalists appeal to the explanatory requirement, but after showing that it is the same as internalism, he never stops to ask why someone would find it appealing (178–80).

70    Dancy, "Why There Is Really No Such Thing as the Theory of Motivation," 3–4.

One consideration that might account for this allure is the fact that we use the term "reason" in both normative and explanatory contexts. Consider the following statements:

Owen's family's tradition of military honor is a reason for him to enlist in the army.

Owen's reason for enlisting in the army is his devotion to his family's tradition of military honor.

The first statement is normative. It asserts that Owen's family's tradition justifies or counts in favor of enlistment. The second statement is nonnormative. It appeals to Owen's family's tradition only in order to explain Owen's enlistment. Such an explanation is perfectly consistent with *any* claim about the normative status of this family tradition. Family tradition may have been Owen's reason even though we all agree it was not a *good* reason. Now, internalists might argue, since the notion of a reason is appropriately employed in both normative and explanatory contexts, then surely something that is a reason in the first sense must at least be capable of being a reason in the second sense. In other words, there must be a close connection between justification and explanation.[71] To deny such a connection would be to allow for the possibility of reasons that cannot be acted upon. But, as E. J. Bond wonders near the beginning of *Reason and Value*, "how can one speak of reasons that cannot be acted on?"[72] Is it not a platitude about reasons they are the sorts of things we can act upon or for?

This argument hardly amounts to an adequate defense of the explanatory constraint. It presupposes that there is only one sort of reason at work in both normative and explanatory contexts. It assumes, in other words, that the term "reason" is univocal. If "reason" is univocal, then something along the lines of the explanatory constraint may indeed follow. If the term "reason" picks out the same sort of thing in both of the above statements, then anything that counts as a reason in normative contexts had better be able to play a role in explanatory contexts as well. But whether "reason" is univocal—whether there is only one sort of reason—is precisely what is at issue here. Philosophers who reject the explanatory constraint often appeal to a distinction between two different sorts of reasons: motivating reasons (which contribute

71    For arguments at least roughly along these lines, see Williams, "Internal and External Reasons," 102–3, Williams, "Internal Reasons and the Obscurity of Blame," 38–9, and Dancy, *Practical Reality*, 2-3. In "Why There Is No Such Thing as the Theory of Motivation," Dancy presents a similar argument designed to show that motivating reasons are not psychological states.

72    Bond, *Reason and Value*, 7.

to explanations of actions) and normative reasons (which justify actions).[73] Moreover, highlighting possible connections between these two sorts of reasons—for instance, that acting for a motivating reason involves treating something as a normative reason—does not show that there is only one sort of reason. Thus, pointing merely to the fact that the term "reason" is employed in both normative and explanatory contexts does not demonstrate that there is an important connection between justification and explanation.[74]

My point is not that we should blindly accept the distinction between motivating and normative reasons. It is rather that internalists cannot defend the explanatory constraint by assuming this distinction away. Merely wondering how there can be reasons that cannot be acted upon does not demonstrate that there are no such reasons. And if the idea of such a reason sounds suspicious, that may only be because we are not attending to the distinction between motivating and normative reasons. A motivating reason we cannot act upon is indeed a bizarre notion, for it is a motivating reason that cannot motivate! It is not at all clear what (if anything) is bizarre about the notion of a normative reason that cannot motivate or explain. The "maxim" that we must be able to act for the reasons we have is therefore far too nebulous to provide any real support to internalism.

A second platitude to which internalists sometimes appeal for support is more significant. This is the idea that reasons for action are *practical*. Here is the *Stanford Encyclopedia of Philosophy* on the appeal of internalism:

> Behind this internalist position lies the idea that practical reason is practical in its issue. Internalists contend that we can make sense of the generation of new intentions through reasoning only if we assume that such reasoning is conditioned by motivational resources that are already to hand.[75]

73    Peter Railton has reminded me that there are actually other distinctions I am eliding here. We might, for example, distinguish between reasons that explain in virtue of being the agent's conscious rationale and reasons that explain in virtue of more unconscious motives (such as greed or envy). A third sort of explanatory reason might have nothing to do with the agent's psychological states. We could, for example, cite someone's bad upbringing as the reason for his rude behavior. The differences among the various sorts of explanatory reasons will not be relevant here, though.

74    Compare Baier, *The Moral Point of View*, 149, 153–4; Bond, *Reason and Value*, 9; Parfit, "Reasons and Motivation," 112–4; and Brink, *Moral Realism and the Foundations of Ethics*, 39–40, 62–3. Contrast Williams, "Postscript," 93.

75    Wallace, "Practical Reason."

The suggestion here is that reasons can be genuinely practical only if they have the capacity to influence our behavior. But again this is exactly what is at issue, and it certainly does not follow merely from the platitude that reasons for action are practical. Externalists agree that reasons for action are practical. Such reasons are practical in their subject matter: they are practical (rather than theoretical) since they are reasons *for action* (rather than for belief).[76]

Internalists reply that the practicality of reasons involves more than a reference to actions. Darwall defends the internalist idea that reasons must be capable of motivating by claiming that

> practical reasoning is practical not just in its subject matter but in its issue. Its question is what *to do*, and its product is, at the least, a judgment about what to do, and more characteristically, an intention or decision to act.[77]

But this further claim cannot settle the matter, since externalists have their own stories to tell about the ways in which practical reasoning yields motivation and thereby influences our actions. They can, for example, appeal to a standing desire to do what we have reason to do. Note that I do not mean to side with the externalists here. My point is merely that the disagreement about the way in which reasons are practical reflects competing conceptions of practical reasoning and agency: internalism is not the only game in town. Until the internalism conception is shown to be superior to the externalist alternatives, the claim that reasons for action are practical will support neither position over the other.

I shall discuss and evaluate these competing conceptions of agency in chapters 2 and 3. But first I want to close this chapter by mentioning a third platitude about reasons for action to which internalists often appeal for support. Reasons, they insist, are *guides*, and a consideration can guide *A* only if that consideration can influence *A* to act. Darwall adopts this approach to defend his version of the explanatory constraint:

> Reasons to act are considerations that guide rational action. It is a necessary condition of something's being a reason, then, that it can play this guiding role; it must be possible, that is, for a person to act for that reason, for the consideration to be a *person's reason* for acting."[78]

---

76    Compare Frankena, "Obligation and Motivation in Recent Moral Philosophy," 50; and Scanlon, *What We Owe to Each Other*, 2.

77    Darwall, *Impartial Reason*, 51.

78    Darwall, *Impartial Reason*, 35.

Again, though, it is hard to see how the platitude that reasons are guides can contribute to a defense of the explanatory constraint. On one understanding of the notion of guidance, $R$ guides $A$ to $\phi$ only if $R$ influences or motivates $A$ to $\phi$. If this is what we mean when we assert that reasons are guides, then something like the explanatory constraint does indeed seem to follow. But in reality we have merely pushed the question back a level. For, as Darwall himself notes, externalists simply deny "that something can be such a guide only if it is capable of motivating the agent."[79] In other words, even if we grant a connection between guidance and the justificatory nature of reasons for action, we are left with questions about the nature of the connection between guidance and explanation (or motivation). To assume that reasons guide in a motivational sense is simply to assume the explanatory constraint. But we are looking for a defense of that constraint, not a restatement of it. What is it about the nature of the sort of guidance offered by reasons that requires a capacity to influence (and hence to explain)?

I think we can move closer to a proper understanding of the explanatory constraint's allure by focusing less on the fact that reasons guide our behavior and more on *the way* in which they guide our behavior. A crucial fact about normative reasons is that they not only guide or direct us but that they do so *with authority*. As we saw in the introduction, any old norm generates reasons that present themselves as guides. The norms of anti-Semitism direct us to hate Jews in much the same way that the norms of morality guide us to respect other persons. The difference between the former and latter, though, is that the latter are supposed to possess an unqualified authority that the former lack. The norms of morality guide or direct with authority; their special normativity just is this authority.

But how does the fact that reasons are authoritative bring us closer to an understanding of the explanatory constraint? I would like to suggest that one of the basic assumptions underlying the explanatory constraint is the idea that authority cannot exist in the absence of governance. Authority and governance always go hand in hand: part of what makes something properly authoritative is its ability to govern your behavior. An authority that cannot govern you has no genuine authority over you at all. This idea goes at least some way towards explaining the attraction of the explanatory constraint. Since normative reasons must be able to govern your actions, they must also be able to explain those actions. This is why internalists maintain that Owen's family is wrong about whether Owen has a reason to join the army.

79    Darwall, *Impartial Reason*, 52.

Considerations of military honor cannot govern his actions, and so they have no authority over him. They are not genuine reasons for him to act.

Of course, merely recasting the explanatory constraint in terms of authority and governance cannot settle the debate between internalists and externalists. It can, however, reveal more clearly what is at issue in that debate. Behind all of the talk about justification and explanation, guidance and practicality, reasons and motives, is the idea that the authority of reasons depends upon their capacity to govern us.

Describing the disagreement in these terms allows us to reconceive the internalism-externalism dialectic. Internalists believe that authority requires governance. As we have seen, this is the metanormative view that being a genuinely authoritative reason consists in part in the capacity to govern, at least under certain conditions. If the specification of these conditions does not involve the appeal to a system of norms that itself calls for justification, then internalism can form the basis of a reductive theory of practical normativity.[80] This is precisely the strategy internalists have pursued over the years. One internalist tradition, stretching back to Richard Cumberland's *A Treatise of the Laws Nature* (1672) and perhaps a bit further to Thomas Hobbes' *Leviathan* (1651), takes normative force simply to be the inescapable practical force exerted by certain motives.[81] More recent members of this tradition include Williams, as well as full-information theorists like Richard Brandt and Peter Railton. A second tradition, with its roots in the work of Joseph Butler and Immanuel Kant, locates the source of normativity in the nature of autonomy and identifies the force of reasons with the force of motives resulting from autonomous practical reasoning.[82] Contemporary followers of this tradition include latter-day Kantians such as Darwall and Korsgaard (in their more recent work). I shall discuss these varieties of internalism in more detail in chapter

---

80    My understanding of internalism leaves little room for the possibility of a nonreductive internalism. There are certainly philosophers who call themselves internalists and who reject reductivism, but I do not count them as internalists, since they either defend internalism in its normative (rather than metanormative) guise or else account for the normativity of reasons in terms of some other, fully normative notion. Contrast Dancy, *Practical Reality*, 18–9.

81    I say "perhaps" to Hobbes, since there is disagreement about whether Hobbes' was an internalist. Darwall has actually defended both interpretations. For his internalist interpretation of Hobbes, see "Internalism and Agency," 162–3; and especially *The British Moralists and the Internal 'Ought'*, ch. 3. For his more recent, projectivist take, see "Normativity and Projection in Hobbes's *Leviathan*."

82    Darwall finds traces of this tradition as far back as Ralph Cudworth and John Locke. See *The British Moralists and the Internal 'Ought'*, chs. 5–6.

3. The important point here is that both traditions understand normative authority in terms of normative governance. Moreover, in these reductive forms internalism promises to put a halt to the problematic regress of justification. If, for example, we can account for reasons for action in terms of a certain inescapable motivational force, then once we have identified the considerations that exert that force, we shall have found unqualifiedly normative reasons. The normativity of these reasons will not depend for justification on any other norms or reasons. Instead it will simply consist in these reasons' capacity to govern us. In other words, there will be no cause to wonder whether the capacity to govern in the specified way really does lend authority to considerations, since—according to internalism—this is just what authority is.

Externalists, on the other hand, reject the idea that an internalist reduction of normativity can halt the regress of justification. They believe that the authority of reasons must ultimately be independent of those reasons' capacity to influence or govern us, and so they regard internalism as a *threat* to the foundations of ethics. Authority, they argue, is not the same thing as efficacy. An investigation into what considerations can influence us is largely irrelevant to questions about the regress of justification, since justification is a matter not of what *can* influence us, but rather of what *should* influence us. There is a difference, these philosophers insist, between what we are moved or influenced to do and what we *ought* to do. The authority of reasons must be able to stand on its own, and internalist attempts to yoke authority to governance thereby sap reasons of any genuine normativity.

Historically, the most prominent externalists have been the ones who speak out against the perceived threat to the foundations of ethics posed by internalism. These externalists include early modern moralists Samuel Clarke and Richard Price (responding to the internalism of Hobbes and Hutcheson) as well as contemporary realists Derek Parfit and T.M. Scanlon (reacting to the more recent internalism of Williams and Korsgaard).

As I understand them, then, internalism and externalism are competing approaches to the foundations of ethics and to the regress of practical justification, and they carry with them competing understandings of the nature of agency and practical reasoning. Internalists hold that, at bottom, the foundations of ethics must involve a connection to the motivational and psychological systems that govern our behavior. Externalists deny this.

Anyone coming to this debate via the many papers about Williams' view would be hard pressed to recognize the dispute as I have just described it. I have tried to show that this is

largely because so many critics (and sympathizers, for that matter) have misunderstood the central assumption underlying Williams' arguments. A close examination of Williams' defense of the explanatory constraint reveals his commitment to a conception of normative authority and agency according to which authority and governance go hand in hand. Although these important ideas have too often been lost in the muddle of the Williams literature, traces of them can be found elsewhere. In the next two chapters, I shall look to these traces as I attempt to assess the relative merits of the internalist and externalist conceptions of normativity.

# CHAPTER 2

## EXTERNALISM

> Everything is what it is, and not another thing, unless it is another thing, and even then it is what it is.
>
> —William K. Frankena, "The Naturalistic Fallacy"

As we saw at the end chapter 1, internalism—at least of the sort that interests me—can provide the basis for a reductive account of normative authority. According to internalists, normative authority is partly constituted by the capacity to govern: an agent's capacity to be moved by a consideration is part of what it is for that consideration to be a reason for him to act. By itself this is not a *complete* account of normative authority, since it delineates only a necessary condition. It does, however, point in the general direction of such an account. In particular, it points toward reductive theories in which the capacity to govern plays a crucial role.

A theory along these lines would, if successful, provide the resources for a solution to the problem of the regress of justification. If normative authority is nothing more than the capacity to exert a certain sort of motivational influence, or the capacity to exert motivational influence under certain specified conditions, then once we identify which reasons actually have that capacity, we will also have identified the reasons that are genuinely authoritative or unqualifiedly normative. A complete internalist theory of normative authority would therefore explain both where the regress of practical justification comes to a stop and why it comes to a stop there rather than elsewhere. It would, in other words, establish the foundations of ethics.

But not everyone interested in the foundations of ethics is an internalist. There is a long tradition of externalist responses to the problematic regress of justification. Philosophers who adopt the externalist approach reject the proposed connection between normative authority and normative governance. They insist that authority can be genuinely authoritative (and hence

unqualifiedly normative) even if it is ultimately independent of our capacity to recognize and respond to it.

Externalist accounts of the foundations of ethics are the focus of this chapter, and internalist accounts the focus of the next. Of course the two approaches cannot be discussed in perfect isolation, since general arguments against one are *ipso facto* arguments in favor of the other. Here my emphasis will be on the flaws of externalism (and hence implicitly on the virtues of internalism). The shortcomings of internalism (along with the saving graces of externalism) are on the agenda for chapter 3.

Adjudicating the dispute between these two approaches is no straightforward matter. Each view offers a conception of normativity with its own theoretical advantages and disadvantages. But choosing a theory requires more than simply tallying the pluses and minuses for each position and then determining which one comes out ahead. In the debates between internalists and externalists an argument from one side or the other is often met with the refrain, "What you say may be true, but then it is not normativity you are talking about!" This raises the rather depressing possibility that internalists and externalists are operating not with different *conceptions* of normativity, but rather with different *concepts*. And if that is the case, our problem of adjudication is difficult indeed, for how are we to choose between two different concepts?

A clash of competing concepts requires us to determine which concept does the most justice to our perhaps inchoate everyday thinking about practical justification and reasons for action. As this chapter and the one that follows will reveal, both sides have something important to say about normativity; however, neither view—at least as it is usually formulated—tells the whole story.

## The Genealogy of Externalism

Externalism has quite an impressive pedigree. Most of today's externalists can trace the lineage of their view through the intuitionists of the early twentieth century and back at least to the eighteenth-century British moralists Samuel Clarke, John Balguy, and Richard Price. In their writings we find the first flowerings of the externalist approach to the foundations of ethics, although of course never under that name.

Consider first the views of Clarke. In his second set of Boyle Lectures, informatively titled *A Discourse Concerning the Unchangeable Obligations of Natural Religion and the Truth and Certainty of Christian Revelation* (1705), Clarke sets out to secure the foundations of morality against the threat posed by Thomas Hobbes' "false reasonings."[1] Gainsaying an "assertion so very absurd, that Mr Hobbes himself could hardly vent it without blushing," Clarke confidently declares that there are "eternal and necessary differences" and relations that give rise to a "fitness or unfitness of certain manners of behavior."[2] For Clarke, practical normativity—the fitness or reasonableness of actions—is entirely a function of these "eternal relations," relations that hold independently of any agent's capacity to be influenced by them and that are antecedent even to "the positive will or command of God."[3] Some things are simply "in their own nature good and reasonable and fit to be done."[4] Thus, that you have a reason to help those in need stems not from any motives or inclinations you have, but rather from "real differences originally, necessarily, and absolutely in the nature of things."[5] Clarke represents these claims as his "attempt to prove and establish the eternal difference of good and evil"; after declaring success in this endeavor, he proceeds to "deduce the *original obligations* of *morality*."[6]

Balguy, a near-contemporary of Clarke, expounds a similar view in *The Foundations of Moral Goodness* (1728–9). He rejects Francis Hutcheson's sentimentalist theory of "the true ground and foundation" of ethics, and in its place he offers an externalist account according to which the origin of moral obligation is "independent of all *instincts*, necessarily fixed, and immovably rooted in the nature of things."[7] Thirty years later, in *A Review of the Principal Questions in Morals* (1758, 3rd ed. in 1787), Price continues the argument against sentimentalism, urging that normative terms "denote what actions *are*."[8] Moreover, on Price's

---

1    Clarke, *A Discourse Concerning the Unchangeable Obligations of Natural Religion*, as reprinted in Raphael, *British Moralists*, § 229. Citations are to section numbers in this edition.

2    Clarke, *Discourse Concerning the Unchangeable Obligations of Natural Religion*, §§ 229, 225.

3    Clarke, *Discourse Concerning the Unchangeable Obligations of Natural Religion*, § 226.

4    Clarke, *Discourse Concerning the Unchangeable Obligations of Natural Religion*, § 228.

5    Clarke, *Discourse Concerning the Unchangeable Obligations of Natural Religion*, § 227.

6    Clarke, *Discourse Concerning the Unchangeable Obligations of Natural Religion*, §§ 227, 251.

7    Balguy, *The Foundations of Moral Goodness*, as reprinted in Raphael, *British Moralists*, §§ 437, 446. Citations are to section numbers in this edition.

8    Price, *A Review of the Principal Questions in Morals*, as reprinted in Raphael, *British Moralists*, § 684, emphasis in original. Citations are to section numbers in this edition.

view, "whatever any thing *is*, that it is, not by will, or decree, or power, but by *nature and necessity*."[9]

Clarke, Balguy, and Price present their externalist accounts of obligation as responses to perceived challenges to the authority of morality. Each sees externalism as crucial to the preservation of a genuinely normative morality. This commitment to externalism is sometimes obscured, however, by the fact that these philosophers are also *internalists*, albeit of a peculiar variety. They all believe that *knowledge* or *awareness* of normative relations is necessarily motivating—in other words, that one could not recognize the rightness or fitness of a particular course of action without feeling moved to act accordingly. (Of course, this necessary motivation can still be overwhelmed by stronger, countervailing motives.) Thus Clarke insists that with knowledge of the eternal relations of fitness, "the *wills* likewise of all intelligent beings are constantly directed, and must needs be determined to act accordingly."[10] Balguy concurs, arguing that "the same necessity which compels men to *assent* to what is *true*, forces them to approve what is *right* and fit," and that this approval is "the *reason exciting* a man to the *choice* of virtuous action."[11] And Price claims that "when we are conscious that an action is *fit* to be done, or that it *ought* to be done, it is not conceivable that we can remain *uninfluenced*, or want a *motive* to action."[12]

Stephen Darwall calls this type of internalism *perceptual internalism* and—at least in recent work—considers it to be a subcategory of *existence internalism*, which has been my focus here.[13] Since I have a rather narrower conception of existence internalism than Darwall, I classify the rational intuitionists as externalists. Let me pause briefly to explain why.

---

9   Price, *A Review of the Principal Questions in Morals*, § 684, emphasis in original. A few lines later, Price concludes that normative properties "must immutably and necessarily belong to those actions of which they are *truly* affirmed" (§ 684, emphasis in original).

10  Clarke, *Discourse Concerning the Unchangeable Obligations of Natural Religion*, § 230, emphasis in original.

11  Balguy, *The Foundations of Moral Goodness*, § 453, emphasis in original.

12  Price, *Review of the Principal Questions in Morals*, § 757, emphasis in original. Plato, of course, held a similar view about knowledge of the Good. This has led many critics to label the view that combines existence externalism with perceptual internalism *Platonism*. See, for instance, Mackie, *Ethics*, 23–4; Smith, "Internalism's Wheel," 334–6; and Korsgaard, "Normativity, Necessity, and the Synthetic A Priori," 2.

13  See Darwall, "Ethical Intuitionism and the Motivation Problem," 248–70, "Reasons, Motives, and the Demands of Morality," 307–310, and "Internalism and Agency," 157–8.

According to Clarke, Balguy, and Price, any essential connection to motivation has nothing to do with the *existence* of obligations or normative reasons. Consider the case of Price. In his statement of his perceptual internalism, he explicitly denies any metaphysical connection between obligation and motivation. He rejects as "defective" the idea that the former somehow requires the latter: motivation is "the *effect* of obligation perceived, rather than *obligation itself*."[14] Price is also careful to distinguish knowledge of one's duty from duty itself. "A man's consciousness that an action ought to be done … cannot, properly speaking, be *obligation itself*," and on Price's view "excitement" or "approval" follows only from the former.[15] For all three of these early modern intuitionists, the connection between obligation and the will plays no part in what actually *makes* some actions obligatory and others not. The existence of obligation is independent of any capacity to motivate agents, and so I hesitate to classify Clarke, Balguy, and Price as existence internalists.[16]

If anything, the internalism at issue here seems closer to *judgment internalism* than to existence internalism.[17] Consider another of Price's comments: "Excitement belongs to the very *ideas* of moral right and wrong, and is essentially inseparable from the apprehension of them."[18] Nowadays "apprehension" is a factive term: a statement of the form "*A* apprehends that *p*" implies that *p* is true. This would suggest that Price is a perceptual internalist rather than a judgment internalist. But there is evidence that Price uses the term "apprehension" nonfactively. Earlier in the *Review* he asks rhetorically: "How strange would it be to maintain … that the apprehension of all beings, on this subject, are alike just, since all sensation must be alike *true* sensation."[19] Here Price expresses bewilderment at the suggestion that all beings would apprehend matters of right and wrong in the same way. But this would be puzzling only if "apprehend" were a nonfactive term like "belief," in which case Price's view that the

---

14    Price, *Review of the Principal Questions in Morals*, § 717.

15    Price, *Review of the Principal Questions in Morals*, § 719. A few lines later, Price writes: "It is not exactly the same to say, it is our *duty* to do a thing; and to say, we *approve* of doing it. The one is the quality of the action, the other the *discernment* of that quality" (§ 719).

16    Interestingly, this was also Darwall's conclusion in some of his earlier work. See *The British Moralists and the Internal 'Ought'*, 325–8.

17    For a reminder of the differences between these two internalisms, see note 7 in chapter 1 above.

18    Price, *Review of the Principal Questions in Morals*, § 757.

19    Price, *Review of the Principal Questions in Morals*, § 682.

apprehension of right and wrong necessarily involves motivation seems like an instance of judgment internalism.

It is crucial to bear in mind here that Price, like nearly all of his contemporaries, is mired in a Cartesian and Lockean "veil of perception" epistemology, according to which mental entities—ideas or perceptions—are the immediate objects of perception.[20] Thus, perceiving or viewing rightness in an action involves perception only of the *idea* of rightness. With this in mind, a number of Price's statements suggest judgment internalism more than perceptual internalism. For instance:

> An affection or inclination to rectitude cannot be separated from the view of it.[21]

> The determination of our minds concerning the nature of actions as morally good or bad, suggests a motive to do or avoid them.[22]

We must resist the temptation to read terms such as "view" and "determine" factively. For Price (as well as for most early modern philosophers), one can *perceive*, *be aware*, *view*, *determine*, and perhaps even *see* that an action is morally right without that action actually being right.[23]

Balguy's defense of his view about rightness and motivation likewise suggests judgment internalism more than perceptual internalism. Consider the comment from Balguy quoted above: "the same necessity which compels men to *assent* to what is *true*, forces them to approve

---

20  Locke's commitment to the "veil of perception" remains a matter of some controversy among Locke scholars. That said, the "standard reading"—which holds that Locke, like Descartes before him, is an indirect or representational realist—still predominates. This interpretation dates back to Berkeley; its classic twentieth-century statement (as well as the phrase "the veil of perception") can be found in Bennett, *Locke, Berkeley, Hume*. For a more recent discussion of this issue, see the following contributions to a symposium on "Locke and the Veil and Perception": Rogers, "Locke and the Objects of Perception"; Yaffe, "Locke on the Ideas of Substance and the Veil of Perception"; Newman, "Locke on Sensitive Knowledge and the Veil of Perception"; Bolton, "Locke on the Semantic and Epistemic Role of Simple Ideas of Sensation"; Lennon, "Through a Glass Darkly"; and Chappell, "Comments." I am indebted to Louis Loeb for a helpful discussion of Locke's view.

21  Price, *Review of the Principal Questions in Morals*, § 757.

22  Price, *Review of the Principal Questions in Morals*, § 758.

23  At one point, for instance, Balguy states (about Francis Hutcheson), "I humbly apprehend he is mistaken" (*The Foundations of Moral Goodness*, § 444). Surely it does not make sense to apprehend "humbly" if one can only apprehend what is true. That said, it is harder to find clear traces of a "veil of perception" epistemology in Clarke and Balguy. Interestingly, Thomas Reid, another rational intuitionist and one of the few early modern opponents of Locke's theory of ideas, also appears to have been a judgment internalist. In his *Essays on the Active Powers of Man* (1788), as reprinted in Raphael, *British Moralists*, Reid claims that "as soon as we have the *conception* of what is good or ill for us upon the whole, we are led, by our own constitution, to seek the good and avoid the ill" (§ 863, emphasis added).

what is *right* and fit."[24] Here Balguy attempts to explain the necessary connection between motivation and the perception of rightness by appealing to the supposedly similar connection between belief and truth. Now, admittedly it is not entirely clear what Balguy intends. He could mean that encountering or perceiving *actual truth* compels assent or merely that apprehending *the idea that something is true* compels assent. A "veil of perception" epistemology certainly suggests that latter, as does the nature of the connection to which Balguy is appealing. After all, merely the *judgment* that something is true is sufficient to compel assent. That is, one need not be correct in one's judgment that *p* is true is order to feel compelled to believe that *p*; one need only *judge* that *p* is true.[25] Thus, the case of truth and assent looks more like one of judgment internalism than perceptual internalism, and this suggests that Balguy has something similar in mind for the case of rightness and motivation.

Ultimately, though, it is not important to my project here whether we count the internalism of Clarke, Balguy, and Price as an instance of judgment internalism or instead as a third view distinct from both judgment and existence internalism. Perceptual internalism is the aspect of rational intuitionism that has come in for the most criticism over the years, and it has been jettisoned even by the contemporary inheritors of the intuitionist mantle.[26] After all, as John McDowell has observed, "it seems impossible—at least on reflection—to take seriously the idea of something that is like a primary quality in being simply *there*, independently of human sensibility, but is nevertheless intrinsically (not conditionally on contingencies about human sensibility) such as to elicit some 'attitude' or state of will from someone who becomes aware of it."[27] It is indeed impossible to take this seriously—at least for me. But what matters here is that after we discard Clarke, Price, and Balguy's utterly implausible brand of perceptual internalism, we are left with a thoroughgoing metaphysical externalism: the *existence* of reasons or obligations does not depend on any capacity to move or influence us.

---

24     Balguy, *The Foundations of Moral Goodness*, § 453, emphasis in original.

25     Nowadays this phenomenon is usually described in terms of belief's *constitutive aim:* belief constitutively aims at the truth. See Darwall, "Ethical Intuitionism and the Motivation Problem," 258–62 for a helpful discussion of Balguy's argument and its relation to contemporary views about belief's relation to truth.

26     See, for instance, Derek Parfit, "Reasons and Motivation," 127, and "Normativity," 336.

27     McDowell, "Values and Secondary Qualities," 132. This is also the aspect of intuitionism that Mackie famously (and rightly, in my view) dismissed as "queer." See Mackie, *Ethics*, 38–42.

## The Externalist Approach

The early modern rational intuitionists were convinced that any attempt to reduce the bindingness of obligation to some sort of motivational or psychological force would ultimately undermine the authority of morality. Underlying this conviction, at least in part, is their shared belief that normative notions like *rightness* and *obligation* cannot be explained or understood in *any* exclusively nonnormative terms. Price, for instance, insists that "our ideas of *right* and *wrong* are simple ideas.... He that doubts this, need only try to give definitions of them, which shall amount to more than synonymous expressions."[28] Price rejects the possibility of finding such definitions. Instead he embraces *nonreductionism*, the view that normative concepts and properties are *sui generis* and hence irreducible. Although we may be able to explain one normative expression in terms of another, we cannot analyze normative expressions into exclusively nonnormative or descriptive components. For Price, as well as for Clarke and Balguy, ethics constitutes an autonomous domain, independent of psychology, politics, and even religion.[29]

Nonreductionism plays a more prominent role in the work of twentieth-century externalists, receiving perhaps its most famous defense in G. E. Moore's *Principia Ethica* (1903). About the word "good" Moore famously declares: "it cannot be defined, and that is all I have to say about it."[30] Moore inherited this nonreductive view from his teacher Henry Sidgwick, who held that the term "right" is "ultimate and unanalysable."[31] In turn Moore bequeathed it

---

28 Price, *Review of the Principal Questions in Morals*, § 672. Later Price maintains that it is "impossible … to express, in any other than synonymous words, what we mean, when we say of certain objects, 'that they have a *fitness* to one another; or are *fit* to answer certain purposes" (§ 707). Clarke never explicitly accepts the irreducibility of normativity. It is, however, an implicit assumption of his criticisms of Hobbes. See also Reid, *Essays on the Active Powers of Man*, § 869.

29 That these externalists also accepted a nonreductive view of normativity is hardly surprising. Externalism and nonreductionism about normativity are a natural fit for one another. After all, as we saw in chapter 1, existence internalism is most compelling as a reductive account of the nature of normative authority. If normativity is *sui generis* and unanalyzable, then there is little motivation for the internalist idea that normative reasons must have the capacity to influence the agents for whom they are reasons. Philosophers who are repelled by reductionism in ethics will therefore naturally find themselves attracted to externalism.

30 Moore, *Principia Ethica*, 58. Moore's claim is, of course, about goodness rather than rightness, obligation, or reasons. But these others terms are definable in terms of goodness, according to Moore, and so he still holds a nonreductive view of fundamental normative notions.

31 Sidgwick, "Some Fundamental Ethical Controversies," 40. According to Sidgwick, one "cannot resolve ['right'] into, or explain it by, any more elementary notions" (40).

to his intuitionist successors, including H. A. Prichard, W. D. Ross, C. D. Broad, and A. C. Ewing, all of them externalists.[32] This view has also found plenty of adherents on the current scene. Foremost among them are the contemporary descendents of the rational intuitionists: Derek Parfit, T. M. Scanlon, Thomas Nagel, Joseph Raz, and Russ Shafer-Landau, as well as David McNaughton, Piers Rawling, Philip Stratton-Lake, William FitzPatrick, Philip Clark, and Graham Oddie.[33] We might also include nonreductive naturalists like Nicholas Sturgeon, Richard Boyd, Richard Miller, and David Brink in this group.[34]

Not all nonreductive externalists are interested in the regress of justification and the foundations of ethics, though. Strictly speaking, externalism is merely the denial of internalism, and some philosophers are externalists because they regard the whole project of finding an end to the justificatory regress as hopeless or pointless (or both). Allan Gibbard, for example, is a nonreductive externalist.[35] Yet he suggests that what those of us who are worried about the justification of our most basic normative commitments need is "not philosophy but therapy."[36] And even though he has criticized internalist accounts of the sources of normativity, he has not offered an account of his own. As the subtitle of his first book attests, the theory Gibbard

32    Prichard, "Duty and Interest," 47, and "Manuscript on Morals," 115; Ross, *The Foundations of Ethics*, 6–7, and "The Right and the Good," 91–3; Broad, *Five Types of Ethical Theory*, 170–1, 176–7; and Ewing, *The Definition of Good*, 132–6.

33    See Parfit, "Normativity"; Scanlon, *What We Owe to Each Other*, chap. 1; Nagel, *The View from Nowhere*, chap. 8; Raz, "Reason and the Will"; Shafer-Landau, *Moral Realism*; McNaughton and Rawling, "Naturalism and Normativity"; Stratton Lake, introduction to *Ethical Intuitionism*; Fitzpatrick, "The Practical Turn in Ethical Theory"; Clark, "Velleman's Autonomism"; Oddie, *Value, Desire, and Reality*; and Enoch, "An Outline of an Argument for Robust Metanormative Realism." See also Hampton, *The Authority of Reason*, chaps. 2–3; and Jonathan Dancy, *Practical Reality*, chap. 7.

        Most of these philosophers shun the "rational intuitionism" label, preferring instead something along the lines of "nonreductive, nonnaturalist realism." Regardless of their philosophical differences (and there are many), they are all nonreductive externalists.

34    See Sturgeon, "Moral Explanations" and "Ethical Naturalism," 98–9; Boyd, "How to Be a Moral Realist"; Miller, "Ways of Moral Learning"; and Brink, *Moral Realism and the Foundations of Ethics*, 37–50, 62–3, 144–70.

35    See Gibbard, "Knowing What to Do, Seeing What to Do, 224–7. Although he allows for the sort of indirect or "oblique" analyses that constitute his expressivism, Gibbard rejects the idea of "*straight* analyses" of normative concepts (*Thinking How to Live*, 185). See Nowell-Smith, *Ethics*, for another expressivist externalist.

36    Gibbard, "Morality as Consistency in Living," 146.

does provide is a theory of normative *judgments*, not of normative *reasons*.[37] John McDowell is another nonreductive externalist who has shown little interest in the foundations of ethics.[38]

On the other hand, many philosophers—including many of those mentioned above—reject internalism because they are drawn to a competing conception of the foundations of ethics. These philosophers adopt a nonreductive externalist account of normative authority, combine it with some form of robust normative realism, and take the resulting view to provide a solution to the regress of justification. In other words, they recognize a genuine philosophical problem here *and* they attempt to address it. In this chapter I shall focus on externalism of this second, more thoroughgoing variety. My target will be the philosophers who embrace an alternative approach to the foundations of ethics (and not those who reject internalism in favor of no approach at all). For the sake of convenience, I shall often use the term "externalism" as shorthand for this more ambitious view. By doing so, however, I do not mean to suggest that all of the philosophers who deny internalism fall prey to the objections aimed at the sort of externalism that interests me.

Unfortunately, the line between these two categories of externalists is not always clear. Prichard, for example, seems to waver between believing that he has given the only possible account of the foundations of ethics and believing that it is a mistake even to offer such an account. He dismisses worries about the regress of justification and maintains we should give up our hopes of proving or demonstrating that we really are obligated to keep our promises. But he also suggests that we can apprehend the source of obligation directly and that this apprehension should put any worries about a justificatory regress to rest.[39] Consequently, it is difficult to say with any certainty whether Prichard is an externalist. Does he provide an account or reject the project out of hand? A similar question arises regarding Ross. Despite the title of his second great ethics treatise (*The Foundations of Ethics*), Ross often appears uninterested in grounding our most basic ethical commitments.

Here, however, my concern is not the question of whether this or that philosopher is an externalist, but rather the plausibility of the externalist approach in general. And there

---

37      The full title is *Wise Choices, Apt Feelings: A Theory of Normative Judgment*.

38      See McDowell, "Values and Secondary Qualities." McDowell's fellow sensibility theorist, David Wiggins, seems to be more ambitious. See Wiggins, "A Sensible Subjectivism?" and "Reply to Peter Railton."

39      See Prichard, "Does Moral Philosophy Rest on a Mistake," 4, 7, and 19, as well as "Manuscript on Morals," 114–7,

certainly are a number of thoroughgoing externalists who take themselves to be addressing the foundational questions that interest me. These philosophers have many normative and metanormative disagreements, but they are united in their view that, in Parfit's words, "when we have a reason to do something, this reason is not provided by, and does not require, the fact that after [sound] deliberation we would want to do this thing."[40] A consideration's status as a reason for me to act does not depend on any role that consideration might play in my rational deliberations. But then on what *does* it depend? In answering successive iterations of this question, these externalists agree that we will eventually arrive at a primitive normative reason that "cannot be explained in non-normative terms": "we cannot say what makes suffering bad."[41]

This view certainly has its attractions, as the distinguished names listed above attest. I shall argue that we should reject it nonetheless.

## The Metaphysical Question

In *The Sources of Normativity*, Christine Korsgaard articulates a powerful objection to the externalist approach. There she argues that a nonreductive approach to justification offers no help to agents struggling with what she calls "the normative question."[42] The normative question, according to Korsgaard, "is a first-person question" confronted by agents in the course of practical deliberation.[43] For Korsgaard, then, the problem of the regress of justification—of finding the foundations of ethics—is a practical problem, one that an agent encounters from his "own, first-person perspective."[44] Moreover, since it is ultimately a problem for deliberating

---

40    Derek Parfit, "Reasons and Motivation," 130.

41    Derek Parfit, "Normativity," 330–1.

42    Korsgaard, *The Sources of Normativity*, 28–42. Here Korsgaard claims that her target is "substantive realism" (34–7). See also Korsgaard, "Realism and Constructivism," 99–122. A number of philosophers have observed that it is difficult to discern any interesting metaethical differences between "substantive realism" and Korsgaard's favored position, "procedural realism" or "constructivism." I suspect that this is because her disagreement with realist philosophers like Clarke, Prichard, and Nagel has little to do with their realism. The genuine dispute concerns the issue of internalism and externalism. It would not be fair to blame Korsgaard for this misunderstanding, though, since she inherited the conflation of realism with externalism from the rational intuitionists themselves.

         For more on Korsgaard's discussions of realism, see Hussain and Shah, "Misunderstanding Metaethics"; and FitzPatrick, "The Practical Turn in Ethical Theory."

43    Korsgaard, *The Sources of Normativity*, 16.

44    Korsgaard, *The Sources of Normativity*, 14.

agents, "the answer must actually succeed in *addressing* someone in that position."[45] Korsgaard's objection to externalism is that it fails to do this. Suppose, for example, I am trying to decide whether to donate money to Oxfam. My friends and family insist that the suffering of people on the other side of the world is a strong reason for me to contribute, but—for whatever reason—at this moment I just do not see it. I know this is what everyone thinks; however, I fail to recognize or feel the force of this consideration. Korsgaard contends—rightly, I think—that the assurances of a nonreductive externalist will do little to sway me. If my friends and family cannot reach me, then an externalist's confident insistence that suffering on the other side of world *really is a reason for me to donate* is hardly going to make a difference.[46] As Korsgaard concludes, "it is a little hard to see how [externalism] can help."[47]

Critics have sometimes interpreted Korsgaard as insisting that a successful philosophical account of the foundations of ethics must be able to *convince* or *motivate* agents struggling with the normative question. G. A. Cohen, for example, suggests that Korsgaard's objection to externalism is that it cannot "convert the radically disaffected."[48] If this is indeed the thrust of her objection, then it implies that a satisfying theory of the sources of normativity must be able to persuade me that the suffering of others really is a reason for me to donate to Oxfam.

But this is an unreasonable demand to make on any philosophical theory. According to externalism, it is a brute fact about the world that certain considerations are intrinsically normative. If, within my practical deliberations, I cannot recognize their normativity, then my problem may not be a philosophical one. I might be rebellious or stubborn. Alternatively, I might be suffering from a psychological disorder like depression or ennui. Or I could simply be dense. It is not a flaw of externalism that it cannot convince me to acknowledge reasons to help others. To see this, compare the case of someone who, while working out a math problem, suddenly begins to wonder whether two and two together actually equal four. He knows that this is what everyone says, but at precisely that moment he cannot see it. Should a

---

45   Korsgaard, *The Sources of Normativity*, 16. Compare Harman, "Moral Relativism Defended," 5.

46   Of course an externalist might be able to explain this particular obligation in terms of some other reason for duty. The point, though, is that before long any explanation he offers will bottom out in an irreducibly normative reason—say, that a person's suffering is a reason for me to help him.

47   Korsgaard, *The Sources of Normativity*, 34.

48   Cohen, "Reason, Humanity, and the Moral Law," 180. See also Parfit, "Normativity," 368–71; and FitzPatrick, "The Practical Turn in Ethical Theory," 682–4.

philosophical theory of mathematics be charged with the task of helping this benighted soul see the light? It is certainly hard to see how a philosopher who adopts a nonreductive approach to the nature of mathematics is going to be able to help, but will theorists of a different stripe be any more successful? This does not seem like a problem for philosophers at all. Imagine instead someone who starts to wonder whether he can pass through solid objects. Sure, a metaphysician with strangely nonreductive views on the nature of medium-sized dry goods is hardly going to be able to convince him otherwise, but will a reductive physicalist be any more helpful? It seems that what we need here is not a metaphysician or even a physicist, but rather a psychiatrist (or a neurosurgeon). The same goes, I think, for radically disaffected agents struggling with the normative question. Externalists may indeed be useless when it comes to persuading recalcitrant agents to act in a certain way, but this is not the sort of problem an account of the foundations of ethics needs to solve anyway.

Korsgaard is best interpreted as identifying a different failing of the externalist approach. Even if a philosophical theory should not be burdened with the task of moving a stubborn or dimwitted agent, it is (or at least should be) burdened with the task of addressing the doubts of a rational and reflective agent.

But reflective agents can have more than one sort of doubt, and what Korsgaard calls "the normative question" is therefore ambiguous. The "why" question she describes is open to more than one interpretation. When I ask why I should donate to Oxfam, the most natural interpretation of my question is as a request for justification. That is, I want to know whether I have reasons to heed the norms that recommend (or demand) charity. Alternatively, we might say that while I know I have a moral reason to donate, I am asking whether I have an independent and more basic reason to be moral. The way to answer such questions is to appeal to some other norms for support. This is the strategy Plato adopts in the *Republic*. There he suggests that the norms of justice derive their authority from the authority of the norms of self-interest. If we interpret Korsgaard's "normative question" along these lines, then it really is a *normative* question—that is, a call for justification.

Internalists and externalists both acknowledge that the process of asking and answering successive *normative* "why" questions will eventually lead us to norms that are intrinsically justified, or norms that do not depend on other norms for justification. In other words, both sides allow that the regress of "why" questions will reach a point at which further justification

is impossible and further normative questions inappropriate. Sooner or later we will arrive at reasons for which there is no further normative explanation. For instance, this point might arrive when we reach considerations of suffering. If I ask what justifies the norms that recommend the avoidance of suffering, it might turn out that *nothing* justifies these norms; they are intrinsically justified. If so, then we will have reached normative bedrock. Suffering is bad for its own sake; we have an unqualified reason to avoid it.

As I have already said, that normative questions are inappropriate once we reach normative bedrock is something both internalists and externalists accept. However, normative questions are not the only sorts of "why" questions. When I ask why I should avoid suffering, I need not be wondering what further norms underwrite the badness of suffering. Explanation can proceed in more than one dimension, and it is their different attitudes towards these other modes of explanation that ultimately set internalists and externalists apart. Internalists offer a *metaphysical* explanation of why our most basic reasons are reasons. If the capacity for a certain sort of motivational influence is constitutive of normative authority—if this is what being a reason consists in—then internalists *can* explain why we have a reason to avoid suffering. Their explanation will be metaphysical rather than normative, but it will be no less of an explanation for that.[49]

Externalists have nothing comparable to offer. They hold that some reasons are unqualifiedly normative and that this is a metaphysically primitive feature of the universe. Some norms simply are authoritative. Some things just are intrinsically good. Some considerations simply are unqualifiedly normative reasons. If we ask why we have a reason to avoid suffering, the externalists' answer is the same regardless of whether our question is normative or metaphysical: We just do.[50]

---

49    I am grateful to Jyl Gentzler for helping me formulate the ideas in this section and especially in this paragraph.

50    See Parfit, "Normativity," 379. One might object that here I am ignoring value-based theories of ethics—that is, theories that ground normative authority in values. (For a few such theories, see note __ in the introduction above.) As I stated in the introduction, I am drawn to the view that *reason* and *ought* are the fundamental normative notions, and that "goodness" should therefore be explained in terms of reasons—reasons to desire or value. (Compare Scanlon, *What We Owe to Each Other*, ch. 2.) Moreover, grounding the normativity of reasons in a nonreductive theory of value will not escape my criticism here. Being told that some things "just are good" is hardly more satisfying than being told that some things "just are right." In both cases, the nonreductive nature of the fundamental normative proprety forecloses any further investigation or explanation.

Many have complained that this is a less than fully satisfying account of the foundations of ethics. Externalists claim to have identified reasons that are intrinsically justified, but they have not provided a metaphysical explanation of why these reasons are the foundational ones. I think that Korsgaard is most charitably interpreted as objecting to *this* feature of externalism. The "why" question that externalists uniquely fail to answer is not really a normative question at all. It is a question about the metaphysical grounds of normativity. If we read Korsgaard's "normative question" as a version of this metaphysical question, we can see why she is so frustrated with the externalist replies. Their answer to the question of why suffering has its foundational status is that *it just does*. But that is not an answer to our metaphysical question so much as a dismissal. It sheds no light at all on the matter that puzzles us.

In the next two sections I shall consider the ways in which externalists respond to this charge.

### The Obvious and the Obscure

I, for one, am in complete agreement with externalists that the suffering of people on the other side of the world is an excellent reason for me to donate to Oxfam. In fact, this year's check is already in the mail. But now, sitting in my philosopher's armchair, I have started to wonder about that reason. I certainly continue to accept that I have one, but where precisely does it come from? What makes it a reason? In virtue of what does it have its peculiar authority? These are not practical questions, for I have already resolved my practical deliberations. They are metaphysical questions: I have come face to face with a phenomenon I want to understand better, for I find—when I really think about it—that I do not understand it well at all. This is precisely the sort of situation that calls for a philosophical theory, and externalism tell me nothing beyond what I already suspect, namely, that I had a very good reason to donate to Oxfam.

Externalists are generally not too exercised by this apparent failing. In fact, they are usually quite unabashed. Parfit's statement (quoted above) that we cannot understand the source of suffering's badness is surprisingly nonchalant.[51] This may be because Parfit, like Clarke before him, thinks that the badness of suffering is obvious. According to Clarke,

---

51    Compare Parfit later in the same piece: "The badness of suffering … needs [no] further explanation" ("Normativity," 379).

fundamental normative truths are "as evident and undeniable" as the basic truths of mathematics.[52]

> These things are so notoriously plain and self-evident, that nothing but the extremest stupidity of mind, corruption of manners, or perverseness of spirit, can possibly make any man entertain the least doubt concerning them.[53]

Clarke is clearly at a loss to understand how anyone could even wonder about the philosophical status of these fundamental truths. As Korsgaard notes, "obviously *he* isn't worried."[54] Neither are most contemporary externalists. At the opening of his chapter on reasons, Scanlon blithely declares that "the idea of a reason does not seem to me to be a problematic one that stands in need of explanation."[55] Like Clarke, Scanlon simply does not recognize any philosophical problem here. He defends externalism's silence on the metaphysical sources of suffering's badness by appealing to the fact that it is simply obvious.

This reaction betrays either a misunderstanding of Korsgaard's objection to externalism or a surprisingly unphilosophical attitude towards metaethics. On the one hand, externalists might be conflating metaphysical and epistemological consideration. We are asking about the metaphysical foundations—the metaphysical first principles, as it were—of basic normative facts, and it is no response to this challenge to say that the truth of these facts is obvious. Externalists answer our request for *metaphysical* first principles by supplying *epistemological* first principles: basic normative facts are self-evident. But that these facts are self-evidently true does not constitute an explanation of them. It may explain how we know them, but that is another matter entirely. Clarke and Scanlon respond to our metaphysical questions with confidence; and confidence, contagious though it sometimes may be, is not an explanation.[56]

52    Clarke, *Discourse Concerning the Unchangeable Obligations of Natural Religion*, § 226.

53    Clarke, *Discourse Concerning the Unchangeable Obligations of Natural Religion*, § 227. Compare Price, who regards basic moral rules as "axioms, the truth of which appears as irresistibly as the truth of those which are the foundation of geometry" (*A Review of the Principal Questions in Morals*, § 747).

54    Korsgaard, *The Sources of Normativity*, 39.

55    Scanlon, *What We Owe to Each Other*, 17. Earlier in the introduction, Scanlon writes: "I do not believe that we should regard the idea of a reason as mysterious, or as one that needs, or can be given, a philosophical explanation in terms of some other, more basic notion. In particular, the idea of a reason should not be thought to present metaphysical or epistemological difficulties that render it suspect" (3). Compare Ross, *The Right and the Good*, 29–30.

56    Compare Korsgaard, *The Sources of Normativity*, 39–41. In "Normativity," Parfit considers seven different interpretations of Korsgaard's normative question, none of which correspond to my metaphysical interpretation. See Parfit, "Normativity," 355–8.

On the other hand, Clarke and Scanlon might mean to suggest not that epistemological first principles somehow answer metaphysical questions, but rather that they make such questions irrelevant: externalism is silent because nothing needs to be said! Why bother constructing a metaphysical theory of normativity when the facts such a theory offers to explain are already so obvious? Since Clarke and Scanlon see no need for a metaphysical explanation of the foundations of ethics (given the self-evident nature of those foundations), externalism's failure to provide one does not strike them as a shortcoming. They justify their lack of a solution to our metaphysical problem by suggesting that there really is no problem worth solving.

As G. J. Warnock says, this approach "consists in a protracted denial that there is anything of the slightest interest to be said."[57] Here I shall pause for a brief *argumentum ad hominem*, since there seems to be something deeply antithetical to philosophy in this quietism about the metaphysical foundations of ethics. Serious or even slight doubt about some range of supposed truths is hardly a requisite for philosophical inquiry into the nature of those truths. Confidence that two and two together really do equal four is not a reason to do away with metaphysical theories of the nature of numbers and mathematics. My certainty that two and two equal four does not lessen my desire to understand the source of that certainty, nor should it. Philosophical curiosity is more than enough to justify an investigation into the nature of mathematics. Of course there is nothing wrong with a lack of curiosity about the nature of mathematical truth, but it would be an odd trait to find in a philosopher of mathematics. Yet an analogous trait is exactly what we find in externalists like Clarke and Scanlon. Here we have a group of philosophers—self-described ethicists and metaethicists—who are apparently uninterested in an investigation into the nature of normative authority. Again, there is nothing wrong with such a lack of interest, at least by itself. For all of us there are surely areas of philosophy that fail to pique our curiosity. That said, it is one thing to lack curiosity about the foundations of ethics; and it is quite another to lack curiosity but to publish papers about those foundations that declare, "I just don't see what the big deal is! Isn't all of this just self-evident? What more do we need to say?"[58]

---

57 Warnock, *Contemporary Moral Philosophy*, 13. Warnock continues: "The effect of this is worse than unhelpful; it is positively misleading" (13).

58 Nick Zangwill responds to Ronald Dworkin's nonreductionism with a similar charge: "Dworkin does not see the question. There is somehow nothing to be explained. But this is just to ignore the philosophy"

Moreover, what if I *am* worried? What if I *do* have doubts about my reasons? What if, in other words, I disagree about the epistemological first principles? If it is only confidence that underlies the claim that "the badness of suffering … needs [no] further explanation," what can we say to the philosophers who lack such confidence? The suffering of people on the other side of the world sure *seems* like a reason for me to donate to Oxfam, but normative authority is not always what it seems. How could it be, when there are so many people to whom suicide bombings or acts of genocide seem perfectly justified? Doubts along these lines will only be magnified by externalist claims that normativity is just not something we should even bother trying to explain. An unwillingness to say anything interesting about a particular phenomenon is at least a sign that the phenomenon may not be all it's cracked up to be. Surely the lack of an explanation signals the need for philosophy, not the end of it.

In response to such worries about the justificatory regress, externalists like Parfit, Scanlon, and Clarke simply point to a consideration and declare, "See that? *That* is an irreducibly normative reason." They are so certain that suffering is bad that they cannot see what all the fuss is about. In this way they are somewhat like Moore, who was so certain he had a hand that he could not bring himself to worry about Cartesian skepticism.[59]

Moore famously offered his hands as a "perfectly rigorous proof" of the existence of things outside of us, and hence as a refutation of Cartesian skepticism.[60] But, as many critics have observed, Moore's "proof" depends crucially on his prior conviction that the skeptical possibility—namely, that he has no hands—is "absurd."[61] If, like Moore, we do not take seriously the possibility that our commonsense beliefs are radically misguided, then of course we will see those famous hands and be immediately convinced of the existence of things outside us. Moreover, nothing the skeptic can say will be able to shake this conviction. A valid argument always leaves us with the choice of either accepting the conclusion or denying one the premises.

("Zangwill Reviews Dworkin").

59    Compare Prichard, "Does Moral Philosophy Rest on a Mistake?" 19.

60    Moore, "Proof of an External World," 145–7.

61    For related criticism of Moore's argument, see Lehrer, "Why Not Skepticism?" and Unger, "An Argument for Skepticism"; as well as Stroud, *The Significance of Philosophical Scepticism*, 83–127. For a defense of Moore's approach to skepticism, see Lycan, "Moore Against the Skeptics." But then see also Conee, "Comments On Bill Lycan's 'Moore Against The New Skeptics,'" for a response to Lycan's defense. Finally, for an entirely different take on Moore's project in "Proof of an External World," see Proops, "Soames on the Metaphysics and Epistemology of Moore and Russell."

And any set of skeptical premises will be less credible or plausible than the nonskeptical beliefs those premises contradict. Thus, if we already regard the possibility that we are radically deceived as incredible, then skeptical arguments will provide us with no good reason to change our mind. But if instead we take the premises of the skeptical argument seriously—if, for example, we find the *a priori* claim that our subjective experiences are consistent with radically different objective realities just as compelling as the commonsense claim that Moore has hands—then Moore's so-called proof carries no weight at all.[62] Ultimately, Moore provides a refutation of skepticism only for those who have already dismissed it as absurd.

There is something deeply unsatisfying about Moore's position, especially for philosophers who do entertain doubts about the justification of our beliefs. My point about Parfit, Scanlon, and Clarke is that their approach is similarly unsatisfying.[63] Motivated as it is by a conviction that the foundations of ethics are self-evident, it offers nothing to those of us who do not share their dogmatic confidence.

As I noted above, to the extent that my comments are directed at the motives behind nonreductive externalism rather than at its truth or falsity, they amount merely to an *argumentum ad hominem*. But when it comes to an issue as abstract as the foundations of ethics, theory choice is rarely as simple as an easy check for truth or falsity. What I have tried to show here is that the attractiveness of at least a certain strain of nonreductive externalism is often symptomatic of either a dogmatic and hence unphilosophical approach to ethics or a simple lack of interest in the nature of normative authority.

---

62   Note that one need not *be* a skeptic in order to find that Moore's argument misses its mark. One need merely regard skeptical worries as serious and genuine philosophical concerns

63   Korsgaard describes what is unsatisfying about this approach by saying that the nonreductive externalist "bring the regress to an end by fiat" (*The Sources of Normativity*, 33). Parfit's rejoinder misses the point in a fascinating way. He contends that it is *internalists* who end the regress by fiat, since they are the ones who think that a motive or an act of will is sufficient for normativity ("Normativity," 380). I shall return to this criticism of internalism in chapter 3 below. What Parfit fails to see, though, is that Korsgaard's charge is metaphilosophical. Strictly speaking, her problem here is not with the plausibility of the claim that some reasons are irreducibly normative. What she objects to are the philosophical theorists who simply declare that such reasons exist and thereby take themselves to have solved the problem of the regress of justification.

Compare Warnock, *Contemporary Moral Philosophy*, 16–7. Warnock argues that only a lack of doubt and curiosity on the part of intuitionists could explain why they are attracted to a view that is so "unilluminating" (17).

To this effect, consider the puzzlement Allan Gibbard once expressed about my dissertation project.[64] He did not think it worthwhile to search for a philosophical explanation of basic normative truths, since even were such an explanation forthcoming, it could never be as plausible or credible as the normative truths themselves. In other words, no account of the foundations of ethics could be as compelling as the claims such an account is supposed to ground. Gibbard had more confidence in his judgment that pain is bad than he would ever have in any philosophical theory, and so he did not see the point of looking for a theory that might support that judgment. After all, what is the point—to borrow a nice turn of phrase from Nagel—of "explaining the obvious in terms of the obscure"?[65]

My response to this challenge is that from the philosopher's armchair, it is *not* obvious that the suffering of people on the other side of the world is a reason for me to help them. Or at least it is not obvious to me. When I am washing dishes it is similarly obvious that I have hands, but this hardly undermines the epistemologist's project of trying to explain how such knowledge is possible. In a way it is the very basis of that project, for philosophy just is the attempt to understand and explain, in a systematic way, things that seem too mundane even to talk or think about during the course of everyday life.[66] Furthermore, once one of these mundane matters is put under the philosophical microscope, it is often revealed to be anything but simple and straightforward. This is certainly the case when it comes to the foundations of ethics: the sources of normativity are deeply mystifying. What is obvious to Gibbard, then, is not obvious to everyone, and that alone is reason enough to search for an explanation, even if it must ultimately be couched in terms of the relatively obscure. For even an obscure metaphysical account of the sources of normativity will be able to explain (obscurely) why the considerations that seem to be reasons *really are reasons*.

Now I am undoubtedly being unfair here, since not all externalists are motivated by the sort of blasé attitude I have been criticizing. Although he remains silent on the subject of

64    Gibbard, in discussion during candidacy seminar, November 17, 2005.

65    Nagel, *The View from Nowhere*, 160. Gibbard makes a related claim in *Thinking How to Live:* "What … does the expressivist's direction of explanation gain us? A kind of vindication for our normative concepts, say I—but why do they need vindication? We wouldn't, after all, give them up if no vindication were to be had" (195). Gibbard continues: "many of our normative judgments—that suffering is worth avoiding, that seeing supports believing—are far more credible than any view could be that they are all nonsense" (195).

66    Contrast Scanlon, who suggests that only the threat of a "genuine skepticism about reasons" would justify a serious investigation into the nature of reasons (*What We Owe to Each Other*, 19).

the foundations of ethics, Gibbard, for one, has plenty to say about the nature of normative judgments and of the psychological states these judgments express. Moreover, explanation must surely come to an end at some point. And for many of the externalists I have been discussing, behind their insistence that we do not need an explanation of normative authority lies the firm conviction that no such explanation is possible. They believe we have struck normative *and* metaphysical bedrock. In the next section I shall evaluate their reasons for thinking that normativity cannot be explained.

## Philosophical Pessimism

At the outset of *What We Owe to Each Other*, Scanlon declares that he will "take the idea of a reason as primitive."[67] According to Scanlon, accounts of what it is to be a reason always end up circling back to the notion of a reason.

> Any attempt to explain what it is to be a reason for something seems to me to lead back to the same idea: a consideration that counts in favor of it. "Counts in favor of it how?" one might ask. "By providing a reason for it" seems to be the only answer.[68]

As we have already seen, Parfit agrees: "normative concepts cannot be explained in nonnormative terms."[69] Externalist Jean Hampton goes so far as to say that normative authority is "ineffable"—that it is "impossible to pin down in a way that makes sense."[70] At best, these externalists contend, we can explain one sort of normative claim (say, that R *is a reason* for you to *ϕ*) by appealing to a different sort of normative claim (say, that R *counts in favor* of you *ϕ*-ing). Attempts to understand normativity in nonnormative terms, however, are bound to miss their target. And so, they conclude, we should not be frustrated with externalists' unwillingness to investigate the nature of normativity, for normativity is just something that cannot be explained.

The *locus classicus* for the claim that normativity is simple and hence irreducible is Moore's open question argument. Moore famously argues that any attempt to define the word "good" in nonevaluative terms is bound to fail. If some definition of the form "goodness is *x*"

---

67    Scanlon, *What We Owe to Each Other*, 17.

68    Scanlon, *What We Owe to Each Other*, 17.

69    Parfit, "Normativity," 330. Compare Nowell-Smith, *Ethics*, 32.

70    Hampton, *The Authority of Reason*, 99.

(where *x* is a nonevaluative property) were correct, then the question whether *x* is good would be conceptually closed. Since no such questions really are closed—since, in other words, it is always an open question whether *x* is in fact good—no such definition can be correct.[71] Goodness, according to Moore, is "simple, indefinable, and unanalysable."[72]

The open question argument has come in for more than its fair share of criticism over the years. It is now widely recognized that the "fallacy" this argument is designed to reveal, if it is a fallacy at all, is to be found only in fairly crude, analytic versions of reductionism. Nowadays most philosophers accept that not all reduction is analytic: developments in twentieth-century metaphysics and the philosophy of language have opened the door to property identity without concept identity. Consequently, we might be able to say something illuminating about the *nature* of normative reasons (and of their normativity) without claiming that this reveals anything about the *concept* of such reasons. Nevertheless, philosophers committed to the irreducibility of normativity continue to look to Moore's argument for inspiration. They argue that even synthetic reductions based on property identity cannot capture the normativity for which they are supposed to account.

To see this Moorean style of argument in action, consider Parfit's discussion of a crude but nonanalytic reductive internalism.[73] According to this view, although the statement that *A* has a reason to *ϕ* is not synonymous with (and hence not analytically reducible to) the statement that *ϕ*-ing would satisfy one of *A*'s desires, these two statements nonetheless report the same fact. Parfit then presents the following example of practical reasoning:

Jumping would satisfy many of my desires.

Therefore, I have a reason to jump.

71 See Moore, *Principia Ethica*, 67–9, 94–5. It is, of course, a matter of great controversy what—if anything—Moore's open question argument actually shows (or is even supposed to how). The classic discussion of the argument (and of the "fallacy" it is supposed to reveal) is William Frankena's "The Naturalistic Fallacy" (1939). For various more recent takes on this subject, see Darwall, Gibbard, and Railton, "Toward *Fin de siècle* Ethics," 115–21; Darwall, "Moore, Normativity, and Intrinsic Value"; Gibbard, *Thinking How to Live*, chap. 2; Rosati, "Naturalism, Normativity, and the Open Question Argument" and "Agency and the Open Question Argument"; Regan, "How to Be a Moorean," 657–61; Sturgeon, "Moore on Ethical Naturalism"; Altman, "Breathing Life into a Dead Argument"; Feldman, "The Open Question Argument"; and Stratton-Lake and Hooker, "Scanlon versus Moore on Goodness."

72 Moore, *Principia Ethica*, 72.

73 See Parfit, "Naturalism," 8–9.

Parfit claims that on the reductive account in question, this reasoning is actually trivial. Since the premise and the conclusion do not report different facts, a more accurate depiction of this practical reasoning would be:

> Jumping would satisfy many of desires.
>
> As another way of reporting this fact, I could say: I have a reason to jump.

Parfit complains that this synthetic reductive internalism "grotesquely distorts" the nature of practical reasoning.[74]

> When I conclude that I should jump, I am not redescribing my argument's premise. I am drawing a quite different, normative conclusion. If my conclusion merely redescribed the [nonnormative premise], it could not be normative.[75]

This is a variation on Moore's argument, cast in terms of normativity. Parfit's claim is that if these two properties really were identical, reasoning from the premise that jumping is what I most want to the conclusion that I ought to jump would be trivial. Since such practical inferences are never trivial, the two properties must not be identical.

Parfit is correct, I think, that this particular reductive view grotesquely distorts our practical reasoning. But this hardly shows that, in Parfit's words, "similar claims apply to all forms of Reductive Internalism."[76] The above account is distortive not because it is reductive, but rather because the particular reduction it posits is so utterly implausible. If we are going to assess the overall prospects for reductionism by evaluating particular reductionist proposals, we should at least consider the strongest and most plausible. After all, reductive internalist accounts of normativity have come a long way since Richard Cumberland's pioneering effort.[77] Instead of focusing on the obvious flaws of the view that reasons can be reduced straightforwardly to desires, we should look to the sophisticated theories of philosophers like David Lewis, Peter Railton, Michael Smith, Christine Korsgaard, or David Velleman.[78] But

---

74    Parfit, "Naturalism," 9.

75    Parfit, "Naturalism," 9.

76    Parfit, "Naturalism," 9.

77    See Darwall, *The British Moralists and the Internal 'Ought'*, ch. 4, for the view that Cumberland was the first reductive naturalist internalist.

78    See Lewis, "Dispositional Theories of Value"; Railton, "Moral Realism," "Moral Explanation and Moral Objectivity," and "Facts and Values"; Smith, *The Moral Problem*, chaps. 5–6; Korsgaard, *The Sources of Normativity*, chaps. 3–4; and Velleman, *The Possibility of Practical Reason*. Parfit does discuss some of these authors. See "Naturalism," 23–5, for his criticisms of Smith and "Normativity," 355–80, for his take on Korsgaard.

even if Parfit can show that all of these proposed reductions fail—even if he can demonstrate that the properties identified by these accounts are not in fact normative—this will not establish the irreducibility of normative properties and the concomitant failure of reductionism as a general strategy in ethics: a reduction that is not grotesquely (or even mildly) distorting might be just around the corner.[79]

Open question considerations can show that reductionism will *always* fail only if they are supported by the assumption that normativity is "simple, indefinable, and unanalysable." Of course with this assumption on board, the Moorean nonreductionist conclusion follows straightaway. Absent such an assumption, though, the open question argument amounts to nothing more than a hurdle that any successful reductive theory must be able to clear.[80] Parfit and his fellow nonreductive externalists may very well be right that no theory to date has been able to clear this hurdle. (In fact, I shall side with Parfit against a number of reductive internalist accounts in chapter 3.) But we cannot determine in advance whether the next reductive account will stumble as well: synthetic reductionism aims to provide an *a posteriori* account of normative facts, and thus we cannot judge its success *a priori*.[81]

If the open question argument does not support the externalists' contention that normativity is irreducible, then what does? Explanation must indeed stop *somewhere*, but why must it stop *here*?

Parfit draws support for his claim that we cannot say anything informative about normativity from the irreducibility of other concepts, notably mathematical, modal, and temporal concepts. Like these concepts, "normative concepts form a fundamental category."[82] Thus, we should no more expect to explain normative concepts in nonnormative terms than we should expect to account for temporal concepts in nontemporal terms. Normative truths, Parfit maintains, "are truths of a distinctive kind."[83]

79   For similar claims, see Railton, "A Reply to David Wiggins," 319; Boyd, "How to Be a Moral Realist," 206–9; and Sturgeon, "Moral Explanations," 241–3.

80   Compare Darwall, Gibbard, and Railton, "Toward *Fin de siècle* Ethics," 177–8; and Railton, "Naturalism and Prescriptivity," 158–9.

81   Compare Schroeder, "Realism and Reduction," 7.

82   Parfit, "Normativity," 331.

83   Parfit, "Normativity," 331. As an aside, it is surprising that a philosopher who defends such a radically reductive account of personhood would be so opposed to reduction in ethics. Given Parfit's conservative approach to metaethics, it is quite jolting to return to the introduction to *Reasons and Persons* and read:

Notice that here Parfit moves from the claim that normative *concepts* form a fundamental category to the claim that normative *truths* or *facts* are of a distinctive kind. This simply ignores the possibility of a nonanalytic reduction. Nonanalytic reductive internalists can agree that normative *concepts* are irreducible, since from this it does not follow that normative *properties* are irreducible as well. Of course our normative concepts will place limits on what counts as a successful reduction of normativity: any proposed reduction must answer to our general understanding of normativity and reasons for action (such as it is). Hence we can probably rule out the theory that normativity is a shade of blue, even before hearing any considerations in its favor.[84] But it is also crucial to bear in mind that a successful reduction need not be true to *every* facet of our current understanding of normativity. Before water's chemical structure was revealed, for instance, it was widely believed to be one of the primary and irreducible elements of the universe. The discovery that water is a compound of two gases contradicted this widespread belief. This did not, however, constitute a failure of the reduction. It merely demonstrated that the older conception of water was, at least in part, in need of revision. Similarly, the scientific discovery that solid objects are largely composed of empty space did not result in the elimination of our concept of solidity.[85] It did require that we jettison a central aspect of this concept—namely, that solid objects are entirely full of matter—but the concept survived even this fairly drastic revision. A successful reduction of reasons for action, then, may require and subsequently survive an analogous modification—this time to our general concept of normativity. To conclude in advance that any such revision would mark an unsuccessful reduction is simply to beg the question against reductionism.[86]

But this point aside, Parfit's defense of the ineffability of normativity rests on the assumption that reductionism is not a live option when it comes to the other "fundamental categories" he mentions. This is simply mistaken. In the philosophy of mathematics, for example, there is the ongoing logicist project—begun by Frege, furthered by Russell,

---

"By temperament, I am a revisionist.… I try to challenge what we assume. Philosophers should not only interpret our beliefs; when they are false, they should *change* them" (x).

84   Compare Parfit, "Normativity, Naturalism, and Noncognitivism," 21.

85   Thanks to Peter Railton for suggesting this example.

86   Compare Stevenson, "The Emotive Meaning of Ethical Terms," 10–1. For more on the reduction of normative properties to natural properties, see Railton, "Naturalism and Prescriptivity," "What the Non-Cognitivist Helps Us to See the Naturalist Must Help Us to Explain," and especially "Reply to David Wiggins."

Whitehead, and Carnap, and promoted today by Crispin Wright and Bob Hale—to reduce mathematics to logic. Hartry Field, meanwhile, defends a more radical, eliminative reduction of mathematics to physics.[87] And in general, despite the difficulties of finding an adequate account of mathematical truth (and of necessary truths in general), philosophers of mathematics continue to plug away. The same is true in the case of modality. Alan Sidelle has proposed a conventionalist and empiricist account of necessity according to which facts about necessity are grounded in "our way of speaking and thinking."[88] J. L. Mackie defends a similar view. David Lewis's modal realism is, despite its name, another example of a reductive theory of modality. And there are a number of philosophers who believe we can analyze necessity in terms of counterfactuals.[89] When it comes to the philosophy of time, matters are admittedly more complicated, but even here there are reductive views in the mix. Kant famously argued that time is nothing more than the form of inner sensibility, and philosophers are still struggling with J. M. E. McTaggart's claim that time is ultimately an illusion.[90] Moreover, our understanding of the nature of time continues to be influenced by work in theoretical physics: consider the effect that Einstein's special theory of relativity has had on our notion of simultaneity.[91] Who knows what future scientific discoveries have in store for our current concept of time?

What matters here is not whether any of these reductive projects can be shown to have succeeded. The point is simply that Parfit cannot appeal to mathematical, modal, or temporal

87    See Hale and Wright, *The Reason's Proper Study*; and Field, *Science without Numbers*. Field, sometimes called a fictionalist about mathematics, denies not only that mathematical truths are necessarily true, but that they are true at all.

88    Sidelle, *Necessity, Essence and Individuation*, 2.

89    See Mackie, "Counterfactual and Causal Laws" and *Truth, Probability and Paradox*; Lewis, *On the Plurality of Worlds*; Roper, "Towards an Eliminative Reduction of Possible Words"; and Kment, "Counterfactuals and the Analysis of Necessity." Thanks to Joshua Brown for a very helpful discussion of these issues.

90    For the view that Kant's theory of time is consistent with modern physics, see Hacyan, "On the Transcendental Ideality of Space and Time in Modern Physics." Gödel actually thought that special relativity *entailed* Kantian idealism about time ("A Remark about the Relationship between Relativity Theory and Idealistic Philosophy").

91    In his assault on reductionism, Parfit also appeals to the fact that normative truths are necessarily true. According to Parfit, necessary truths could never successfully be reduced to contingent natural truths. But this simply ignores the possibility of *a posteriori* property identifications. That water is $H_2O$ may very well be a necessary truth. If it is, then a reductive account of normativity can allow for similarly necessary truths. See Putnam, *Reason, Truth and History*, 205–11.

concepts in order to support his claim that normativity is irreducible, since reductionism is a live option in those disciplines as well.

Let me be clear, though. Normativity is certainly strange; there is no doubt about that. At first blush normative properties do not seem to have much in common with any natural properties. But not all reductions are obvious. Prior to the advent of modern chemistry, who would have thought that a liquid (water) could be reduced to two gases (hydrogen and oxygen)? My point is simply that the evident strangeness of normativity does not by itself justify the nonreductive externalists' pessimism about arriving at an illuminating reductive account.

Ultimately, it may be just the dogmatic confidence in normative authority discussed above that underlies externalists' pessimism about the prospects of reductionism. Consider Philip Stratton-Lake's discussion of the open question argument. Stratton-Lake argues that there was room for reduction in the cases of heat (reduced to molecular kinetic energy) and water (reduced to $H_2O$) only because the concepts of heat and water are "metaphysically incomplete and superficial"; neither concept tells us anything about the *nature* of its referent.[92] The concept of goodness, however, "does not seem to be metaphysically superficial or incomplete" in this way, and so no "metaphysically deeper" account is necessary.[93] But this argument is wrong on both counts. First, prior to Lavoisier's discovery the concept of water *did* purport to tell us something about the nature of its referent, namely, that it is one of the basic and irreducible elements of the universe. This was no barrier to Lavoisier's proposed reduction, which simply required revision of the concept. Second, Stratton-Lake suggests that the concept of goodness is completely perspicuous. But what precisely does this concept tell us about the nature of its referent? That it is good? To the extent that it tells us anything informative (say, that goodness seems to have an attractive force or pull), it does so in a reductionist spirit, since this information naturally raises the question of what goodness could be such that it has that kind of pull. Stratton-Lake insists that the concept of goodness "does not merely describe certain properties goodness has" (as the concept of water does); it "aspires to tell us what goodness

92    Stratton Lake, introduction to *Ethical Intuitionism*, 9.

93    Stratton-Lake, introduction to *Ethical Intuitionism*, 10. Compare Parfit, who urges that the concepts of right and wrong, unlike the concepts of heat and water, "do not … I believe, have gaps that are waiting to be filled, in a way that would allow these concepts to refer to one or more natural properties" ("Normativity, Naturalism, and Noncognitivism," 6).

is."[94] But if that is its aspiration it fails miserably, given that all we can do, at least according to externalists like Stratton-Lake, is explicate goodness in terms of other, equally mysterious normative concepts.

It is difficult to avoid the conclusion that behind Stratton-Lake's insistence that normativity is irreducible is just his confidence the normative notions do not *require* or *call for* any explanation. Heat and water are reducible only because the concepts of heat and water are mysterious and incomplete. The concept of goodness, however, is transparent, and it therefore leaves no room for reduction. Clearly, like his fellow externalists Clarke, Scanlon, and Parfit, Stratton-Lake just "isn't worried."[95] Like them, he mistakes epistemological first principles for metaphysical ones.

For at least some externalists, then, it is not the irreducibility of normativity that accounts for their lack of interest in further investigation into the foundations of ethics. Rather, it is the other way around: How, they wonder, could a reductive account be successful when normative concepts like goodness are already "metaphysically complete"? Ultimately, whether what underlies the externalist approach is this unphilosophical lack of concern or instead an unfounded pessimism about the prospects for a successful reduction, this approach has little to offer us as a philosophical theory of the foundations of ethics. This is reason enough to look closely at the internalist alternatives, which I shall do in chapter 3.

But there is more to be said about externalism here, for—as I shall argue in the next section—there are other respects in which the externalist approach simply shrugs off the call for further explanation.

94    Stratton-Lake, introduction to *Ethical Intuitionism*, 10.

95    Elsewhere, Stratton-Lake defends his claim that there is nothing mysterious about the simple property of goodness with an analogy: "no naturalist will claim that the basic elements of the material world are mysterious just because they are simple" (introduction to *The Right and the Good*, xxiv). But they *are* mysterious, in the sense that they call for further explanation. This is precisely what prompted scientists to attempt to understand the nature of oxygen and hydrogen even though they were thought to be basic elements.

      For an especially helpful discussion of what motivates pessimism about reductionism, see Schroeder, "Realism and Reduction."

## Explanatory Impotence

It is accepted by philosophers of nearly all stripes that normative facts (or properties) supervene on descriptive or nonnormative facts (or properties).[96] Unfortunately, there are as many ways of defining supervenience as there are stripes. For my purposes here, the following rough statement of the supervenience claim will suffice: any two agents, actions, or considerations that are identical in all of their nonnormative features must be identical in normative respects as well. In other words, there can be no normative difference without a nonnormative difference.[97]

According to externalists like Parfit and Scanlon, the property of being a reason is distinct from any of the nonnormative properties in virtue of which considerations are reasons. Thus, although I have a reason to donate to Oxford *because* doing so would alleviate suffering, the fact that I have this reason is something over and above the fact that my action would alleviate suffering. This aspect of the externalist approach makes the supervenience of the normative on the nonnormative difficult to explain. If normative authority is just an ineffable feature of norms (or if rightness is just an irreducible feature of actions), then it is not clear why this authority (or rightness) can never vary without a corresponding variance in some descriptive property. Externalists do not seem to have the resources to explain why there cannot be normative differences without nonnormative differences.

Actually, externalists do have one recourse: they can explain supervenience by appealing to the fact that normative truths are *necessary* truths. If, for example, the right-making properties of actions are *necessarily* right-making, then supervenience follows straightaway. There could not be two actions with all of the same nonnormative properties that differ in terms of rightness, since nonnormative properties that are right-making in one case are necessarily right-making (and hence right-making in all cases). But this solves one mystery only by creating another. Now externalists face a new question: Why are normative truths necessary truths? If the rightness of my action cannot be reduced to the fact that it alleviates suffering, why could

---

96    For the one exception with which I am familiar, see Griffin, "Values: Reduction, Supervenience, and Explanation by Ascent" and *Value Judgment*, chap. 3. But see Smith, "Does the Evaluative Supervene on the Natural?" for a persuasive response.

97    For a more thorough and precise statement, see Smith, "Does the Evaluative Supervene on the Natural?" 217.

there not be another action that alleviates suffering (and that is alike in all other nonnormative respects) but that is nevertheless wrong?[98]

Another general feature of normativity externalists cannot explain is normative authority's limited jurisdiction. Not everything in the universe is subject to the authority of reasons. There are, for example, no reasons for volcanoes to erupt; it does not make sense to say of a volcano that it ought to erupt, at least on any normative understanding of the word "ought." Volcanoes, like rocks and electrons (and perhaps flowers and toads), fall outside the jurisdiction of reasons. But why is that? Why does the arm of normative authority not reach as far as volcanoes? Why can the property of rightness not apply to the movements of plants as well as to those of agents? Externalists have no answers to these questions. According to externalism, normative authority is irreducible and ineffable, and so externalists cannot explain why we find rightness in such a determinate pattern, attached only to considerations that stand in certain relations to agents.[99]

These may seem like unreasonable objections. After all, externalists might observe, the necessity of normative truths (and hence the supervenience of the normative on the nonnormative) follows from the very *concept* of normativity. It is, for example, part of the concept of rightness that properties that are right-making in one case will be right-making in every possible case. Were we to meet someone who classified actions as "right" in a way that was inconsistent with the supervenience thesis, we would conclude that this person was not employing our concept of rightness but rather some other concept. And if the supervenience of the normative is conceptually guaranteed, then surely, externalists will continue, it is no failing

---

98    For further discussion of the externalist's claim that normative truths are necessary, see Korsgaard, "Normativity, Necessity, and the Synthetic *A Priori*," 6–8.

       Nonreductive naturalists often attempt a further explanatory move. They try to explain the supervenience of irreducibly normative natural properties upon nonnormative natural properties by claiming that the former are *constituted* by the latter. See, for example, Brink, *Moral Realism and the Foundations of Ethics*, 175–7. Shafer-Landau, who calls himself a nonnaturalist, nevertheless adopts a nearly identical strategy for explaining supervenience (*Moral Realism*, 63–5, 72–8, 84–9).

       The problem with this approach, briefly, is that once these philosophers admit that normative properties are constituted by natural properties, it is difficult to see in what respect their position remains nonreductive. See Darwall, Gibbard, and Railton, "Toward *Fin de siècle* Ethics," 171–2.

99    Existence externalism combined with expressivism (as in Gibbard's view) or sensibility theory (as in McDowell's view) *does* allow for an explanation of these facts about normativity. But neither expressivism nor sensibility theory is an account of the foundations of ethics in the sense that interests me, and so I shall consider neither theory here.

of externalism that it cannot explain this. On the contrary, this is precisely what we should expect of any theory of normativity. And the same goes for my point about the jurisdiction of reasons. It is part of the very idea of a reason for action that reasons can be had only by agents, and so no explanation of the fact that only agents have reasons is need.[100]

This response misconstrues my objection to externalism. Conceptual truths do not amount to metaphysical explanations, and it is metaphysical explanations that the externalists cannot provide. Let me elaborate with an example. Suppose we are considering a theory of the nature of water that cannot explain why water is clear. And suppose I object that this explanatory failing is a strike against the theory: all other things being equal, we should opt for a theory that can explain why water is clear. It would be no response to my objection to say that the clarity of water is conceptually guaranteed and therefore stands in no need of explanation. I am asking our imaginary theorist to explain why water is clear, and he is responding: if it were not clear, then it would not be water! But this is the wrong sort of explanation. Reminding us that our concept of water applies to the stuff in Lake Michigan partly in virtue of the fact that that stuff is clear does not explain why the stuff in Lake Michigan is clear. The water theorist has given us a partial explanation of why our concept of water applies to the stuff in Lake Michigan. He may also have explained why the stuff in Lake Michigan would *not* be water if it were *not* clear. He has *not* explained, however, why the stuff in Lake Michigan is clear, and *that* is what a theory of water should explain.

What about a theory of normativity? The externalist maintains that rightness is not reducible to any nonnormative property. In the words of Ronald Dworkin, rightness "exercises its own sovereignty."[101] But why would a property that exercises its own sovereignty supervene on properties that are outside its domain?[102] Why would rightness always attach to the same nonnormative properties? Conceptual truths cannot answer these metaphysical questions. I am asking the externalist to explain why this property of rightness he has described supervenes, and he is responding: if it did not supervene, then it would not be rightness. As in the case of the water theorist discussed above, this is the wrong sort of explanation. Observing that

---

100    Thanks to Stephen Darwall for pressing me on this point.

101    Dworkin, "Objectivity and Truth," 128.

102    Compare Gibbard, *Thinking How to Live*, 184; and Railton, "What the Non-Cognitivist Helps Us to See," 298 n. 2.

nothing would satisfy our concept of rightness unless it supervened no more explains why the property of rightness supervenes than observing that nothing would satisfy the concept of water unless it was clear explains why water is clear. The conceptual truth that rightness supervenes may explain why our concept of rightness does not apply to properties that fail to supervene. It does not, however, explain why the property of rightness supervenes. Nor does the conceptual truth that normative authority has a limited jurisdiction explain why the property identified by externalists as normative authority has a limited jurisdiction.[103]

Note that I am not suggesting that externalism is inconsistent with these facts about normativity. I am merely arguing that it lacks the resources to explain them.

Internalism, on the other hand, is not plagued by this explanatory impotence. Recall that internalism is most plausible as a reductive theory, and reduction offers a straightforward account of supervenience. If, for example, the property of being a reason just is the property of being a consideration with a certain internal connection to one's motives, then there cannot be considerations that differ with respect to their normativity without also differing with respect to this internal connection. Internalists have a similarly ready answer to questions about the jurisdiction of reasons. According to internalism, normative authority is a function of normative governance, and so the former only ever extends as far as the capacity for the latter. Since volcanoes cannot be normatively governed, they fall outside the jurisdiction of reasons. In other words, because volcanoes cannot erupt for reasons, they cannot have reasons to erupt.

The failure to account for important and interesting aspects of normativity is an unfortunate characteristic of the externalist approach to the foundations of ethics. What is more, externalism's explanatory stinginess contrasts sharply with internalism's munificence, and this counts strongly in internalism's favor. What might count in externalism's favor? I have already considered two possible defenses of externalism—the self-evident nature of basic normative truths and the suspicion that normativity is irreducible—and both have been found wanting. In the next section I shall consider a third possible motivation for the externalist approach.

---

103    I am indebted to Mark Schroeder for helping me formulate the ideas in this paragraph.

## Phenomenology

The central tenet of the externalist approach is that a norm or reason can be authoritative independently of any capacity to move or influence us. I suspect that this view owes much of its initial attractiveness to the sense that when we contemplate a normative property, that property seems to apply to its object merely in virtue of what the object itself is. When we confront an object that matters, for instance, its mattering feels like something we detect—an aspect or feature of the object in question. Put another way, the phenomenology of normative contemplation is akin to the phenomenology of perception, and it is nothing like the phenomenology of desire. Our role in the awareness of normativity seems more like that of a passive observer than that of an active participant.

Many externalists seem to think that the nature of normative experience tells us something about what sort of property normativity is. If the phenomenology of normative experience is one of detection rather than participation, then it seems to follow that considerations are reasons (or actions are right, or things are good) merely in virtue of the nature of the considerations (or actions, or things) themselves. In other words, fundamental normative properties must be *intrinsic properties* rather than *relations* (or *relational properties*). Translated into the language of reasons, this amounts to the thought that normative authority (or reason-givingness) is a property that certain considerations have regardless of what the rest of the world looks like. Considerations are not normative *for a particular agent*; they are normative *simpliciter*. Evidence of this way of thinking about normative authority can be found in the work of a number of externalists. Recall that for Clarke, acts are "*in their own nature* good and reasonable and fit to be done."[104] We find a similar view in Moore, who claims that whether something is good "depends solely on the intrinsic nature of the thing in question," as well as in Ross, who accepts Moore's conception of goodness and defends a similar account of rightness.[105]

The notion that normativity is an intrinsic quality naturally yields the externalist conclusion that there can be reasons for agents who cannot be influenced by those reasons. A consideration's lack of influence is not one of its intrinsic features, and so it does not bear

104    Clarke, *Discourse Concerning the Unchangeable Obligations of Natural Religion*, § 228, emphasis added.

105    Moore, "The Conception of Intrinsic Value," 286; and Ross, *The Right and the Good*, chs. 1 and 4. Compare Mackie, *Ethics*, ch. 1; and Sosa, "Pathetic Ethics," 289.

on the question of whether that consideration is a reason. This view of normativity also has a number of strange consequences, though. Suppose, for example, that we are visited by Martians who lack the concept of suffering (perhaps because they lack the capacity to suffer). Now, if normativity is an intrinsic quality, then the fact that an action would lessen my suffering will be a reason even for these Martians. After all, if it is a reason *simpliciter*, then *a fortiori* it is a reason for them. Although the Martians are unable to register the authority of this consideration, they are subject to it nonetheless. Perhaps an even stranger consequence is that there could be genuinely normative reasons or values even in a universe with no agents at all. These would be reasons or values "for the universe," so to speak.[106]

We can avoid these sorts of conclusions if, following William Frankena, we regard normativity as a relational quality.[107] To say that something ought to be done implies that it ought to be done *by someone*. Similarly, to say that $R$ is a reason to $\phi$ implies that $R$ is a reason *for some agent* to $\phi$.[108] Normativity is not something that certain actions or considerations simply emanate on their own. Rather, it is a relation in which the agent is a crucial participant. This is why facts about what agents are capable of doing have normative significance—why, for example, *ought* implies *can*.[109] A relational account of normativity also to avoids the difficult conclusion that my suffering is a reason even for Martians who lack the concept of suffering.

Internalists, of course, embrace the relational nature of normativity. Since normativity is a relation between a consideration and an agent, it makes sense to try to understand normativity by first understanding the nature of agency and action for a reason. After all, the agent is one of the relata, and so it seems reasonable to expect that his limitations (motivational or otherwise) will play a role in determining whether the relation in question obtains.[110] This is precisely what internalists deliver. They hold that whether a consideration is a reason depends crucially

---

106    Moore presumably has something along these lines in mind when he says that whatever has intrinsic value "ought to exist for its own sake" (*Principia Ethica*, 34).

107    See Frankena, "Obligation and Value in the Ethics of G. E. Moore."

108    John Skorupski contends that *being a reason* is actually a five-place relation. The relata are a consideration, an agent, an act-type, a measure of the strength of the reason, and a time. See Skorupski, "The Ontology of Reasons," 113–5; and Skorupski, "Propositions about Reasons," 27–31.

109    For an internalist argument that proceeds from the premise that *ought* implies *can*, see Velleman, "Is Motivation Internal to Value?"

110    Compare Railton, "Facts and Values," 47–9; and Rosati, "Internalism and the Good for a Person," 313–4.

on whether a particular explanatory relation obtains: $R$ is a reason for $A$ to $\phi$ only if $R$ could explain $A$'s $\phi$-ing.

It would seem, then, that the implausibility of the view that normativity is an intrinsic quality counts in favor of internalism and is a strike against the externalists who embrace this view. To be fair, though, it is not necessarily a strike against externalism in general. Externalists can accept that being a reason is a relational quality. They can, in other words, allow that whether a consideration counts as a reason for an agent $A$ to $\phi$ depends on all sorts of features of $A$, including whether $A$ is physically capable of $\phi$-ing and whether $A$ possesses the relevant concepts, and so forth. (They had better accept this if they believe, as most philosophers do, that "*ought* implies *can*" is a conceptual truth.) Technically, all that externalism must *disallow* is that the relational property of being a reason has anything to do with the *motivational* capacities of the agent (or with the *explanatory* capacities of the reason). In other words, externalists need only reject one particular account of the relational nature of normativity. And it may therefore appear that the claim that normativity is relational poses no threat to their view.

But I would like to suggest otherwise. Once the externalist accepts that reasons for action are relational, he undermines one of the primary motivations for rejecting internalism. To accept that reasons are relational is to acknowledge that we cannot derive conclusions about the nature of normativity simply from the phenomenology of normative experience. And that is one reason fewer to adopt the externalist approach. Moreover, if reasons are not authoritative "all on their own," then there seems to be no metanormative justification for ruling out the possibility that their authority depends on our motivational capacities. Given that $R$'s status as a reason for $A$ to $\phi$ depends on various characteristics of $A$, why exclude $A$'s motivational attributes? That is, if other relations between $A$ and $\phi$ (such as that $A$ *can* $\phi$) are necessary for the normativity of $R$ for $A$, then on what grounds do externalists reject a similar claim on behalf of the explanatory relation identified by internalists? Once externalism has opened the door to the idea that normativity is relational, there seems to be no principled reason to rule out internalism.

Externalists therefore face an unpalatable choice. Their first option is to adopt the view that normativity is an intrinsic quality, in which case they remain true to the phenomenology of normative experience but only at the cost of accepting the implausible consequences mentioned

above. And that cost is almost certain to be too steep, especially given the unreliability of phenomenology. If faithfulness to the phenomenology of experience were worth more, we would be more inclined to accept the naïve views that color, humor, and deliciousness are intrinsic properties of objects. After all, when I look at my new car, I experience the color as an intrinsic property of the car's surface.[111] Similarly, when I bite into a Zingerman's sandwich, I experience the deliciousness as a property purely of the sandwich itself. I certainly do not think, "Wow, this sandwich is disposed to produce the experience of deliciousness in me under normal circumstances." But few philosophers regard color as an intrinsic property, and fewer still regard deliciousness as one.[112] This suggests that we should not place too much emphasis on the character of normative experience. Externalists' second option is to embrace the view the normativity is a relational quality, in which case then they open the door to internalist accounts of that relation. Either way, the phenomenology of normative experience provides little support for externalism.

The arguments up to this point have been directed at some of the considerations to which externalists appeal in order to defend their view. I have tried to show that these defenses fail, and that without them externalism is laid bare as a theory with little or no explanatory potential. None of this demonstrates that externalism is false, however. If externalists were to demonstrate that internalism could not provide an acceptable account of normativity, then externalism—dissatisfyingly obscure though it is—would be our best option.

In the remaining sections I shall argue that externalism is more than merely obscure. The more carefully we scrutinize its depiction of normative authority, the more apparent it becomes how little this picture resembles our general understanding of normativity and reasons for action.

**Flesh and Bones**

In chapter 1 I suggested that the question of whether normative authority requires the capacity to govern is central to the dispute between internalists and externalists. Internalists suppose that an authority that cannot govern really is really no authority at all. A norm that cannot

---

111   I am indebted to my mother for selling me her car and thereby inspiring this example.

112   There are a philosophers are do regard color as an intrinsic property, although it is definitely the minority view. See, for instance, Campbell, "A Simple View of Color."

govern may have all of the trappings of genuine authority, but it exerts no genuine force—no real push or pull. Internalists often appeal to these sorts of considerations for support, but it is not at all clear how much support they can actually provide. In particular, there is the danger that the attractiveness of this appeal derives solely from the political contexts in which the words "authority" and "governance" are often used. Externalists can freely acknowledge the plausibility of the claim that *political authority* requires the capacity to govern. And this does seem quite plausible: a ruler with no ability to enforce his dictates possesses no real political authority, no matter how enlightened he may be. But normative authority is not the same as political authority. Moreover, what is commonly thought to distinguish the two is that the former, unlike the latter, does not depend on any sort of capacity for coercive force. Our challenge, then, is to clarify these competing conceptions of normative authority without simply presupposing a political or second-personal conception of authority.

A consideration can be a reason for me to act—that is, it can justify or count in favor of my action—only in relation to a norm or standard. This is what drives the regress of justification, for a consideration will count as a genuine or unqualified reason for action only if the norm in relation to which it justifies is itself justified. And about any norm we can ask what reason we have to heed or follow it. What we are looking for, then, are norms that do not depend on other norms for justification. In other words, we are looking for norms that are intrinsically justified. As we have seen, internalists and externalists agree that there are norms that fit this description. Most agree, for example, that the norm directing one to avoid suffering does not draw its authority from some other, even more basic norm. The fact that some course of action will cause me to suffer is an unqualified reason not to perform that action.

Internalists explain the special, intrinsic authority of this norm in terms of the central role that considerations of suffering play (or least can play) in practical deliberations. According to internalism, this capacity is just what it is for a norm and its associated reasons to be genuinely authoritative. As we have seen, externalists reject this sort of reductive explanation. For them, what makes one consideration rather than another authoritative for me—and hence an unqualified reason for me to act—is entirely independent of any capacity I may or may not have to respond to this consideration. I am subject to a norm's authority, in the sense that its

dictates apply to me, regardless of whether I can be governed by it. It has its authority all on its own, as it were; its authority is somehow "built into the fabric of the world."[113]

But what would it be for a norm to have authority all on its own? A number of internalists have tried to answer this question. W. D. Falk, for example, describes the externalist position this way:

> They presuppose, not unnaturally, that when someone 'ought' or 'has a duty' he is subject to some manner of demand, made on him without regard to his desires; and they imply that this demand issues essentially from outside the agent: that, whether made by a deity or society or the 'situation' (if this means anything), it has an objective existence of its own depending in no way on anything peculiar to the agent's psychological constitution.[114]

This has struck many as a reasonable characterization. And since most externalists reject the possibility that genuine authority could come either from the commands of a deity or from social norms, we are left the question of where these demands originate? The only possible answer seems to be: "the situation," or the world.

As we saw in the introduction, we are awash in a sea of norms. We are bombarded on all sides by recommendations and commands: Avoid suffering! Love thy neighbor! Keep your elbows off the table! Hate Jews! Go with the flow! Externalism, at least as Falk characterizes it, suggests that amidst the cacophony of these competing imperatives is something different

---

113   Hare, "Nothing Matters," 40. Parfit denies that his realist externalism commits him to any metaphysical or ontological conclusions about "the fabric of the world." He thinks that truths about external reason are necessary truths like mathematic or logical truths. "These two kinds of necessary truth are not part of 'the fabric of the world,' with which we causally interact" ("Normativity, Naturalism and Noncognitivism," 35). Nor, Parfit concludes, are normative truths. As I argued above, though, philosophers of mathematics who deny that mathematical propositions have ontological truth conditions owe us an account of mathematical truth. And they have offered a number of competing theories. Parfit owes us a similar account, but—as we have seen—he is unwilling to provide one. (Compare Scanlon, *What We Owe to Each Other*, 62–63.)

   Of course Parfit might mean to take refuge in a minimalist conception of truth. I do not have sufficient space to consider fully the consequences of such a maneuver here. I shall simply observe that minimalism would seem to leave little room for a genuinely metanormative account of the nature of normativity, since under such a conception it would be difficult to distinguish the claims of such an account from simple, first-order normative claims.

   Minimalism is the route Dworkin favors (although perhaps without realizing it) in "Objectivity and Truth," and as Gibbard (*Thinking How to Live*, ch. 9) observes, it is hard to see how the position that results is different from Gibbard's quasi-realist expressivism. See also Blackburn, "Blackburn Reviews Dworkin"; Zangwill, "Zangwill Reviews Dworkin"; and, for a reply that largely misses the point, Dworkin, "Dworkin Replies to Critics." Finally, for an illuminating general discussion of the impact of minimalism on metaethics, see Dreier, "Meta-ethics and the Problem of Creeping Minimalism."

114   Falk, "'Ought' and Motivation," 32.

and special: a voice that speaks with genuine authority. In addition to the various voices in our lives—our conscience, our friends, our culture, our laws, and so forth—there is also the voice of the world, or at least of the normative part of the world. And the difference between the norm that directs us to avoid suffering and the one that advises us to pursue suffering is that only the former has the backing of this authoritative voice. We may not always be able to hear or recognize this voice, but it is there nonetheless, telling us what we ought to do: not what we ought to do according to this norm or that, this person or that, but rather what we *really* ought to do.

If the externalist approach really does entail something along these lines, then externalism is going to be difficult to swallow. One struggles even to make sense of the idea that an impersonal world can issue commands, make recommendations, demand some courses of action but not others, and so forth. This is undoubtedly the reason why a number of philosophers have found the very idea of external authority unintelligible. Consider the following analogy. Were I to posit the existence of creatures without limbs but that can nonetheless walk, you would (I hope) quite reasonably complain that there is something incoherent about my proposal. It would be no response to such a complaint for me to say, "Ah, but you are not allowing for the possibility of something that walks even though it has no legs. You see, there *is* a sort of walking that does not require limbs!" Yet this is precisely how Parfit replies to critics of externalism who complain that they cannot comprehend the notion of an external authority. After providing his own, internalist account of reasons, for instance, Bernard Williams struggles to come to terms with the externalist's interpretation of the claim that *A* has an external reason to φ. In the end, he throws up his hands and denies that "the sense of external reason statements is in the least clear."[115] Parfit replies that the sense of such statements is perfectly clear: the externalist means that "despite *A*'s motivational state, *A* has a reason to do this thing."[116] But Williams already knows *that*. His complaint is that he does not understand *how A* could have an authoritative reason that is independent of his motivational state in this way. He cannot comprehend what this could possibly involve.

R. M. Hare expresses similar bewilderment when he considers the notion of external normative authority. While discussing the exclamation, "Nothing matters!" Hare quickly

---

115    Williams, "Replies," 215.

116    Parfit, "Normativity," 335.

dismisses the externalist interpretation: As if mattering were "(some activity or process) that things did."[117] Parfit's reply is succinct: "There is, I believe, a use of the word 'matters' which Hare does not understand."[118] In one sense this is clearly true: Hare does not understand the way in which Parfit uses the word 'matters'. But that may not be Hare's fault. Hare is suggesting that, as Parfit uses the word, mattering must be something that objects do, or at least it must be a state in which they can exist all by themselves. And Hare finds this possibility barely intelligible. As Mackie puts it, Hare "cannot frame for himself any clear, detailed picture of what it would be like for [normative authority] to be part of the fabric of the world."[119]

Mackie himself claims to have no such problem. He is sufficiently confident in his understanding of the idea of external authority that he feels comfortable denying that there is any such thing.[120] Mackie goes so far as to say that the idea of "external authority" or "objective, intrinsic prescriptivity" or "authoritative prescriptivity" is an essential ingredient of ordinary, everyday moral judgments.[121] But even still Mackie has difficulty describing just what this authoritative prescriptivity involves. At one point he suggests that authoritative reasons counting against a particular course of action would require that action to have "not-to-be-doneness somehow built into it."[122] But just what is not-to-be-doneness, and how could it be built into the world, or into a certain course of action? That such a notion is an aspect of our everyday normative reasoning does not establish its legitimacy; the ideas of external authority and intrinsic "not-to-be-doneness" would hardly be the first everyday notions to be revealed as nonsense upon close inspection. Perhaps because of such worries, Mackie immediately attempts a more concrete description: a situation counts in favor of an action when that situation has "a demand for such-and-such an action somehow built into it."[123] Now we are

---

117    Hare, "Nothing Matters," 37.

118    Parfit, "Normativity," 326.

119    Mackie, *Ethics*, 24.

120    Mackie thinks that any normativity worth taking seriously must be external, and so he qualifies as an externalist.

121    Mackie, *Ethics*, 34, 35, and 39.

122    Mackie, *Ethics*, 40.

123    Mackie, *Ethics*, 40.

back where we started, for how can we even make sense of a demand without a demander?[124] Of course there is a superficial sense in which these words are intelligible, just as there is a sense in which we can understand my idea of an entity that walks without limbs. The problem in both cases arises when we actually try to come to terms with these notions. The more we try to understand what a demand without a demander is, the more we realize we have no idea what is being proposed or posited.[125]

I have been assuming that Falk's description of externalism in terms of demands and instructions is accurate. But externalists may object that it is more of a caricature than a characterization, and that Falk unfairly ascribes to them a second-personal conception of authority.[126] Even if, as Stephen Darwall has recently suggested, such a conception is crucial to an account of the normativity of *morality* and of peculiarly moral notions such as *obligation* and *accountability*, it may not be necessary for an account of normativity in general.

> There is … an important difference between the idea of an authoritative or binding norm in the familiar sense of a valid ought that entails genuine normative reasons, on the one hand, and that of an authoritative (second-personal) claim or demand on the other.[127]

Thus, although we might need to appeal to the notions of a demand and a demander in order to explain the authority of morality, there is another, "familiar" conception of authority that requires nothing of the sort. If Darwall is correct, then Falk is indeed stacking the deck against externalism by describing authoritative norms in terms of commands, demands, and directions. Normativity, externalists will insist, involves nothing like *that*. All it requires is for some consideration to be a reason in favor of an action, and for it to do so even though it may have no influence on our practical deliberations at all. Put another way, externalism does not posit demands without a demander; it merely claims that there are considerations that count in favor

---

124   Compare Garner, "On the Genuine Queerness of Moral Properties and Facts": "It is hard to believe in objective prescriptivity because it is hard to make sense of a demand without a demander, and hard to find a place for demands or demanders apart from human interests and conventions" (143).

125   Parfit might insist that we are struggling to understand the notion of normative authority only because we have forgotten that normativity is irreducible. Normativity is ineffable, and here we are trying to understand it in terms of commands, demands, and directions. But this cannot be the source of our struggle, for we are not trying to explain normative authority in nonnormative terms. Falk's depiction of external normativity does not do away with the irreducible notion of authority. He and Williams and Hare are merely attempting to apply that notion in ways consistent with their understanding of it.

126   I am indebted to Stephen Darwall for repeatedly pressing me on this point.

127   Darwall, *The Second-Person Standpoint*, 13.

of actions. Surely, the externalists insist, we can conceive of *this* sort of normativity without imagining that the world itself somehow issues commands or instructions in an authoritative voice.

We must tread carefully here, for the externalist may simply be helping himself to the notion of a reason that justifies or counts in favor all on its own. I am not even sure what this sort of justification could be, and I am therefore inclined to conclude, along with James Dreier, that the notion "does not even make much sense."[128] It is difficult to grasp just how a bare consideration can count in favor of anything. Much like demanding or advising, *counting* does not seem like something an impersonal world can do; nor, for that matter, does *favoring*. Counting for or against is something *agents* do, as when I count the fact that you like Mahler in your favor. Again, externalists might simply insist that there *is* a familiar sense in which considerations can stand in favoring or justifying relations to actions all on their own. But my point is that merely insisting that this is an intelligible notion does not make it so. The move from peculiarly second-personal notions like *demands* to more the "familiar" notion of *counts in favor* does not immediately remove the difficulty. It is hard even to imagine how an impersonal world could do *any* of these things.

Perhaps I am being too literal, though. An externalist might reply that "counting" and "favoring" are just metaphors. In the face of demands for an elucidation of their notion of a reason for action, externalists have complied by providing helpful metaphors. But it would be a mistake, their response goes, to infer that "counting" is supposed to be what reasons actually do: no one is suggesting that the universe can somehow do what only agents can do. It is telling, however, that the metaphors to which externalists appeal always involve agential notions like *counting* and *favoring* (or *recommending* and *guiding*, or *binding* and *requiring*). The temptation to describe normativity in these agential terms may suggest that the internalists are correct to seek the sources of normativity in the conditions of agency. It may, in other words, suggest that the fact that some $R$ is a reason for $A$ to $\phi$ has at least something to do with $A$'s capacity to count $R$ in favor of $\phi$-ing—that is, with $A$'s capacity to take $R$ to be a reason to $\phi$. If normativity is anything like counting or recommending, then reasons are not as external to the motives of agents as externalists suppose.

128    Dreier, "Moral Relativism and Moral Nihilism," 242.

Confronted with this sort of pressure, externalists can refuse even metaphorical explanations of normativity. Such a refusal leaves us back at our starting point with the claim that the notion of a reason is primitive and unanalyzable. Every attempt to put flesh on this skeletal externalist account of normativity has ended with an implausible picture of reasons for action, and so externalists are left with nothing more than the bare bones. But if the externalist can say no more about reasons than that are primitive, we are left with a puzzle. How can we interpret their claim that the authority of reasons is independent of any capacity to govern? When they make this assertion about reasons, is there any way for us insure that we have the right idea clearly in mind? They use the terms "reason" and "norm," but there seems to be no way to know whether their theory is a theory about the reasons and norms that are discussed in everyday practical discourse. These are pressing concerns, for if externalists are *not* working with the same ideas, then their view does not have any distinctive (or at least interesting) content. It does not even contradict internalism.

This may explain why internalists have had so much trouble formulating arguments against the externalist approach. It is practically impossible to get a handle on such an underspecified view. How does one criticize a view that is nearly empty of content, except by pointing out that it is nearly empty of content? Parfit, for one, openly acknowledges the opacity of his view. Knowing that he can shed no light on the nature of normative authority, he admits to being a polemicist who devotes nearly all of his energies to attacking his philosophical opponents. "Given my beliefs about normativity," he explains, "I have no alternative."[129] But notice how this shapes the dialectic. Parfit's primary critique of internalist views is always the same: they do away with genuine normativity. Yet he and his fellow externalists refuse to specify in any way what the internalists are doing away with. They suppose that most of us have at least a basic grasp of what they are talking about, but the resistance of so many internalists over years is good evidence that this simply is not the case. Internalists like Falk, Hare, and Korsgaard just do not understand the externalist conception to normative authority.

This suggests that the fears I expressed at the outset of this chapter have been realized: externalists and internalists are actually operating with different concepts of normativity. In the remaining part of this chapter I shall argue that we nevertheless have good reasons to reject

129   Parfit, "Normativity," 332.

the externalist view. Their concept of reasons for action simply cannot answer to our general understanding of those reasons and of the role they play in our practical lives.

## The Normative and the Practical

The externalist concept of normative authority suggests a particular account of normative deliberation. To see this, recall the example of my students' papers from chapter 1. I know that my students are anxiously awaiting their graded essays, which I promised to return to them tomorrow morning, and that I have plenty of time to finishing my grading this evening. But I also know that the Vienna Philharmonic is playing Mahler tonight at Hill Auditorium, and that I would derive a great deal of pleasure were I to attend the concert. I have all of the relevant nonnormative facts at my disposal, and I am left only with the pressing normative question of what ought I to do, or of what I have most reason to do. On the externalist view, when I ask such questions I am wondering which of my two options is ultimately supported by authoritative norms and reasons. I know that some norms and reasons are authoritative and that others are not, and I am simply trying to figure out which are which. One norm permits me to break promises for the sake of greater pleasure, and another does not. According to the externalists, my struggle to figure out what I ought to do is simply an attempt to determine which of these norms is invested with (or justified by other norms and reasons that are invested with) genuine normative authority.

For externalists, then, normative deliberation is a species of theoretical deliberation. To ask what you ought to do or what you have reason to do is to ask whether certain norms have the property of being authoritative. It is, in other (and perhaps less charitable) words, to ask how the normative parts of the world are arranged.

A number of internalists have argued that answers to these sorts of questions cannot settle the question of what you ought to do. P. H. Nowell-Smith, for example, contends that propositions about whether a norm is endowed with genuine authority are descriptive propositions about what is; and, as Hume famously argued, we cannot derive normative conclusions about what one *ought* to do from nonnormative from premises about what *is*.

> I have noticed the right-making characteristic and the rightness; and I feel the emotion of obligation. Does it follow that I ought to do the action towards which I

feel the emotion?... From statements to the effect that these [authoritative norms] exist no conclusions follow about what I *ought to do*."[130]

Thomas Nagel lodges a similar complaint regarding Moore's externalism: it "permits "someone who has acknowledged that he should do something and has seen *why* it is the case that he should do it to ask whether he has any reason for doing it."[131] Comparable criticisms are offered by Richard Brandt and Christine Korsgaard, among others.[132]

Externalists, according to these philosophers, have made normative authority so external that it now seems irrelevant to questions about what we really have reason to do. There may indeed be "special" norms and reasons endowed with genuine authority—considerations whose normativity is somehow built into the fabric of the world—but why, Nowell-Smith wonders, should we care?

> A new world is revealed for our inspection; it is mapped and described in elaborate detail. No doubt it is all very interesting. If I happen to have a thirst for knowledge, I shall read on to satisfy my curiosity, much as I should read about new discoveries in astronomy or geography. Learning about "values" or "duties" might well be as exciting as learning about spiral nebulae or waterspouts. But what if I am not interested? Why should I *do* anything about these newly-revealed objects? Some things, I have now learnt, are right and others wrong; but why should I do what is right and eschew what is wrong?[133]

Nowell-Smith acknowledges that these questions sound absurd. Of course we ought to do what is right. His point is that such questions would not in fact be absurd if our normative terms were used in the way that externalists suppose.

Here we can see the two competing concepts of normativity at work. Nowell-Smith and his fellow internalists have their own concept of what a reason for action is, and this is the concept they deploy when they ask how the properties described by externalists could provide genuine reasons for action. Parfit responds, rightly I think, that this objection does not take the externalist notion of normative authority seriously. "If these acts had the nonnatural property

---

130    Nowell-Smith, *Ethics*, 35–6.

131    Nagel, *The Possibility of Altruism*, 9.

132    See Warnock, *Contemporary Moral Philosophy*, 15–6; Brandt, *A Theory of the Good and the Right*, 155; and Korsgaard, "Skepticism about Practical Reason," 317. Falk objects to externalism on similar grounds, although his argument is marred by his outright conflation of reasons and motives ("'Ought' and Motivation," 28–25). For further discussion of Falk's views about reasons and motives, see Darwall, review of *Ought, Reasons, and Morality*, and "Internalism and Agency," 155–6.

133    Nowell-Smith, *Ethics*, 36.

of being the right thing to do, they would have the nonnatural property of being what we ought to do."[134] And so even on the externalist view, the questions are absurd (as they should be). Parfit could offer a similar reply to Nagel. If there are facts about what we should do that are true in virtue of an external normative authority, then that same authority establishes what we have reason to do. The externalist view is that normative questions just are questions about the normative structure of the world: normative deliberation has the aim of determining which norms and reasons are authoritative. Therefore, norms whose authority is built into the fabric of the world really would settle questions of what we ought to do or what we have reason to do. According to externalists, that is exactly what such questions are about.

Of course, this cannot resolve the matter, since what the internalists are actually questioning is whether the "nonnatural property" to which Parfit appeals really is the property of "being the right thing to do." They doubt that this is the property captured or expressed by the concept of rightness. How do we resolve this dispute?

We can make progress, I think, by recasting the internalist objection in terms of the practical question of *what to do*. If we accept the externalist account of our normative concepts, then facts about external authority may indeed resolve our *normative* questions. They do not, however, settle our *practical* questions. As Nagel puts it, if our normative concepts were "merely classificatory then a conclusion about what one *should* do would by itself have no bearing on a conclusion about what *to do*."[135] Considerations about my reasons for action would be utterly irrelevant to my practical deliberations, unless I happened to care about acting in accordance with the world's distribution of external normative authority. And what Nowell-Smith and Nagel are wondering is why anyone would ever care about that? As Nowell-Smith acknowledges, one may very well have a purely academic or theoretical interest in the distribution of normative authority, much as Nowell-Smith has a theoretical interest in the nature of waterspouts. But why would this merely academic concern play any role at all in one's practical decision making? I shall call this *the practicality objection to externalism*.

134    Parfit, "Normativity," 347. Compare Shafer-Landau, *Moral Realism*, 180; and Frankena, "Obligation and Motivation in Recent Moral Philosophy," 53–5.

135    Nagel, *The Possibility of Altruism*, 109.

The merely contingent practical relevance of externally authoritative norms seems to follow from the externalists' conception of normative reasoning as a branch of theoretical reasoning. As Nowell-Smith observes,

> Moral knowledge is represented by intuitionists as knowledge that a certain object has a certain characteristic. To learn a moral truth is like learning that Henry VIII had six wives or that α Centauri is 4½ light years away.[136]

Nowell-Smith confines his remarks to externalists about morality, but they apply equally well to externalists about reasons for action (and about normativity in general). For such philosophers, normative judgments are just "descriptions of features of the universe"—albeit features of a special kind.[137] When we judge that some consideration is a reason we are merely assigning it to a special category; we are designating it as normatively significant. And this sort of classificatory judgment has no more intrinsic relevance to our practical deliberations than any other classification. Parfit, of course, denies that externalism portrays normative deliberation as nothing more than designation. According to Parfit, even when normative judgments are practically irrelevant, "they could be more than merely classificatory. They could have *normative* force. Perhaps these people *should* act differently."[138] Nowell-Smith's point, however, is that even saying this amounts to nothing more than a classification. To have normative force is merely to fall into a particular category.[139]

Although recasting these challenges in terms of the practical question *what to do* brings us closer to the nub of the issue, the practicality objection is crucially ambiguous. Consider again Nagel's wording: judgments couched in normative concepts that are merely classificatory would "have no bearing on a conclusion about what *to do*." The ambiguity here resides in the phrase "bearing on." Whether one thing bears on another can be either a normative or a descriptive matter. On the one hand, when I ask whether *sui generis* normative properties bear on (or are relevant to) my question of what to do, I may be asking whether my judgments

---

136  Nowell-Smith, *Ethics*, 34. Compare Nagel, *The View from Nowhere*, 146–8. Nagel develops an account of normative inquiry that draws heavily on his account of theoretical inquiry. Although he denies that normative reasoning should be understood on the model of perception of the external world—"The subject matter of our investigations is how to live"—he cannot escape the language of "discovery" and "perception" (148).

137  Nowell-Smith, *Ethics*, 34.

138  Parfit, "Reasons and Motivation," 111–2.

139  Normative judgments may be classifications of a special sort, but then so are judgments about color.

about such properties *should* play a role in my practical deliberations. To say that *x* is relevant to *y* is often to suggest that anyone thinking about *x* *ought* at least to consider *y*. On the other hand, the question of whether externalist considerations bear on my question of what to do may be a question not about whether these considerations should play a role in my practical deliberations, but rather about whether they *actually do* play such a role.

So there are actually two distinct objections here. Both hinge on the claim that externalism cannot do justice to the practical significance of normative judgments—that is, to the intimate connection between normative deliberation and practical deliberation.[140] Where these two objections differ is over the nature of this connection. One takes it to be normative or *de jure*; the other takes it to be descriptive or *de facto*. What is more, each objection threatens externalism: it seems clear that normative deliberation *should* play a central role in practical deliberation and that it *does* play such a role. I shall therefore consider both, beginning with the *de jure* objection.

### The Practical Conception of Ethics

Reasons for action are practical. Yet as we saw in chapter 1, this is a platitude accepted by both internalists and externalists. The *de jure* practicality objection appeals to more than the obvious fact that reasons for action are about actions. It rests on the stronger claim that, as Darwall puts it, questions about reasons for action are "practical in the further sense that they are the questions that confront an agent deliberating about what to do."[141] The normative perspective is inextricably intertwined with the practical perspective, and the normative question of what I have most reason to do (or of what ought I to do) has a special, intrinsic connection to the practical question of what to do. I shall call this *the practical conception of ethics*.[142]

140  I have borrowed the phrase "practical significance" from Mele, *Autonomous Agents*, 18.

141  Darwall, "Reasons, Motives, and the Demands of Morality," 307.

142  Nowell-Smith signals his acceptance of the practical conception of ethics when he treats "What shall I do?" as "the fundamental question of ethics" (*Ethics*, 233). In the postscript to *Ethics*, Nowell-Smith adds: "Moral philosophy is a practical science; its aim is to answer questions in the form 'What shall I do?'" (277–8). Compare Hare, *Moral Thinking*, 32; Darwall, "Reasons, Motives, and the Demands of Morality," 307; and Gibbard, *Thinking How to Live*, 9–10. Gibbard actually claims that the slogan "Ought questions are question of what to do" is "a form of expressivism" ("Knowing What to Do, Seeing What to Do," 212–3). I disagree. What distinguishes expressivism is its oblique approach to the explanation of normativity. One can accept something along the lines of Gibbard's slogan but also pursue a straight-on account of normativity. This is precisely my strategy in chapter 4.

To make this claim more precise, we can borrow a page from one of Pamela Hieronymi's recent papers. According to Hieronymi, we should think of reasons as *considerations that bear on a question*.[143] Moreover, what distinguishes one sort of reason from another are the questions on which they bear. The practical conception of ethics, then, is a view about what makes something a reason *for action*. It holds that reasons for action are considerations that bear on the question of *what to do*. We ask normative questions in order to help us answer our practical questions. When I wonder what I have reason to do, I am looking for considerations that are relevant to my question of what to do. Similarly, when we ask what *you* have reason to do (or what *Caesar* had reason to do), we are searching for considerations that bear on *your* (or *Caesar's*) practical deliberations.[144]

The *de jure* practicality objection is that the externalist approach to the foundations of ethics is inconsistent with (or at least in serious tension with) the practical conception of ethics. Why would judgments about a property of the sort externalists have described be intrinsically relevant to my deliberations about what to do? Externalists contend that normative authority is an external and *sui generis* property that attaches to norms (and hence to reasons). Some norms are authoritative, others are not, and this is a brute fact about the world. But if that is all authority is—if normativity could be so utterly disconnected from the sort of person I am and the sorts of things I care about—why should considerations of this authority play any role in my practical decision making? Even if there are objects or actions that simply matter in some *sui generis* sense, as the externalists suppose, why should they matter *to me*?

---

143    Hieronymi, "The Wrong Kind of Reason." Compare Shah, "How Truth Governs Belief," 466–7.

144    We could cast the practical conception of ethics in terms of a very weak version of judgment internalism: if *A* judges that *A* has most reason to $\phi$ (or that *A* ought to $\phi$), then *A* normally has defeasible motivation to $\phi$ insofar as *A* is rational. The rationality proviso is necessary to capture the *de jure* nature of this version of the practicality objection. This would be a version of what Shafer-Landau calls "hybrid judgment internalism" (*Moral Realism*, 144), and it is defended by R. Jay Wallace in "Moral Motivation." (For the distinction between different sorts of internalism, see note 7 in chapter 1 above.)

    I am reluctant to apply the "judgment internalism" label, though, since debates about judgment internalism invariably collapse into debates about whether a certain sort of agent (the amoralist or the anormativist) is possible. The idea is that if we can find just one rational agent who sincerely uses normative language without having any of the required motives, we shall have shown that normative deliberation does not have the practical weight it is supposed to have. (See, for example, Smith, "The Moral Problem, ch. 3; and Shafer-Landau, *Moral Realism*, ch. 6.) Even Dreier's "modest internalism" is too strong. (Dreier, "Internalism and Speaker Relativism," 9–14.) The practical conception of ethics is meant to capture the idea of a much more global connection between normative deliberation and practical deliberation, one that may not be preserved at all times or in every agent.

In a way, the *de jure* practicality objection is nothing more than the open question argument turned back against externalism. The property that externalists call "being a reason" cannot be the normative property of being a reason, the objection goes, since it is an open question whether considerations that have this external property *really are reasons*. Put differently, the externalist property of mattering just does not seem to matter all that much, and so externalists must have missed the normativity they were trying to capture.

I can think of three externalist responses to the *de jure* objection. First, externalists might simply reject the practical conception of ethics and the supposed connection between normative and practical questions. That is, they might just deny that ethics is practical in the sense proposed here.

This is a rather extreme response. Rejecting the practical conception of ethics involves more than merely saying, as Parfit does, that the fundamental question of ethics is not "What *shall* I do?" but rather "What *should* I do?" It also involves suggesting that the latter has nothing intrinsically to do with the former—that by itself the fact that I ought to jump does not bear at all on the question of whether to jump. Of course this fact *in conjunction with a commitment to do what I ought to do* will be relevant to my decision, but the practical conception of ethics posits a tighter connection than that. And so according to this extreme response, normative questions are relevant to practical ones only insofar as one has a practical commitment to act in accordance with our reasons.

This may actually be Parfit's conception of ethics. If it is, then his reasons for adopting it are weak. Parfit argues that were ethicists to address the question of what to do, they would be in the business of making our decisions for us. "If moral philosophy had the aim of answering such a question," Parfit contends, "it could not possibly succeed. Philosophy cannot make our decisions."[145] But this is just a caricature of what I have been calling the practical conception of ethics. The central idea is not that ethicists or ethical considerations must somehow make our decisions for us. It is, rather, that ethical or normative considerations are ingredients out of which *we* make our decisions. The practical conception of ethics does not imply that moral philosophers do the hard work of decision making; it proposes that what moral philosophers tell us should inform *our* decision making. And surely it should. The central question of ethics

---

145    Parfit, "Normativity," 333.

may indeed be "What should I do?" but this is not an idle question. We ask it when we are trying to figure out what to do.[146]

A second and more plausible externalist response acknowledges that normative questions bear on practical questions and emphasizes the normative nature of this connection. Thus, an externalist might observe that the relevance of the normative for the practical is itself a normative matter and can therefore be secured by appealing to an external normative fact: *A* is *rational* only if *A*'s normative judgments about what *A* ought to do influence *A*'s practical judgment about what to do. However, this sort of appeal cannot answer the practicality objection. That objection asks why what matters in some externalist sense should matter in our more familiar sense, and the externalist has answered merely by drawing our attention to something else that matters in his externalist sense. In other words, the externalist's response will carry weight only if we have already decided that what matters in his sense *really matters*. But that is precisely what is in doubt. If we are unsure as to whether we should bother to act as the world's distribution of normative properties dictates, identifying another instance of one of those properties is hardly going to help.

This leads naturally to the third and most difficult externalist response to the *de jure* practicality objection, namely, that the concept of a reason for action (or of practical rationality) *just is* the concept of something that bears on practical deliberation. That is, the concept of what matters *just is* the concept of something that should matter to me. And so the practicality objection seems to be based on some sort of conceptual confusion. To suppose that external reasons do *not* bear the question of what to do is to suppose that something which *by definition* bears on practical questions does not bear on such questions, and of course that is a contradiction. It therefore seems that the very concept of a reason for action guarantees that externalism will be consistent with the practical conception of ethics.[147]

This line of argument is similar to the one mentioned above as a possible response to the charge of explanatory impotence. Recall that externalists might attempt to justify their theory's failure to explain the supervenience of the normative on the descriptive and the limited

---

146   It is important to note that the practical conception of ethics is not inconsistent with the claim that *other* normative questions might be idle. When I ask "What should I have done?" or "What should Caesar have done?" my questions cannot have any immediate practical relevance, since the practical questions upon which they bear are in the past. I discuss the idleness of such questions in chapter 4 below.

147   See Scanlon, *What We Owe to Each Other*, 60–1, for an example of this sort of response.

jurisdiction of reason by claiming that these are conceptual truths. If it is analytically true that normative properties supervene, they might argue, then there is nothing really to explain. My reply above was there are also *metaphysical* questions about why *those* properties—which externalists claim to be normative—supervene on more familiar descriptive properties. Merely answering with a conceptual truth does not address the underlying metaphysical question. The third externalist response to the practicality objection misses the point in a similar way. Externalists respond to the charge that their view is inconsistent with the practical conception of ethics by pointing out that any theory which was inconsistent with this conception would not be a theory of normativity. But internalists agree that the practicality of reasons is conceptually guaranteed: the concept of a reason for action is the concept of a consideration that bears on practical questions. The thrust of the practicality objection is that the reasons externalism posits do not satisfy this concept! Put more precisely, the irreducible property identified by externalists as the property of being a reason does not seem like the sort of thing that ought to play a role in practical deliberations, at least not to someone who has yet to adopt the externalist perspective. Consequently, this *sui generis*, putatively normative property does not seem to be a normative property at all. Given doubts about whether external authority is genuinely relevant to practical deliberation, externalists must *earn* the right to apply normative concepts to the relations and properties they posit. Merely observing that there are irreducible properties and then christening these properties "rightness" and "goodness" does not show that the properties in question *are* rightness and goodness.[148]

Externalists might reply that their critics are begging the question by assuming that the internalist concept of reasons for action is the one that best matches our ordinary concept. In that case the *de jure* practicality objection would amount to no more than the claim that the externalist concept of normativity is inconsistent with the internalist one, and we already knew *that*. I do not think that the practical objection presupposes internalism in this way, though. The practical conception of ethics is not simply a restatement of internalism. Instead it is a crucial aspect of our everyday concept of reasons for action. And the more we realize what the externalist picture of normative authority actually looks like, the more we begin to wonder how anything in that picture could serve as the property of being a genuinely practical reasons.

148    Compare Wallace, "Normativity and the Will," 75–6; and Copp, "The Ring of Gyges," 95–96.

I say this is something that *we* begin to wonder, but I do not pretend to have provided an argument that will convince entrenched externalists to decamp for the internalist lines. If your concept of being a reason is the externalist one, then you will undoubtedly continue to believe that there is an externalist story to tell about the practical significance of normative deliberation. The *de jure* practicality objection is not intended to refute this story. Rather, its purpose is to show that someone not already ensconced in the externalist worldview has a good reason to reject that approach to the foundations of ethics.

## The Desire to Do as We Ought

Recall that there is a second, *de facto* version of the practicality objection. This version revolves around the observation that normative considerations play a central role in our actual practical deliberations. Whether such considerations are ultimately relevant or not, we all seem to *take* or *treat* them to be relevant. Even if there are agents for whom "ought" judgments are practically irrelevant, such agents are notably few and far between.[149] This poses a problem for the externalist approach. If normativity is external and has no essential connection to our motives and commitments, it is hard to see why normative considerations are so universally influential. If the fact that something is a reason is just the fact that it possesses some irreducible property, why do our judgments about reasons exert such a strong and consistent pull? If normative authority does not require governance, why do we care so much about being governed by our reasons? Put simply, why does normativity matter so much to us?

Perhaps the most tempting response to these questions is to explain why we care about our reasons by appealing to the fact (discussed above) that we *should* care. I have already attempted to demonstrate that externalists have trouble accounting for that fact as well. But even if we grant that we should give external reasons weight in our practical deliberations, this cannot explain why we do in fact give them this weight. The claim that we ought to give them weight (or that it is rational to do so) is just another external normative consideration, and it would explain why we actually give them weight only if we already had a general understanding

---

149  This version of the practicality objection also relies on a weak form of judgment internalism: if *A* judges that *A* has most reason to $\phi$ (or that *A* ought to $\phi$), then *A* usually has defeasible motivation to $\phi$. No rationality proviso is necessary, since we are dealing with a descriptive rather than a normative claim. For reasons sketched above in note 144 I am reluctant to use the term "judgment internalism."

of why external normative considerations are influential. Absent such an understanding, the appeal to claims about rationality only pushes the need for an explanation back a step.

A different answer to the *de facto* practicality objection is provided by externalist Russ Shafer-Landau. According to Shafer-Landau, it is a conceptual truth that normative judgments are intrinsically motivating.

> The near-perfect alignment of [normative] judgment and (some degree of) motivation can be explained by the intrinsic, prima facie motivational power of [normative] judgment.[150]

The idea here is that normative beliefs can have practical consequences *even in the absence of any related desire*. The mere belief that one ought to $\phi$ can generate defeasible motivation to $\phi$. One possible problem with this approach is that it contradicts what is commonly called the Humean theory of motivation—namely, the view that only desires or conative states are intrinsically motivating. But I do not want my critique of externalism to rest on that controversial theory.[151] A second and more serious problem mirrors one we have seen in this chapter a couple of times before. The *conceptual* truth that normative judgments are intrinsically motivating cannot explain the *metaphysical* fact that our normative judgments consistently have practical consequences. Put another way, the fact that judgments about properties of the sort externalists envision can be classified as normative judgments does not explain why those judgments are so often motivating. Any explanation involving this conceptual truth would have to proceed the other way around: judgments about external reasons count as normative at least partly because those judgments exert such a consistent pull.

The explanation issue is more complicated this time around, though, since there *are* genuinely metaphysical (and biological) explanations to which externalists could appeal in order to establish that normative judgments are intrinsically motivating. For instance, they might adopt a theory of the evolution of our normative governance system along the lines of the one proposed by Allan Gibbard in *Wise Choices, Apt Feelings*. There Gibbard speculates that the practical force of normative judgments can be explained by their evolutionary origin,

---

150 Shafer-Landau, *Moral Realism*, 157. Shafer-Landau is actually writing about *moral* judgments, but any version of moral judgment internalism is surely more plausible as a version of normative judgment internalism. T. M. Scanlon defends a similar view of normative judgment in *What We Owe to Each Other* (37–41), as does Jonathan Dancy in *Practical Reality*, ch. 4.

151 See note 26 in chapter 1 above.

and in particular by their crucial role in the coordination of actions and intentions.[152] I do not want to delve any deeper into the details of Gibbard's theory, since I think that the dangers for externalists trying to appropriate this kind of account are clear. Even if externalists accept such an explanation of the practicality of our normative judgments, they face the seemingly insurmountable task of squaring this story with their claim that our normative judgments are *about* external properties that have no essential tie to our motives and intentions (and thus to any interest me might have in coordinating our actions and intentions). I do not see how concepts developed for the purposes Gibbard outlines could ever come to be about the properties posited by externalists, except perhaps by sheer coincidence. After all, those properties are *sui generis* and irreducible.[153] They therefore have no features to which we could appeal in order to explain how judgments about them might play a role in the coordination of action. A Gibbard-style evolutionary theory of the practicality of normative judgments is much more conducive to an internalist account of the content of those judgments, however, since such an account leaves plenty of room for our motives and commitments to play a prominent role.[154]

This leads me to conclude that the intrinsically motivating nature of normative judgments cannot help externalists rebut the *de facto* practicality objection.

The most common externalist response to this objection is to contend that the close connection between normative and practical deliberation is not as close as it seems: normative judgments are only *contingently* motivating. According to this line of argument, we are consistently moved to act according to our reasons not because there is something special about the nature or content of normative judgments, but rather because nearly of all of us have a standing desire to act in accordance with our reasons. That is, we generally desire to do what we

---

152   See Gibbard, *Wise Choices, Apt Feelings*, chs. 4 and 7.

153   This is probably what led rational intuitionists like Clarke and Balguy to adopt their strange perceptual internalism. They needed to account for the practicality of normative deliberation somehow, but they could not appeal to any specific features of normative properties, since they held that such properties are *sui generis* and hence irreducible. Consequently, they settled on the frankly mysterious view that the awareness of genuine normativity is intrinsically motivating. This chalks the practicality of normative knowledge up to the nature of normative properties themselves. That is, it is normativity itself that is somehow intrinsically motivating on their view. This would provide a response to both versions of the practicality objection, but only at the huge cost of perceptual internalism's terrible implausibility. This view is, as I noted above (in agreement with McDowell), impossible to take seriously.

154   For more on why existence internalism fits well with the broad sort of judgment internalism I am defending, see Smith, *The Moral Problem*, 62, 177–81.

ought to do. W. D. Ross is perhaps the most famous adherent of this view. In the midst of his discussion of the psychology of moral action, he notes:

> When I ask myself why I do my duty (*when* I do it, and do it conscientiously), the truest answer I can find is that I do it because, then at least, I desire to do my duty more than I desire anything else.[155]

Many other externalists have followed Ross in arguing that the practical upshot of normative deliberation can be explained by appeal to a nearly universal desire to act as one ought.[156]

There are, I think, two significant problems with this response to the practicality objection. First, explaining the practicality of normative deliberation by appealing to a desire does not do justice to the phenomenon of clear-eyed akrasia or weakness of the will.[157] To see why, return once again to the example of my decision whether to mark my students' essays or to attend the Vienna Philharmonic performance. And suppose I decide that I have most reason to skip the concert—that I ought to work through the evening so I can keep my promise to

---

155   Ross, *The Foundations of Ethics*, 206. Compare Ross, *The Right and the Good*, ch. 7.

156   See, for instance, Prichard, "Duty and Interest," 38; Parfit, "Reasons and Motivation," 113–5; Shafer-Landau, *Moral Realism*, ch. 6; Stratton-Lake, "Why Externalism Is Not a Problem for Ethical Intuitionists," 88–90; and Brink, *Moral Realism and the Foundations of Ethics*, ch. 3. Compare Raz, "The Amoralist"; and for a particularly illuminating discussion, Svavarsdóttir, "Moral Cognitivism and Motivation" and "How Do Moral Judgments Motivate?" A complication arising with respect to most of these authors is that their arguments concern the practical upshot of *moral* judgments rather than of normative judgments more broadly. That said, as I suggested in note 150 above, I think that the plausibility of internalism about moral judgments rests on the plausibility of internalism about normative judgments (together with the view that moral judgments express normative requirements). It is, in other words, the normativity of morality that accounts for morality's "magnetism." (The term "magnetism" comes from Stevenson, "The Emotive Meaning of Ethical Terms," 13.)

It is important to note that this response does not entail the implausible claim that normative considerations figure in our practical deliberations via a conscious awareness of our desire to do what we ought to do. This desire need not be present in the foreground of practical deliberation in order to exert its usual force. See Pettit and Smith, "Backgrounding Desire," for more on how desires can exert their usual influence even from the background of deliberation.

157   Not all philosophers accept the possibility of clear-eyed akrasia. In his most recent book, Gibbard defends the view that normative judgments just are practical judgments, that is, that ought questions are simply questions about what to do. (See Gibbard, *Thinking How to Live*, chs. 1 and 7.) If the normative question of what I ought to do is identical to the practical question of what to do, then it is impossible for me to give one answer to the former question but a different answer to the latter one. And thus it is impossible for me to act against my better judgment: autonomous actions are necessarily a reflection of my better judgment.

Externalists might be tempted to sidestep my objection by accepting Gibbard's equation of normative and practical judgments. After all, if there is no such thing as clear-eyed akrasia, then they cannot be criticized for failing to account for it. However, this maneuver cannot help externalists like Parfit and Scanlon, since, as Parfit acknowledges, judgments about *external* reasons are *not* themselves judgments about what to do.

my students. But now imagine that I go to the concert nevertheless. Despite my considered judgment about what I ought to do, I make a fully autonomous decision to hear a Mahler symphony instead. Most philosophers agree that such cases of clear-eyed akrasia are possible, and externalists who explain the practicality of normative judgments with an appeal to the desire to do as we ought have a straightforward story to tell. In this case, my desire to do what I ought to do is simply not strong enough to overwhelm my longing for Mahler.

The problem with this final externalist response to the practicality objection, then, is not that it cannot account for cases of clear-eyed akrasia, but rather that it cannot explain why such cases are so philosophically and psychologically significant. The externalist explanation of such cases is *too* straightforward; it actually makes them mundane. If we explain the practical significance of normative judgments by appealing to a desire to act as we ought, then akrasia is just a matter of one desire being stronger than another. Philosophically, there is nothing remarkable about akrasia on this view, nothing that would make it worthy of the special philosophical attention it has received over the millennia. Nor, for that matter, is there anything that would make it worthy of being called weakness *of the will*. What is weak in akratic cases is not the *will*, at least according to the externalist response we are considering, but simply one of our desires. This also makes it difficult to explain why akrasia is usually accompanied by psychological consequences: a sense of frustration, tension, or weakness. After all, nearly all of our nonakratic decisions involve stronger desires trumping weaker ones, yet none of these decisions involve the psychological struggle that characterizes akrasia.[158]

Darwall expresses a related worry about the externalist approach when he claims that its account of the practicality of normative judgments does not comport with the central role those judgments play in our practical lives.

> If something's being a reason is simply a … property of it of which we take notice in judging the consideration to be a reason, then the desire to act for reasons is in no sense integral to the self. It is a fascination with a … property that one may have or lack without any change in the self. So understood, the desire to act for reasons is not in itself intelligible. We cannot see it as essential to us.[159]

---

158    Externalists cannot appeal to the irrationality of weakness of will in order to explain the psychological significant of akrasia, since—as I argued above—this is just another externalist normative consideration with psychological and practical consequences that need to be explained.

159    Darwall, *Impartial Reason*, 57.

Externalists might reply that the desire to do what we ought to do *is* integral to the self. Perhaps it is a higher-order desire or an essential component of agency itself. Perhaps the desire to act as we ought is actually *constitutive* of agency.[160] But this leads naturally to the second problem with the final externalist response to the practicality objection: externalists cannot explain why the desire to do as we ought is so commonplace. Where does the nearly universal commitment to do what we ought to do come from? Authoritative norms earn their special status independently of any connection they have to our practical concerns. Why then do we care so much about this status? Given that what we ought to do is a function of the world's distribution of external normative authority, it is hard to fathom why so many of us would have that commitment. The above suggestions—that the desire to act as we ought is a higher-order desire or constitutive of agency—only make the need for an explanation more pressing. They invite the same objections that sapped the explanatory potential of the claim that normative judgments are intrinsically motivating. Any compelling story about how we might have evolved a desire to act as we ought—or about why this desire is an essential component of agency—is going to be in serious tension with the externalist view that our normative judgments are about a *sui generis* external property.

If these arguments are correct, then the externalist approach to the foundations of ethics founders on the practicality of normative deliberation and judgment.[161] Where does this leave us? As I suggested above, Parfit is correct that externalists and internalists operate with different and competing concepts of normative authority. I have tried to show that the externalist concept is not the one most of us employ when we are making everyday normative judgments. The problem with the externalist concept of normativity is that whatever property it picks out is not what we are thinking about when we are thinking about reasons. The sort

---

160  David Velleman discusses the idea that the motive to act as we ought is constitutive of agency in "The Possibility of Practical Reason," 176–8.

161  At this point we might reasonably wonder why so many philosophers are attracted to a view that misses its mark in this way. Here I am inclined to defer to G. E. M. Anscombe, who famously argued that the notion of external (moral) authority is a holdover from the now defunct philosophical framework that made this notion intelligible: the divine law conception of ethics. The idea of an external reason derives its force from such a theory, since it presupposes an authority that is entirely independent of and external to us. Severed from the divine law conception, however, the externalist idea has "no reasonable sense" ("Modern Moral Philosophy," 8). Of course Anscombe's conclusion was even more severe: she claimed that her criticisms applied to all normative notions of *ought*, *rightness*, and *obligation*, external or otherwise. In chapters 3 and 4 below we shall see whether internalists have any more success making "reasonable sense" of these notions.

of authority envisioned by externalists—whether or not it matters in some external sense—is not the sort that matters to us. And so externalism cannot be the best account of normative authority and reasons for action.

At one point Parfit, the archexternalist, considers this possibility. He often refers to what is normative "in my sense" or to normativity "as I conceive it," and he acknowledges the possibility that what is a reason in his sense may not reflect what is a reason in everyone else's sense.[162] He makes the further point that even if his theory is not a theory of normativity as it is ordinarily understood, it may identify an important and interesting aspect of the world that our everyday notion of normativity simply ignores. This is a surprising concession, for if his theory turns out not to be an account of normativity as it is ordinarily understood, then his arguments against internalism collapse. Recall that Parfit's primary criticism of internalist views is that they do away with the normativity. But if it turns out that this merely amounts to the charge that they do away with normativity *as he conceives it*, then internalists can plead guilty with impunity. They will be perfectly content to offer accounts "merely" of normativity *as almost everyone else conceives it*.[163] The upshot of all this, of course, is that if externalism is not a theory of normativity as it is normally understood, then it cannot provide an answer to the questions we have set out to investigate: it cannot establish the foundations of ethics, at least as ethics is ordinarily understood.

Admittedly, most externalists are unwilling to make this concession. Most believe that they *are* offering an account of our shared concept of reasons for action. Against these externalists, I have attempted to demonstrate that to the extent we can actually grasp their notion of normative authority, that notion is in significant tension with our shared understanding of normativity.

Notice that for the most part my objections to externalism are neither ontological nor epistemological. Let me grant that there is an ineffable, even nonnatural property of being authoritative. Let me also grant that we have access to this property through a special faculty of intuition. The point of my arguments is that even if this strange sort of authority exists, it is not

---

162     See, for example, Parfit, "Normativity," 354, and "Normativity, Naturalism, and Noncognitivism," 19, 25.

163     Compare Parfit's criticism of Brandt in "Normativity, Naturalism, and Noncognitivism," 26. Here I am drawing on the debate about Parfit's views at PEA Soup, a collaborative blog devoted to the discussion of philosophical ethics. This particular entry, "Parfit on Normative Irreducibility," can be found online at http://peasoup.typepad.com/peasoup/2007/07/parfit-on-norma.html.

*normativity*. It is not the sort of authority that underlies our unqualified reasons and that can therefore put a principled stop to the regress of justification. It is not, in other words, the kind of authority that can provide the metaphysical foundations of ethics. For that, we shall have to look elsewhere.

I do not mean to suggest, however, that our common understanding of reasons for action bears no traces of the externalist viewpoint. On the contrary, the externalists' emphasis on independent authority captures an absolutely crucial aspect of normativity—one that internalists have difficulty accommodating, as we shall see in chapter 3. Where externalism goes wrong is by emphasizing this authority to exclusion of all else, and in particular to the exclusion of the practical significance of reasons for action. Internalists do not make this mistake; they assign a central role to the practicality of reasons. It remains to be seen, however, whether this sort of account leaves room for genuine authority. If it does not, then internalists shall find themselves similarly unable to provide an account of the foundations of ethics.

# CHAPTER 3

## INTERNALISM

> On the one hand you admit that men cannot be guided through their intelligence, they are ruled by their passions and their instinctual demands. But on the other hand you propose to replace the affective basis of their obedience to civilization by a rational one. Let who can understand this. To me it seems that it must be either one thing or the other.
>
> —Sigmund Freud, *The Future of an Illusion*

The regress of justification presents a challenge tailor-made for reductionism. The assigned task is to find norms that somehow resist the demand for justification—norms that are immune to those seemingly incessant "Why?" questions. Perhaps the most promising strategy for completing this task is to provide a reductive account of justification itself. Such an account can explain why the request for further justification is inappropriate: justification naturally comes to an end when we arrive at the property that *just is* the property of being justified. It would no more make sense to ask why things with this property are justified than it would make sense to ask why water is $H_2O$. "That is just what water is!" would be an acceptable reply to the latter question; similarly, "That is just what justification is!" would be a conclusive reply to the former.

But how do we begin to develop an account of the nature of justification or normativity itself? Peter Railton has a proposal: "Justification has to start somewhere, and if it is to be justification *for us*, it had better start where we are."[1] This is the intuition that animates internalism. Unlike the externalists discussed in chapter 2, internalists take seriously the metaphors we use to describe normativity. We speak of the *force* of reasons and the *binding* nature of obligation. Yet only something *internal* to an agent—a motive or faculty—can bind him to a particular course of action. Normativity must resonate from within the practical

---

1    Railton, "On the Hypothetical and Non-hypothetical," 314.

perspective: only norms with the capacity to govern us can have genuine authority over us. With these thoughts in mind, internalists set about their studies of the nature of normative authority by investigating the phenomenon of normative guidance or action for a reason. Their ultimate goal is to develop a theory that locates the foundations of normativity in the nature of agency.

Of course, the outline of any theory that results from such an investigation will be shaped largely by the conception of agency at work in the background. Following Stephen Darwall, we can therefore divide internalists into two broad camps: *empirical naturalist internalism* and *autonomist internalism*.[2] Empirical naturalist internalists operate with an account of agency that accords special significance to our motives and to motivational force. Consequently, the theories of normative authority these internalists develop treat normative force as a kind of motivational force. Darwall traces the empirical naturalist strand of internalism as far back as Richard Cumberland's *A Treatise of the Laws of Nature*, published in 1672.[3] More recent members of this tradition include Bernard Williams, Richard Brandt, and Gilbert Harman.[4] Autonomist internalists, as their name suggests, bring to their study of normativity an account of agency that focuses on our status as free and autonomous agents. And so these internalists end up identifying the force of reasons with the psychological force of motives resulting from autonomous practical reasoning. This tradition stretches at least as far back as Bishop Butler, although its most famous exponent is undoubtedly Immanuel Kant.[5] Kant's autonomist descendents include Darwall and Christine Korsgaard, and perhaps W. D. Falk and William Frankena as well.[6]

---

2   See Darwall, *The British Moralists and the Internal 'Ought'*, 14–20. Throughout this chapter I draw heavily on Darwall's writings on internalism.

3   See Darwall, "Internalism and Agency," 163–5; and *The British Moralists and the Internal 'Ought'*, ch. 4.

4   See Williams, "Internal and External Reasons," "Internal Reasons and the Obscurity of Blame," "Replies," and "Postscript"; Brandt, *A Theory of the Good and the Right*, pt. 1; and Harman, *The Nature of Morality*, 125–6.

5   See Darwall, *The British Moralists and the Internal 'Ought'*, chs. 9 and 11, as well as "Internalism and Agency," 166–8. Darwall finds even earlier traces of autonomist internalism in the work of Ralph Cudworth and John Locke. See *The British Moralists and the Internal 'Ought'*, chs. 5 and 6.

6   See Darwall, "Reply to Terzis," "Autonomist Internalism and the Justification of Morals," and "Internalism and Agency"; Korsgaard, *The Sources of Normativity*, "The Normativity of Instrumental Reason," "Internalism and the Sources of Normativity," and "Realism and Constructivism in Twentieth-Century Moral Philosophy"; Falk, "Action-Guiding Reasons" and "Morality, Self, and Others"; and Frankena, "On

In this chapter I shall discuss both empirical naturalist internalism and autonomist internalism at length. Rather than attempt to assess each strand as a whole, though, I shall discuss what I consider to be representative figures from the contemporary wings of both traditions: Williams for the empirical naturalists and Korsgaard for the autonomists.[7] I shall begin with empirical naturalist internalism and its champion, Williams.

## Williams Redux

Recall Williams' statement of his internalism: *A* has a reason to $\phi$ only if "*A* could reach the conclusion that he should $\phi$ (or a conclusion to $\phi$) by a sound deliberative route from the motivations that he has in his actual motivational set—that is, the set of his desires, evaluations, attitudes, projects, and so on."[8]

Before I begin to weigh the merits of Williams' theory, I need to review a pair of modifications I proposed in chapter 1. First, recall that Williams defends internalism only as a theory about a *necessary* condition for an agent to have a reason. And as I observed, a theory about only a necessary condition cannot provide a complete account of the foundations of ethics, since it cannot explain what having a reason itself consists in. Williams believes, however, that his internalism actually states a necessary *and sufficient* condition.[9] Whether his arguments support such an account or not, Williams clearly takes himself to have provided a complete theory of reasons for action, and that is the sort of theory I am interested in evaluating here. And so even though he never explicitly defends the sufficiency claim, I shall treat his internalism as though it includes such a claim. Second, recall that Williams presents his view so that it describes only what it is for an agent to have a reason and not what it is for *some consideration to be that reason*. I therefore proposed the following reformulation: *R*

---

Saying the Ethical Thing" and *Thinking about Morality*. For the view that Frankena was (or at least ended up as) an autonomist internalist, see Darwall, "Remembering Frankena, 698–705.

7    Of course, I do not want my criticisms to rest on the idiosyncrasies of their individual views, and so I shall discuss the work of other philosophers as the need arises. Even so, I am aware that it may seem unfair to draw conclusions about internalism in general based on my evaluation of such a small sample. This is a legitimate concern. Although I do regard Williams and Korsgaard as representative of their respective traditions, and although I shall level what I take to be serious criticisms at their respective theories, I do not want to suggest that either the empirical naturalist or the autonomist approach is hopeless. The account of normativity I defend below in chapter 4 draws heavily on both traditions.

8    Williams, "Internal Reasons the Obscurity of Blame," 35.

9    See Williams, "Internal Reasons and the Obscurity of Blame," 35–6.

being a reason for *A* to *φ* consists in the fact that *A*'s sound deliberation on *R* and his actual motivational set could defeasibly motivate *A* to *φ*. This is the interpretation of Williams' internalism I shall consider.

Most of Williams' critics have focused on the arguments he provides in defense of his view.[10] I have already discussed those arguments at length, and so I shall not rehash them here. Rather, my goal is to determine the extent to which Williams' internalism can supply a plausible account of the nature of practical normativity.

We can start with the aspect of normativity that derailed externalism in chapter 2: the practicality of reasons for action. At first glance, Williams does not seem to face any serious problems on this front. He proposes that a consideration can be a reason for me to act only if it can find purchase in the motives I already have. And if it can do that, its practical relevance appears to be undeniable. According to Williams, to deliberate about my reasons just is to deliberate about considerations that are somehow relevant to my concerns and commitments. Hence, there is no mystery at all about why that sort of deliberation would play such a prominent role in practical decision making.

Yet reasons are supposed to be more than merely practical. That some consideration has practical consequences does not establish that it is a reason for me to act. A consideration's practical relevance may account for the role it plays in governing my actions. But even if governance is a crucial ingredient of normativity, it is not the whole normative enchilada. Normativity is also supposed to be about *authority*, and we might reasonably wonder whether the motivational relation identified by Williams manages to capture this authority.

One way to check would be to see whether the normative implications of Williams' theory are inconsistent with our standing normative beliefs. If they are, then we might conclude that Williams has discovered foundations, but for the wrong edifice. As it happens, Williams' internalism does contradict at least some of our prior normative commitments. For example, it implies that a man with a sufficiently malicious motivational set would have no reason to stop beating his wife. And this might tempt us to conclude that Williams' account cannot be correct. His critics have often succumbed to this temptation, but I shall try to resist it here. For one thing, this argumentative strategy presupposes that all of our prior normative beliefs are correct. This is an unwarranted assumption, especially given that we do not yet have a clear

---

10    Two notable exceptions are Korsgaard, "Skepticism about Practical Reason"; and Parfit, "Normativity."

notion of what normativity really is. When we discover the nature of a certain property, we must often jettison some of our prior beliefs about what does and does not have that property. It would therefore be hasty to reject Williams' theory because of these beliefs. Moreover, we must not overstate the amount of revision his internalism actually requires. As Williams observes, the conclusion that a man has no reason at all to stop beating his wife follows only on the implausibly strong assumption that he has absolutely no motives that could be subserved by stopping.[11] We are unlikely to encounter such a wholly contemptible individual in real life. And so although Williams' internalism may have unsavory implications in theory, those implications will seldom turn up outside philosophical thought experiments.

A more effective way to determine whether Williams has captured the authority of reasons is to appeal to Moorean, open question considerations. Williams is particularly vulnerable to this sort of challenge, since he takes himself to be defending an account of the "sense" of statements about reasons for action—that is, of what such statements "mean."[12] With this in mind, we could ask whether we have reason to do what we could be defeasibly motivated to do via sound deliberation from our current motives? If this is an open question, then that would seem to suggest that Williams has not succeeded in identifying the *meaning* of reasons statements. However, as I mentioned in chapter 2, the open question argument has come in for a great deal of criticism over the years, and it is by no means clear whether it succeeds even as an argument against analytically reductive views.[13] That said, we can avoid some of these issues by treating Williams' view as a *nonanalytically* reductive account of reasons for action. This makes irrelevant the matter of whether the above question is conceptually open or closed.

However, even if Williams is defending a synthetic theory of what reasons *are* (rather than an analytic theory of what reasons statements *mean*), open question worries remain.[14] Williams claims that reasons are considerations that can engage our existing motives, but it is a pressing question whether a consideration becomes a reason merely by standing in this particular

11    See Williams, "Internal and External Reasons," 25.

12    Williams, "Internal Reasons and the Obscurity of Blame," 40, and "Internal and External Reasons," 106. He also describes his view as an "interpretation of reasons statements" ("Internal and External Reasons," 101).

13    See chapter 2, note 71 above.

14    For a discussion of the ways in which the open question argument can pose a challenge even to synthetically reductive accounts of normativity, see Railton, "Naturalism and Prescriptivity," 157–9. See also Darwall, Gibbard, and Railton, "Toward *Fin de siècle* Ethics," 116–8; and Rosati, "Agency and the Open Question Argument," 490–501.

motivational relation. We can, in other words, acknowledge that a certain course of action would subserve one of our aims without taking ourselves to have settled anything of normative significance. The normative question remains open, even when the motivational facts have been settled. Moreover, once we start challenging Williams' account in this way, we might begin to wonder whether he is talking about normativity at all. His focus on a motivational connection suggests that he is addressing entirely the wrong question. We seek an account of what we *ought* to do or of what we have *reason* to do, but Williams has answered the question of what we are *likely* to do or of what we are *going* to do. Reasons are supposed to be considerations that guide and recommend, and Williams has identified considerations that seem able only to predict.

Here is another way to put the same point. We want to understand the normative *force* of reasons, or what early modern philosophers called the *bindingness* of obligation. The metaphors of "force" and "bindingness" are appropriate, since genuinely authoritative norms are not optional like the rules of a game. We are (we think) *bound* by the laws of morality in a way that we are not bound by the rules of chess. Normativity is supposed to be somehow inescapable. But Williams seems to take these metaphors a bit too seriously. Motives certainly exert force, and—if they are sufficiently strong—they can even bind us to perform certain actions. But this is the wrong sort of force. Strong motives push and compel us; they do not guide us. The laws to which Williams appeals are ultimately psychological laws, a subset of the physical laws of nature.[15] But there is nothing authoritative about physical laws. They are not normative, nor is motivational or psychological force the same thing as normative force.[16]

Suppose, for example, that you are addicted to heroin: you have a strong and often overwhelming desire for it. Williams view seems to entail that you therefore have a reason to shoot up. Since the consideration that the stuff in front of you is heroin can engage your addictive desire, this consideration counts as a reason for you to take it. Of course you may have other motives that generate reasons to abstain, but if your desire for heroin is stronger than these other motives—and let us suppose that it is—then your reasons to abstain will be

---

15    Davidson, of course, denies that there can be strict psychophysical laws. See "Mental Events" and
       "Psychology as Philosophy."

16    Here I am following Parfit, "Reasons and Motivation," 126–30; and "Normativity," 332–7. See also
       Korsgaard, "The Normativity of Instrumental Reason," 220–34.

outweighed by your reasons to indulge.[17] This is where the open question challenge comes into play, since the motivational force of your addictive desire does not appear to be authoritative in any way. Yes, it is strong, and it may even *bind* you in a certain sense, but not in the right sense. After all, our motives themselves, even our noninstrumental desires, can be subject to normative authority. Sometimes we have reasons not to want the things we want, even when we want them for their own sake. But if our noninstrumental desires can be the subjects of normative authority, then they cannot also be the source of that authority.

These open question challenges are hardly conclusive, though, since—if Williams is correct—the questions are not as open as they seem. Still and all, they constitute a *prima facie* case against Williams' proposed reduction.

Recall, however, that Williams' internalism includes a proviso meant to forestall this objection: it is only motivational force produced by "sound deliberation" that counts as the normative force of reasons. In chapter 1 I referred to this as the *rationality proviso* because it requires reasons to exert not merely motivational influence but also rational influence. Sometimes our desires influence us in nonrational ways; action motivated by a desire is not always action for a reason.[18]

In Williams' case it is hard to see how the rationality proviso is going to help. This is because his broadly Humean theory of motivation is also his theory of action for a reason. And a Humean account of agency may not have the resources to distinguish the merely motivational force of an addictive desire from the rational influence exerted by a reason. Of course, Williams' theory is only *broadly* Humean, and his conception of sound deliberation is not crudely instrumental as it is on a narrowly Humean view. Still, Williams' account of agency leaves him stuck with the thought that all of our noninstrumental desires are capable of exerting some rational influence, and thus he is also stuck with the conclusion that the considerations which engage those desires count as reasons.

---

17    Williams hints that his internalism need not be an account of the strength of competing reasons. (See "Internal Reasons and the Obscurity of Blame," 35–6.) However, if the motivational relation he identifies is meant to constitute the relation of being a reason (and that is how I am treating it), then there is no room for an alternative explanation of the strength of a reason. The only available account seems to be the straightforward one: the stronger the motivational force, the stronger the normative force.

18    See, for instance, the example of the nefarious psychologists and neurosurgeons in chapter 1 above.

This may be a flaw of Williams' overall view, but it need not count as a strike against his *internalism*, since we can easily substitute a more sophisticated but still broadly Humean theory of agency for the one Williams proposes. That is, we can retain Williams' account of what it is to be a reason for action but supplement it with a theory of action for a reason that does not count addictive desires as sources of rational influence. Consider, for example, Harry Frankfurt's hierarchical model of agency.[19] Frankfurt's model also rests on a broadly Humean foundation, since it takes desires, or conative states more broadly, to be the source of rational influence. Where Frankfurt differs with Williams is over which desires have that special status. For Williams, it is all noninstrumental desires, whereas for Frankfurt, it is second-order desires. On Frankfurt's view, then, your addictive desire does not exert the influence of a reason, no matter how strong or overwhelming it is, since it is not backed up by a second-order desire. You do not want to want heroin, and that is why your addiction is not reason-generating. And so even though Williams' theory as he presents it cannot meet the open question challenge stemming from addictive desires, a modified version that retains the original's Humean spirit can.

The underlying problem remains, however. A more sophisticated version of internalism calls for a more sophisticated open question challenge. Our original worry was that motivational force is not the same as normative force. Simply delimiting a specific kind of motivational force does not make that force any more normative. It is more like the physical force of gravity than the normative force of reasons. Of course, my modified version of Williams' view does identify the subset of motivational influence that counts as rational influence. But recall again that rational influence is simply the influence exerted by a consideration when it is the reason for which one acts. Rational influence is an explanatory notion, yet Williams seems to want to turn it into a normative one. On Williams' view, if $A$ can $\phi$ for reason $R$, then $R$ is a reason for $A$ to $\phi$; if $R$ can explain $A$'s $\phi$-ing, then $R$ also justifies $A$'s $\phi$-ing. This seems to conflate justification and explanation. Granted, the motivation for internalism is the idea that justification cannot be isolated from explanation. But even if they

19    See Frankfurt, "Freedom of the Will and the Concept of a Person." Frankfurt does not lay out his theory in terms of action for a reason. Instead his emphasis is on freedom of the will. We can, however, translate his view into an account of autonomous action or action for a reason and still be in keeping with the spirit of his approach. Whether Frankfurt would ultimately endorse such a translation is not important for our purposes here.

are closely related, it is hard to see how justification could be *no more than* the capacity to explain.

Warren Quinn presses this point forcefully.

How can a … state whose central significance in this context is to help explain our tendency to act toward a certain end, or in accordance with a certain principle, *rationalize* our pursuit of the end or our deference to the principle? How can the fact that we are set up to go in a certain direction make it (even prima facie) rational to decide to go in that direction? How can it even contribute to its rationality?[20]

Quinn's worry is that empirical naturalist views account for normativity only by changing the subject. We want to know about justification, and internalists respond with explanation. Alternatively, we are interested in authority, and internalists reply with governance. Even if, as internalists propose, authority and governance must go hand in hand, being governed is not the same thing as being governed authoritatively.[21] There is a difference between acting for a reason and acting for a *good* reason, a difference Williams' internalism seems to ignore.

Here is another way to describe the same problem. On Williams' view, if *A* acts for reason *R*, then *R* is exerting rational influence, in which case *R* is capable of exerting rational influence, in which case *R* is a reason for *A*. Of course, there are desires that exert *nonrational* influence, and when such desires dictate one's behavior one will not be acting for good reasons. But since *ex hypothesi* these desires exert only nonrational influence, in such cases one will *not be acting for reasons at all*. Williams' theory therefore seems to rule out the possibility of acting for a bad reason: either we act for a good reason, or we act for no reason at all. Yet the possibility of error—of wrong or bad or irrational action—is central to our concept of normativity. Normativity involves the idea of a standard of correctness or norm. And the notion of a norm implies the possibility of something being in violation of that norm, just as the notion of a standard of correctness implies the possibility of incorrectness. This error constraint is close to an axiom of metaethics, and Williams' internalism seems to run afoul of it.[22] The moment your action becomes incorrect (say, because it is driven by an addictive desire that does not exert rational influence), it ceases to count as a full-blooded action, in which case it is not

---

20  Quinn, "Putting Rationality in Its Place," 236.

21  Normative governance may require that one *regard* one's reasons as authoritative, but that is different from those reasons actually *being* authoritative.

22  For a list of philosophers who accept the error constraint, see Lavin, "Practical Reason and the Possibility of Error," 425 n. 1. Lavin himself questions whether this really is a constraint on theories of normativity.

really incorrect at all. Put another way, you cannot act contrary to your reasons without first leaving their jurisdiction, in which case you are not really subject to their authority in the first place. This is an unacceptable conclusion for a theory of normativity. Surely it must be possible to contravene the authority of our reasons for action and thereby act incorrectly. If so, then Williams' view—at least as I have stated it so far—must be wrong. Normative force must be more than merely the force exerted by a reason when we act for that reason.

Williams himself is worried about this problem, and his attempt to solve it involves the addition of another proviso. Up to now I have omitted any mention of this proviso in order to emphasize its importance. I have tried to demonstrate that without it, Williams has little claim to a satisfying theory of normative authority and reasons for action. Consequently, if his internalism is to capture the authority of reasons, this second proviso will have to do the heavy lifting.

## The Truth Proviso

Williams notes that reasons statements are about more than what an agent is motivated to do. They are also "concerned with the agent's rationality."[23] He acknowledges that if an account of reasons is to have "normative force" it must be able to "go beyond what that agent is already motivated to do."[24]

> '*A* has a reason to *ϕ*' means more than '*A* is presently disposed to *ϕ*.' One reason why it must do so is that it plays an important part in discussions about what people *should become* disposed to do.[25]

Among other things, an account of reasons needs to preserve our practice of citing reasons as a way of giving advice. We often tell people that they have reasons to act in some way even when they do not believe they have such reasons, and even when they are not disposed to act in that particular way at all.

To make room for the normativity of reasons, Williams incorporates a truth proviso into his conception of sound deliberation. According to Williams, an agent is deliberating soundly

---

23     Williams, "Internal and External Reasons," 103.

24     Williams, "Internal Reasons and the Obscurity of Blame," 36.

25     Williams, "Internal Reasons and the Obscurity of Blame," 36, emphasis added.

only if the motives from which he is deliberating do not rest of false beliefs.[26] To see the truth proviso at work, consider Williams' familiar example of the petrol and tonic. Suppose I desire to drink the contents of the glass in front of me because I believe those contents to be gin and tonic. In fact, however, they are petrol and tonic, a somewhat less savory combination. Williams observes that an account of reasons for action that does justice to the authority of those reasons cannot imply that I have a reason to drink the contents of my glass. This is where the truth proviso comes into play. Since considerations count as reasons only if they can engage motives through a process of sound deliberation, they will count as reasons only if the motives they can engage are not based on false beliefs.

The addition of this new proviso yields the right answer in the petrol and tonic case. Here my desire to drink the contents of my glass is based on a false belief about those contents, and so the considerations that trigger this desire do not count as reasons. The truth proviso also thereby allows for the possibility of action for a bad reason: I act for a bad reason when the desires driving my action are based on false beliefs. Therefore, Williams concludes, the addition of the truth proviso is sufficient to capture the authority of reasons. When we apply the notion of sound deliberation to the question of whether an agent has a reason, "we are allowed to change—that is, improve or correct—his beliefs of fact and his reasonings in saying what it is he has reason to do. This is already enough for the notion to be normative."[27]

With the addition of the truth proviso, Williams' internalism can also make sense of the practice of citing reasons in order to give advice. I may not be offering you advice when I remind you that you already want to $\phi$, but I *am* offering you advice when I point out that you would want to $\phi$ if only some of your false beliefs about $\phi$-ing were corrected.

Does all of this show that Williams has captured the normativity of reasons? Derek Parfit insists that it does not. That a consideration can be cited in order give advice is not sufficient for it to count as normative.

> If I say, "Your wine is poisoned," this claim may be advice. But that would not make this claim, or the fact that it reports, normative. For a claim to be normative, it must use at least one normative concept.[28]

---

26     This may explain why Williams opts for the phrase "sound deliberation" rather than "valid deliberation."

27     Williams, "Internal Reasons and the Obscurity of Blame," 36. See also Williams, "Postscript," 92–3.

28     Parfit, "Normativity, Naturalism, and Noncognitivism," 13.

Notice, however, that Parfit's explanation supports only his rejection of an analytically reductive view. That the assertion about your wine being poisoned can be treated as advice does not, Parfit claims, make the *fact* reported by this assertion normative. But he defends that conclusion only by pointing out that *the assertion itself* is not normative, since it does not employ any normative concepts. But this is simply beside the point if we are treating Williams' view as a nonanalytically reductive theory. Such a theory claims not that the statement "Your wine is poisoned" is itself normative, but rather that this statement identifies a reason for you not to drink the wine and that it does so in virtue of the fact that this consideration can engage one of your motives.

The failure of this particular objection does not mean that Williams is out of the woods, though, for we might wonder what licenses the addition of the truth proviso in the first place. Williams' primary justification for including it is that it yields the right results in cases like the petrol and tonic case. But this does not show that the truth proviso is consistent with the motivation behind internalism. Recall that in chapter 1 I argued that internalism seeks the sources of practical normativity in the conditions of agency. And thus I suggested that provisos *not* rooted in the nature of agency have no place in a genuinely internalist theory, at least not if that theory seeks to understand the nature of normativity itself.[29] What, then, of the truth proviso? What justifies its inclusion? Note that I am not questioning whether the truth proviso in fact produces the right results. The idea that we ought not to act on desires that are based on false beliefs is independently plausible. But internalism is not supposed to rest on any independent normative principles. It seeks to understand the nature of practical normativity, and so it cannot appeal to the freestanding normative claim that one ought to act only on one's true desires. Were internalism to allow such claims, then someone could defend an internalist position that included a "morality proviso" according to which only morally sound motives count as reason generating. As I argued in chapter 1, such a view would not be internalist at all.[30]

Williams is aware of this problem as well, and he attempts to spin an internalist story in support of the truth proviso. The grounds for including it, he argues, "are quite simple":

---

29    See chapter 1 above, especially pages 30–37.

30    Compare Williams, "Internal Reasons and the Obscurity of Blame," 36–7.

> Any rational deliberating agent has in his [subjective motivational set] a general interest in being factually and rationally correctly informed…. The basic point … is that on the internalist view there is already a reason for writing, in general, the requirements of correct information and reasoning into the notion of a sound deliberative route, but not a similar reason to write in the requirements of prudence and morality.[31]

According to this story, the truth proviso is not an independent standard at all. It derives its normative standing from the source that generates *all* of our reasons: our motives.

We could certainly quibble with Williams' claim that "any rational deliberating agent" has an interest in being factually correct. He even admits that there may be the occasional agent who needs to believe what is false, and that such an agent would therefore "have reason to acquire false beliefs."[32] But there is a more fundamental problem with Williams' response. The point of adding the truth proviso is to back away from the suggestion that the normative force of reasons derives entirely from the psychological force of motives. The addition of the truth proviso suggests the bindingness of reasons is not purely motivational. Instead motives derive some of their normative force from the truth of the beliefs upon which those motives rest. But now it turns out that *that* force itself depends on some further desire. And so we are back to our original and objectionable picture of internalism in which the force of reasons is not qualitatively different from the force of the laws of physics. Our problem with that original picture was that we could not see what is authoritative about our noninstrumental, second-order desires. Recognizing this problem, Williams' provided the truth proviso as a way to explain which desires are authoritative and which are not. Yet his official story is that the authority of this proviso is rooted in another motive. And this naturally raises the question: Whence comes the normative authority of *this* motive? What makes matters even more difficult for Williams is that our "general" desire to be rid of false beliefs often seems to be merely an instrumental desire. We know that false beliefs will prevent us from satisfying the rest of our desires, and we therefore want to be rid of such beliefs. But would most agents feel so negatively about having false beliefs if they knew that such beliefs would not prevent the satisfaction of their desires? Of course, we philosophers with our inquiring minds might desire to know the truth for its own sake, but we should hesitate before presupposing that the rest of

31    Williams, "Internal Reasons and the Obscurity of Blame," 37. See also Williams, "Postscript," 91–2.

32    Williams, "Internal Reasons and the Obscurity of Blame," 37.

the world is like us in this (or any other) respect. Thus, if the only justification for including the truth proviso is our general desire not to have false beliefs, then the proviso threatens to derail Williams' internalism, at least as an account of the nature of practical normativity and reasons for action.

Fortunately for Williams, there may be a better way to defend the truth proviso. Most philosophers agree that only facts (or states of affairs, or true propositions) can be normative reasons for action. This seems to follow from the concept of a reason: a false proposition or a nonexistent state of affairs cannot count in favor of an action or bear on the question of what to do. If this is correct, then this alone should be sufficient to account for the truth proviso. For in cases in which one's desire is based on a false belief, a full specification of the putative reason will include the proposition that is falsely believed. But since a false proposition cannot be a normative reason, the putative reason will not be a reason at all. Return for a moment to the petrol and tonic case. Recall that our worry was that without the truth proviso, Williams' view would entail the implausible conclusion that you have a reason to drink the petrol and tonic. But what would that reason actually be? The consideration that triggers your desire would have to be something like: "The glass in front of you contains gin and tonic." By hypothesis, however, that consideration is false, and so it cannot be a reason for you to do anything.[33]

In more complicated cases, it may not always be obvious that a full specification of the reason includes a false proposition. Consider, for example, a modified version of my story from

33  Here some philosophers might claim that your *belief* is your reason: that you believe your glass contains gin and tonic is a reason for you to drink the contents of your glass. And this consideration is true. You really do believe that it is a gin and tonic. This possibility raises a couple of difficult issues in metaethics and the philosophy of action. First, there is the question of whether reasons for action are psychological states, the content of psychological states, or the states of affairs to which those contents refer. This question is most pressing when it comes to reasons in explanatory contexts—that is, to the reasons for which we act. Here there are good reasons for thinking both that reasons are psychological states and for thinking that they are contents or states of affairs. For the former view, see Davidson, "Actions, Reasons, and Causes"; and Smith, *The Moral Problem*, ch. 4. For the latter view, see Dancy, *Practical Reality*, ch. 4–5; and Collins, "The Psychological Reality of Reasons."

Second, there is the matter of whether the belief that your drink is gin and tonic does, in some limited sense, count in favor of imbibing that drink. Some philosophers hold that even if your belief does not provide or constitute a *reason*, it does provide a *rational* or *normative requirement*. See, for instance, Broome, "Normative Requirements." This does not pose a threat to the truth proviso, though, since if we accept that your belief rationally requires you to drink the petrol and tonic, this requirement will stem not from the false proposition that your drink is a gin and tonic but rather from the true proposition that you believe it to be a gin and tonic.

chapter 1 about having to choose between marking essays and attending a concert.[34] Suppose that although I believe tonight to be my only chance to hear the Vienna Philharmonic play Mahler this year, there is a second performance tomorrow night of which I am unaware. And so if I stay late in the office tonight, I can finish my grading, keep my promise to my students, *and* attend the concert tomorrow night. Suppose further that although my motivation to attend the concert tonight is currently much stronger than my motivation to mark essays, I would be much more inclined to skip tonight's concert if only I knew about tomorrow's repeat performance. Now, what would a version of Williams' internalism without an explicit truth proviso say about this case? On the one hand, it might seem that I have most reason to attend the concert tonight, since that is currently what I most want to do. The strongest desires in my actual motivational set, together with my beliefs, are pushing me in the direction of indulging my love for Mahler. However, a full specification of this supposedly overriding reason would include the false proposition that there is not another concert tomorrow. And so this cannot be the strong reason it seems to be. Of course, even absent this false proposition the pleasure I would receive from tonight's concert might still be a reason to attend. But in the absence of my false belief that I cannot hear the Vienna Philharmonic tomorrow, my motivation (and hence my reason) to attend tonight would be much weaker. If this is correct, then the truth proviso does not need to be derived from a general aversion to false beliefs. It follows from the very concept of a reason. The truth proviso turns out to be consistent with the motivation behind internalism after all.

But does it solve the problem it is meant to solve? Does its addition make internalism more plausible as an account of the authority of reasons for action? Ultimately I think it does not, as I shall attempt to demonstrate in the next section.

## The Authority of Motives

Williams runs motivational force through the filter of our epistemic reasons and calls the product normative force. Yet even though this product *is* genuinely normative, we seem to have arrived at the wrong sort of normativity.

Recall that Williams acknowledges that our noninstrumental motives—even when they are supported by second-order motives—are not authoritative on their own. That something

---

34    See page 15 in chapter 1 above.

can be my reason for acting (a psychological and explanatory fact) does not establish that it is a reason for me to act (a normative fact). Adding the sanction of epistemic reason is what, according to Williams, transforms his view into an account of practical normativity and reasons for action. This suggests that it must be the authority of our epistemic reasons that is pulling all of the normative weight. In other words, what transforms my noninstrumental motives into authoritative reason generators is just the fact that these motives are not based on any false beliefs. But how can the mere absence of epistemic error be sufficient to turn psychological force into normative force?

The picture we are evaluating now looks like this: a consideration counts as a practical reason merely because the motive it engages depends only on true beliefs. The correctness of our actions therefore derives entirely from the correctness of our beliefs. Yet the open question objection discussed above shows that there is nothing normative or authoritative about the psychological force of our motives, at least when they are considered in isolation. How can adding only the proviso that those motives are not based on false beliefs supply the missing normativity? None of our open question worries are even remotely allayed by Williams' assurance that the motives in question do not rely on false beliefs or faulty theoretical reasoning. If there is nothing authoritative about a motive, then there is nothing authoritative about a motive that rests only on true beliefs. Consequently, true beliefs are not sufficient to guarantee that when we act for a reason we are acting as we ought. If this picture accurately reflects Williams' view, then the proper conclusion is not that he mistakes motivational force for practical normativity, but rather that he simply does away with practical normativity. There are no reasons for action; there are only motives and reasons for belief. In other words, there are no incorrect actions, only uninformed ones.

Let me put my point another way. I stated above that Williams' internalism seems to have no trouble capturing the practical significance of reasons for action. But close examination of his view reveals that I was too hasty to credit him on this matter. According to Williams, practical significance is simply the felt psychological force of motives. But the psychological force even of our second-order motives does not settle our practical questions. It has no more intrinsic practical relevance than addictive or nonrational desire.[35] Just as we can wonder whether we *should* want the things we want, so can we wonder whether our second-order

---

35    See Korsgaard, *The Sources of Normativity*, 15–6, for a similar argument.

desires are directed as they *ought* to be. We can, in other words, struggle against this force in our practical deliberations or even shrug it off with a casual "So what?"[36]

As a consequence, Williams' account of practical normativity leaves us unable to contemplate the correctness of our noninstrumental commitments. Sure, we can wonder whether these commitments are based on false beliefs; but if they are not, then we must simply accept them as correct come what may. Yet this does not do justice to the ways in which we actually reason about what to do. Sometimes we ask normative questions that our epistemic reasons cannot answer. We ask, for example, whether what we ought to care about is the same as what we do care about. In the face of even overwhelming motives to perform some action, we may still wonder whether there are reasons to do something else instead. That is, we may wonder whether something we do not desire might nonetheless be desirable or worth desiring. We also reflect critically on our commitments, sometimes even on the ones that seem most central to our identity. Admittedly, we do not confront these sorts of questions on a regular basis. It is not every day we have to decide what career path to follow or what sort of person to be. But we *do* sometimes face such choices, and when we do we must weigh considerations for and against our various alternatives.[37] We must, in other words, engage in practical reasoning. But Williams' internalism does not allow for this. According to Williams, knowing how all of the relevant considerations relate to your noninstrumental motives (once those motives have been cleansed of falsehoods) *just is* to recognize all of your reasons for action. There is nothing left to figure out—no more reasoning to be done. But these other questions *do* arise in practical deliberation. If Williams' internalism cannot accommodate them, then it cannot fully capture the practicality of reasons for action.[38]

All of this suggests that despite appearances, the truth of our beliefs does not have the normative muscle Williams attributes to it, and so he must still assign some normative force to our noninstrumental motives. According to this revised picture, the authority of reasons is a concoction made up of the putative normative force of our motives and the normative force of our epistemic reasons. The picture now is one in which our motives have a *prima facie* authority

---

36   Compare Rosati, "Naturalism, Normativity, and the Open Question Argument," 49.

37   Here I am following Darwall, "Autonomist Internalism and the Justification of Morals," 260. Compare Darwall, *Impartial Reason*, 15–6; and Gibbard, *Wise Choices, Apt Feelings*, 19–21; as well as Platts, "Moral Reality and the End of Desire."

38   Compare Scanlon, *What We Owe to Each Other*, 52–5.

that can be undermined only by a dependence on false beliefs. If our motives do not rest on any such beliefs, however, then their status as authoritative is assured. And now the open question challenges return with full force. Once again we can ask: What is authoritative about a noninstrumental motive, even one that is supported by the relevant second-order motives? How are the psychological laws relating motives to actions any more normative than the laws of motion?

As I noted above, though, the open question argument may not be conclusive, since Williams can always maintain that the normative questions only *seem* open, much as the question of whether water is $H_2O$ may have seemed open to Lavoisier's contemporaries. If Williams' proposed reduction succeeds, then his internalism does *not* do away with practical reasoning at all, since then practical reasoning *just is* theoretical reasoning about whatever subserves our noninstrumental aims and commitments. In this case, however, Williams must regard practical reasoning about our noninstrumental commitments as either confused or incoherent. Being a consideration that engages an informed noninstrumental motive *just is* being a reason. Consequently, it cannot be appropriate to ask whether any such considerations really are reasons, nor whether any considerations that do not engage any of our motives might be reasons nonetheless.

This is an unacceptable consequence of Williams' reduction. The externalists discussed in chapter 2 may have been too quick to condemn reductionism in general. However, I believe they are correct to reject a view that dismisses such an important aspect of what we commonly call practical reasoning. My objection here is not simply that Williams' internalism entails that some of our settled normative beliefs are false or confused. I believe that any reductive view is likely to have this consequence. Rather, my objection is that Williams' reduction is inconsistent with our general understanding of what practical reasoning involves. It is true that we often deliberate from a settled conception of our aims and commitments. But often we step back from those aims and attempt to assess their merit in comparison to competing aims—aims we have not yet adopted as our own. To do this we need to be able to weigh reasons that count for and against our noninstrumental motives. If Williams' internalism does not provide us with the resources to do this, then it cannot be a complete account of the normativity of reasons for action. Even with the boost from the truth proviso Williams cannot clear the open question hurdle.

# Full Information

At least a few philosophers have thought that the problem with this sort of internalism is not that there is *too much* emphasis on the authority of our true beliefs, but rather that there is *not enough* emphasis. These philosophers replace the truth proviso with a stronger *full information proviso*. They hold that in order for our motives to exert normative rather than merely motivational force, we must do more than merely cleanse these motives of any reliance on false beliefs.[39] In addition, we must also determine whether our motives would survive exposure to *all* of the relevant (nonnormative) facts. As a result, only motives that could withstand our possession of "full information" count as authoritative. What someone *ought* to do, according to these full information theorists, is what he *would* do were he aware of all of the relevant facts.[40]

Two observations are appropriate here. First, it is unclear whether my defense of the truth proviso's internalist credentials can be extended to cover the full information proviso.[41] The truth proviso fits comfortably into an internalist approach only because desires based on false beliefs generate reasons that, when fully specified, contain a falsehood themselves. And since reasons cannot be false, desires resting on false beliefs do not generate genuine reasons. Can we make the same claim on behalf of desires based on less than full information? Perhaps. One could argue that the full specification of any reason generated by such desires must also include a falsehood. Suppose, for example, that I am planning an early afternoon trip to the local market because I need to pick up some bread. Unbeknownst to me, however, my nemesis

---

39    According to Velleman, this actually turns the intuition behind the truth proviso "on its head, so as to say, not that one shouldn't want what one wouldn't want if one knew more, but rather that one *should* want whatever one *would* want if only one knew more" ("Brandt's Definition of 'Good,'" 371).

40    The most famous full information internalist is probably Richard Brandt, who defends his revisionist account of rationality in part 1 of *A Theory of the Good and the Right*. For a somewhat different approach, see Hubin, "Hypothetical Motivation."
         Most other full information theorists apply their view only to the notions of well-being and prudential value. See, for instance, Railton, "Facts and Values" and "Moral Realism," 173–8; Griffin, *Well-Being*, chs. 1 and 2; Hare, *Moral Thinking*, 105–5, 214–6; Rawls, *A Theory of Justice*, 407–24. For criticisms of such views, see Velleman, "Brandt's Definition of 'Good'"; Sobel, "Full Information Accounts of Well-Being"; and Loeb, "Full-Information Theories of Individual Good." As I suggest below, I believe that the full information approach is more plausible when applied to the notion of well-being than when applied to reasons for action and practical normativity more broadly.

41    It is worth noting that Williams' defense of the truth proviso is *certainly* not available to full information internalists. Most agents decidedly *lack* a general desire to possess *all* of the relevant information. Of course, I argued above that Williams' defense fails even for the truth proviso, and so this is probably no great loss for full information theorists.

at work—let us call him Shel—is planning a trip to the same market at the same time. I cannot stand Shel. Whenever I see him he says something insulting that invariably makes me feel terrible about myself. Thus, if I knew that he was going to be at the market this afternoon, my motivation to go would instantly evaporate. I would much rather live without bread than encounter my nemesis. In this case, the motive pushing me to go to the market is not based on any false beliefs. I do not believe that Shel *will* be there, nor do I believe that he *will not* be there. Frankly, the possibility just has not occurred to me. Nonetheless, a full information theorist might contend that a complete specification of my reason would include the false thought that my trip will not coincide with Shel's.

The problem with this strategy is that it makes it impossible for me fully to grasp my reasons for action. To see this, suppose that Shel is planning a trip to a *different* market. Assuming that no other unpleasant surprises await me, I now have a reason to go to the market for bread, according to the full information theorist. But in order for me to grasp this reason, I must know not only that Shel will not be there, but also that I will not be crushed by a falling piano on my way to the market, that the ceiling will not cave in on my head while I am there, that a meteor will not pulverize me as I walk home, and so forth. In other words, if a full specification of my reason includes all of the relevant facts, then I must actually grasp all of these facts in order to grasp my reason. And since I can never be aware of *all* of the relevant facts, I can never grasp the reasons for which I act.

But surely I *do* sometimes grasp the reasons for which I act, and I am therefore inclined to conclude that the full information proviso is not consistent with the underlying motivation for internalism. I do not mean to suggest that this proviso is not independently plausible. On the contrary, I find the ideas behind full information accounts extremely compelling. But recall that internalism, as I understand it, is a reductive approach to the nature of practical normativity itself, one that seeks to understand reasons for action in terms of our capacity to act for reasons. And thus internalists cannot defend aspects of their view on independent normative grounds and still remain internalists. Consequently, our failure to find a nonnormative defense of the full information proviso threatens to disqualify full information theories as accounts of what practical normativity itself consists in.

I shall turn now to my second—and ultimately more damaging—observation about the full information approach. The truth proviso was insufficient to transform Williams' view

into an account of the authority of reasons for action not because this proviso was too weak but rather because it sought to ground that authority entirely in the authority of our reasons for belief. Williams simply went looking for practical normativity in the wrong place. This suggests that merely strengthening the truth proviso into a full information proviso will not help. Full information theorists are *still* looking in the wrong place. Admittedly, they do capture an important aspect of practical reasoning. When we are trying to decide what we ought to do, sometimes we *do* consider what we would care about if we had more (or even full) information. But this does not exhaust the scope of our practical reasoning. To see this, consider another example. Suppose you cannot decide whether you should attend law school or become a horse jockey. You are fully aware that a legal career would be more lucrative, lead you to higher standing in the community, and provide plenty of intellectual stimulation. But it has always been your dream to become a jockey. To feel the thrill of guiding your horse into the home stretch several times a day—what could be better? Suppose further that you have an accurate sense of how your life will turn out given one career or the other. Both options are sure to subserve many of your desires and cares, and so your choice is not between a life of satisfaction over a life of misery. Yet you also know deep down that the life of a jockey is going to be the one that satisfies more of your strongest noninstrumental commitments. It promises more of what you most want and care about for its own sake. Finally, suppose that no further information— about jockeying, lawyering, or any other career—would significantly affect the strength of your commitments. The life of a jockey is the one you would pursue even if you were fully aware of all of the relevant facts.[42]

The full information theory implies that when you reach this realization, your practical reasoning is effectively complete. You know which career is going to subserve your strongest informed motives, and so you know what you have most reason to do. You may even be able to feel the psychological force of these motives—the pull of the racetrack overwhelming the urge to study law. But surely, even in the face of this tangible force, there are further questions you can ask. You are aware that becoming a jockey is what you are most motivated to do, yet you might still wonder whether the life of a jockey is the best for you. Are there not aspects of the law you care little about now but that are nonetheless worth promoting? Even though

---

42    For more examples of this kind, see Gibbard, *Wise Choices, Apt Feelings*, 19–22; and Velleman, "Brandt's Definition of 'Good,'" 360–2.

you love the thrill of racing more than you love torts and contracts, are there not independent reasons to pursue a life of the mind over a life in the saddle? Does it matter that were you to choose a legal career you could devote your life to helping those in need (or to helping multinational corporations evade their liability)? These are questions not about what would subserve your informed desires and commitments but rather about what is *worth* wanting and being committed to for its own sake. The full information view entails that such questions are confused.[43]

Again, a full information theorist could simply stick to his guns and insist that practical reasoning *just is* reasoning about what we would be motivated to pursue under conditions of full information. But he must then show that our informed motives can properly play the authoritative role they have been assigned. Richard Brandt presents such a case in *A Theory of the Good and the Right*.[44] According to Brandt, judgments about what would subserve our fully informed aims have "recommending force" in virtue of certain features of our motivational system. The awareness that our immediate desires are not in line with our informed aims creates a feeling of discomfort akin to the uneasiness of cognitive dissonance. "One is made uncomfortable by the awareness, and is [therefore] motivated to remove its source."[45] It is this additional motivation that constitutes the recommending force of facts about our informed aims and desires. It should be clear, though, that this is no response to the open question objection facing the empirical naturalist internalists. Brandt acknowledges that the recommending force of informed desires "depends on the prevalence of other desires."[46] But since we are challenging the very authority of desires, it is no response to this challenge to

---

43    Compare Velleman, "Brandt's Definition of 'Good,'" 355–9.

44    Brandt acknowledges that he has not captured every facet of our everyday sense of practical normativity. He considers large portions of this "sense" to be nonsense, and he presents his own account of practical rationality in an openly revisionist spirit. (He calls it a "reforming definition." See *A Theory of the Good and the Right*, 10–1. For a similarly revisionist view, see Railton, "Naturalism and Prescriptivity.") Consequently, it may seem unfair to criticize his view for failing to capture the normativity of our everyday notions of rationality and reasons for action. That said, Brandt does "want to be sure" that his proposal "does not overlook something important that is asked about in traditional ordinary-language normative questions," and thus he devotes an entire chapter—chapter 8—to showing that his account captures the "recommending force" of normative judgments (14). Compare Gibbard, *Wise Choices, Apt Feelings*, 18–22, 183–8.

45    Brandt, *A Theory of the Good and the Right*, 157.

46    Brandt, *A Theory of the Good and the Right*, 159.

appeal to further desires. We can simply reapply the open question argument: whence comes the authority of the desire to avoid the discomfort of "dissonance"? Can we not legitimately wonder whether we ought to be so averse to aims that diverge from what our fully informed desires would be?[47]

If we can, then Brandt's theory cannot be a complete account of practical reasoning and reasons for action. Even if considerations that engage our fully informed motives are authoritative, they are only a subset of the range of considerations to which we can appeal when we are deliberating about what we ought to do.[48] Now, Brandt suggests at several points that his account is only meant to capture those aspects of practical reasoning that are "important" or that have "an identifiable point."[49] And so perhaps he thinks that the practical questions I have been discussing are unimportant or pointless. But he never argues for this conclusion. As David Velleman notes, Brandt wants to jettison the bits of practical reasoning that escape his full information account. "Yet the bits to be jettisoned are, in reality, particular thoughts and particular questions that can be framed in our old vocabulary but not in Brandt's new one."[50] Unless Brandt can demonstrate that these thoughts and questions are incoherent, this counts as a fairly decisive consideration against his full information theory of reasons for action.[51]

This does not mean that Brandt's theory has no advantages over Williams'. For one thing, the move from the truth proviso to the full information proviso makes it plausible to argue that at least some of the apparently open practical questions are in fact closed. In the above example concerning your career choice, I proposed that it is an open question whether the life of a jockey *is the best life for you*. However, I believe that full information theorists can make a strong case that this is not an open question—that facts about what is good *for you* or about what would contribute to your *well-being* are indeed exhausted by facts about what would subserve your informed aims and commitments. Full information theories are strongest, I think, when

47    These seem like genuinely open questions, and the fact that Brandt closes them belies his claim that his account "does not import any substantive value judgments into the concept of 'rational'" (13). Parfit presses a similar objection in "Normativity, Naturalism, and Noncognitivism," 25–31.

48    This suggests that if considerations about our fully informed motives are indeed normative, it is likely in virtue of some more fundamental feature they share with other normative considerations. This is consistent with the account of reasons for action I defend in chapter 4 below.

49    Brandt, *A Theory of the Good and the Right*, 14, 127.

50    Velleman, "Brandt's Definition of 'Good,'" 356.

51    But see note 44 above.

they are offered as accounts of prudential value and of what would make the best life for you. However, even if some such theory is correct, there are plenty of other normative questions that are left wide open. And so although the full information approach is well suited to the domain of prudential reasons, it is not a good fit for a theory of reasons for action and practical normativity in general.[52]

Where does this leave the empirical naturalist approach to internalism? Empirical naturalist internalists start with the proposal that normative force is simply the psychological force of our noninstrumental motives. Confronted with open question arguments that challenge the authority of these motives, these internalists introduce authority via provisos that incorporate the uncontroversial normativity of our epistemic reasons.[53] I have suggested that this amounts to a reduction of the normativity of reasons for action to the normativity of reasons for belief, and I have argued that this vitiates empirical naturalist internalism's claim to be an account of practical reason at all.[54] There must be more to the normative force of reasons than the motivational force of our aims, desires, and commitments. And while the addition of

[52] Brandt's full information proviso actually requires more than that our motives be fully informed. They must also be able to survive a process of supposedly "value-free reflection" that Brandt labels "cognitive psychotherapy." Cognitive psychotherapy requires not just an awareness of all of the relevant information but also a repeated and maximally vivid presentation of that information. I have not discussed this aspect of Brandt's view because I do not think it is consistent with spirit of internalism. By itself, this is not a criticism of Brandt, since he does not claim to be a member of the internalist tradition. I have been treating him as one merely to see whether his ideas can contribute to a plausible internalist account of reasons for action. That said, what makes Brandt's account of cognitive psychotherapy inconsistent with internalism is the fact that it is normatively loaded. It presupposes that a particular sort of reflection is normatively privileged.

    Most of Brandt's critics have rejected the claim that the process of cognitive psychotherapy is privileged in this way, and many have argued it is not a desirable sort of reflection at all. See, for instance, Velleman, "Brandt's Definition of 'Good,'" 356–62, 364–71; Rosati, "Brandt's Notion of Therapeutic Agency"; Sobel, "Full Information Accounts of Well-Being," 792–3; Loeb, "Full-Information Theories of Individual Good," 4–11; and Gibbard, *Wise Choices, Apt Feelings*, 20–1. Railton presents a more compelling version of the full information approach in "Facts and Values" and "Moral Realism." For general criticisms that encompass Railton's account as well as Brandt's, see Sobel, "Full Information Accounts of Well-Being," 796–810; Loeb, "Full-Information Theories of Individual Good," 11–17; and Rosati, "Persons, Perspectives, and Full Information Accounts of the Good."

[53] Compare Darwall, "Internalism and Agency," 165.

[54] It is important to recognize that even if the arguments I have presented so far are sound, we cannot conclude that *all* empirical naturalist internalist reductions will succumb to them. Although the views considered above have, in Connie Rosati's words, "failed to fasten upon the right content, and, thereby, to capture" normativity, some future empirical naturalist proposals may have more luck ("Naturalism, Normativity, and the Open Question Argument," 47). Such a proposal may appear as soon as chapter 4 below.

a truth or full information proviso admittedly provides *some* normativity, it is not normativity of the right kind. We must be able to reason *practically* about our final aims and commitments, and this is something that none of the empirical naturalist views I have considered allow.[55]

Most internalists who embrace the empirical naturalist approach acknowledge that their view fails to capture certain aspects of our ordinary notion of reasons for action. They also, however, tend to believe that we cannot give a complete and satisfactory account of this notion. Yet unlike the externalists I discussed in chapter 2, they find our inability to explain this notion troubling. As Darwall puts it:

> On the one hand, they assert that there is no such thing as normativity as traditionally conceived, but they argue, on the other, that many of the traditional conception's functions are served just as well by a sufficiently close substitute that can be based in a philosophically respectable way.[56]

Motivation—perhaps tempered by theoretical reason—seems the best (and perhaps the only) candidate to replace our traditional understanding of normative force. There is nothing the source of reasons could be other than the goal-directed force of motivation.

> For the philosophical naturalist, concerned to place normativity within the natural order, there is nothing plausible for normative force to be other than motivational force, perhaps when the agent's deliberative thinking is maximally improved by natural knowledge.[57]

For the empirical naturalist internalists, the only alternative to this motivational account— aside from outright normative nihilism—is the externalist view that some considerations "just are" reasons. Williams, probably speaking for most internalists, considers this kind of externalism to be little more than "bluff and brow-beating," and he is therefore reluctant to embrace it.[58] I believe that this reluctance is admirable. If our only choices for an account of

---

55    At this point one might object that my conception of internalism is too narrow, and that were I to count philosophers like Michael Smith (in *The Moral Problem*) as internalists, I could evaluate internalist views that are better able to capture the normativity of reasons for action. I concede that theories like Smith's can indeed come closer to capturing the authority of practical reasons. However, as I argued above in chapter 1, this is only because such accounts rely on an independently normative (albeit procedural) conception of rationality. Since I am interested in internalism as an account of the "nature of normativity itself," to borrow Darwall's phrase once again (*The British Moralists and the Internal 'Ought'*, 10), I classify views that incorporate an independently normative rationality proviso as externalist.

56    Darwall, *The British Moralists and the Internal 'Ought'*, 320–1.

57    Darwall, "Internalism and Agency," 168.

58    Williams, "Postscript," 95.

the nature of normativity are an openly revisionist empirical naturalist internalism on the one hand and the externalist view criticized in chapter 2 on the other, I am inclined to opt for revisionism.[59]

But it remains to be seen whether these are in fact our only alternatives. We are still on the hunt for an internalist account that does justice to both the authority and practicality of reasons, and this is precisely what defenders of autonomist internalism promise to deliver. Let us therefore turn to their approach.

### Autonomist Internalism

The second internalist tradition—as Darwall's label indicates—assigns a central role in its theory of normativity to the notion of autonomy. Autonomist internalists hold that "a free rational agent can only be bound by constraints emanating from his own will," and thus that the normativity of reasons for action "derives from self-imposition in the activity of autonomous living."[60] "How else," Darwall asks, "*could* demands necessarily bind any rational agent unless they were somehow rooted in autonomous rational will itself?"[61]

Recoiling from the empirical naturalists' reduction of practical reason to theoretical reason, autonomist internalists believe that from the standpoint of a rational or autonomous agent, reason has a *purely practical* dimension. This alone promises improvements over the empirical naturalist approach. If pure practical reason is possible, then fundamental practical questions about what sort of life to lead and what sort of motives to have—questions that were closed off by Williams and Brandt—will be wide open, as it seems they should be. An autonomist internalism will close only the normative questions whose answers are dictated by the nature of autonomous willing itself.

Many autonomist internalists employ a conception of autonomy that retains an explicitly normative element.[62] These internalists are generally content to account for the normative

---

59    Philosophers of a different philosophical disposition will regard revisionism of this sort as abandoning rather than rescuing ethics. See Parfit, "Normativity," 339.

60    Darwall, "Autonomist Internalism and the Justification of Morals," 263.

61    Darwall, *The British Moralists and the Internal 'Ought'*, 20.

62    For a helpful discussion of such theorists, see Darwall, "Autonomist Internalism and the Justification of Morals," 264–5.

authority of morality in terms of more basic rational or deliberative ideals.[63] Compelling
though such views can be, they will not concern us here, for they cannot halt the regress of
justification. They incorporate plausible but ultimately independent normative standards or
assumptions that lack the required immunity to calls for further justification. Consequently,
our focus will be on autonomist theories that purport to provide an account of the nature of
normativity itself. Such foundational theories must invoke a conception of autonomous willing
that is empty of any normative content. This conception can then serve as the basis for the
derivation of our authoritative reasons and obligations.

The foremost autonomist internalist currently on the scene is undoubtedly Christine
Korsgaard. In *The Sources of Normativity*, she develops an account of normative authority that
locates the foundations of ethics in our capacity for autonomous willing.

> The source of the normativity of moral claims must be found in the agent's own
> will, in particular in the fact that the laws of morality are the laws of the agent's own
> will and that its claims are ones she is prepared to make on herself. The capacity for
> self-conscious reflection about our actions confers on us a kind of authority over
> ourselves, and it is this authority which gives normativity to moral claims.[64]

Thus, on Korsgaard's view, it is *our* authority as reflective agents that is the source of
normativity. When we judge that certain considerations count in favor of an action—that
is, when we reflectively endorse an action in light of such considerations—we *make* these
considerations into normative reasons for that action. Normativity, according to Korsgaard, is
created by our will.

The centerpiece of Korsgaard's theory is her first-person conception of normativity.
As with the empirical naturalist internalists, Korsgaard's starting point is the practicality of
reasons for action. She too considers normativity to be, "among other things, a psychological
force."[65] However, she rejects the empirical naturalist attempt to explain the practical

63    See, for instance, Falk, "Action-Guiding Reasons" and "Morality, Self, and Others," as well as Darwall, "Reply
      to Terzis," 119–24; and "Impartial Reason," 118–9. In these two papers (the texts of which overlap), Darwall
      puts an autonomist spin on the theory he defends in his first book, *Impartial Reason*. Darwall went on to
      develop the idea of autonomist internalism in a series of papers (see especially "Autonomist Internalism and
      the Justification of Morals" and "Internalism as Agency"), and also in *The British Moralists and the Internal
      'Ought'* (especially chs. 1, 5, 7, 9, and 11). However, perhaps because of the sorts of doubts discussed in
      chapter 4, note 59 below, Darwall never recast his own theory of reasons in fully autonomist terms.

64    Korsgaard, *The Sources of Normativity*, 19–20.

65    Korsgaard, "Self-Constitution," 1.4, here cited by lecture number and then page number.

significance of reasons in terms of the merely hydraulic force of motives. Instead, she attacks the problem of normativity from an avowedly first-person perspective: "The normative question is a first-person question that arises for the moral agent who must actually do what morality says."[66] Since the nature of an appropriate answer is determined by the nature of the question to which that answer is directed, the answer to the normative question must "spring from the position from which the normative question arises, the first-person position of the agent who demands justification of the claims which morality makes upon him."[67] Hence, according to Korsgaard, considerations are authoritative insofar as they recommend an action "from the agent's own, first-person perspective."[68] They justify your actions only insofar as they justify them "from your own point of view."[69]

Korsgaard's first-person approach seems to ensure the practical significance of reasons. By treating the psychological force of reasons as a force that confronts agents from within practical deliberation, she essentially builds what I referred to in chapter 2 as the practical conception of ethics—the idea that reasons for action are considerations that bear on the question of what to do—into her account of the sources of normativity. If some consideration does not present itself as relevant from within practical deliberation, then it is not a reason. The practicality of reasons is therefore assured.

That leaves the authority of reasons, and it is here that Korsgaard—like her fellow internalists—runs into trouble. A number of Korsgaard's critics have observed that what is justified from your own point of view and what is actually justified seem to be two different matters. Thus G. A. Cohen concludes that "it looks as though what she has investigated is the experience or phenomenology of obligation, not its ground or authenticating source."[70] Similarly, Rachel Cohon argues that Korsgaard confuses having a reason with merely taking something to be a reason.[71] Allan Gibbard complains that she seems to conflate value and

---

66    Korsgaard, *The Sources of Normativity*, 16.

67    Korsgaard, *The Sources of Normativity*, 16.

68    Korsgaard, *The Sources of Normativity*, 14.

69    Korsgaard, *The Sources of Normativity*, 15.

70    Cohen, "Reason, Humanity, and the Moral Law," 183.

71    Cohon, "The Roots of Reasons," 68–9.

valuing.[72] The latter is surely a first-person phenomenon, but the former need not be. Finally, Parfit contends that Korsgaard's internalism falls into the same trap that snared Williams. In response to her claim that normativity is a sort of psychological force, Parfit claims that "normativity, so understood, is a kind of unavoidable and irresistible motivation."[73] Whether someone feels the psychological force of a particular consideration says nothing about whether that consideration exerts any *normative* force. An agent's failure to feel the force of an authoritative reason may merely be a reflection of poor moral training or some overwhelmingly strong, self-centered desires. "Psychological necessity, though it may have normative significance, is not normative necessity."[74]

Parfit's objection reveals a misunderstanding of at least one aspect of Korsgaard's view. She never asserts that normativity involves *motivational* necessity. Motivation, after all, is a causal matter, and Korsgaard is quite clear that normative force cannot be reduced to causal force. Causal necessity makes a certain outcome inevitable. The sort of necessity involved in normativity, which Korsgaard calls "rational necessity," does not.[75] In cases of acedia or akrasia, we may be confronted by the overwhelming normative force of a set of reasons and yet fail to act. Rational necessity—being essentially first-personal—involves only a feeling or experience of necessity. It is, in Korsgaard's words, "a necessity we confront" in deliberation.[76] This is why Korsgaard can claim that a "practical reason must function both as a motive and as a guide, or a requirement."[77] The normative or binding force of reasons is distinct from their causal, motivating force. And thus Korsgaard actually accepts Parfit's contention that normativity is a matter of justification, not motivation.

Yet this leaves the general objection leveled by her critics still standing. From the fact that Korsgaard's theory can distinguish normative from motivational force it does not follow that she has captured the peculiar *authority* of reasons. Even if Korsgaard's notion of normative force is not reducible to motivating force, it remains unsatisfyingly psychological. Consider her

72    Gibbard, "Morality as Consistency in Living," 150.

73    Parfit, "Normativity," 373.

74    Parfit, "Normativity," 376.

75    See Korsgaard, *The Sources of Normativity*, 226 n. 11, and "Normativity, Necessity, and the Synthetic *A Priori*," 7–8. Compare Railton, "Normative Force and Normative Freedom," 324–6.

76    Korsgaard, "Normativity, Necessity, and the Synthetic *A Priori*," 7.

77    Korsgaard, "The Normativity of Instrumental Reason," 219.

claim that the term reason refers to a sort of "reflective success."[78] Reflective success may be required for us to *act for a reason*, but—her critics argue—we can have a reason to act without successfully reflecting (or even reflecting at all). If my friend is in pain, then I have a reason to run to his aid, even if his need never enters my mind. It seems that a good reason has normative force whether we confront that force during deliberation or not. Moreover, we can successfully reflect on a set of considerations and then act accordingly without having a good reason to do so. Thus, reflective success seems to be neither a necessary nor a sufficient condition for a consideration to count as a reason for action. The objection, then, is that full-blooded normativity or justification involves more than justification from the agent's own, first-person perspective; and if Korsgaard's theory reduces the former to the latter, she has lost normativity in her attempt to discover its source.[79]

This objection can be stated more broadly, as a challenge to the autonomist approach in general. For it is not at all clear how principles that stem from our own will can ever possess genuine authority over our will. After all, any law that the will legislates it can just as soon overturn. Yet the sort of authority we are attempting to ground cannot be so easily escaped. As David Velleman has remarked, "no amount of rethinking"—or, for that matter, *rewilling*—can make genuinely authoritative reasons "irrelevant."[80] We are in the market for reasons whose authority outstrips our own. "Our power to construct [such a] framework of reasons would therefore undermine the whole point of having one."[81]

Korsgaard naturally disagrees, and in a chapter entitled "The Authority of Reflection" she argues that reasons for action acquire genuine normative authority through a process of reflective endorsement. In the following sections, I shall attempt to show that this argument fails and, moreover, that it fails for reasons which threaten to undermine the any autonomist internalism with foundational ambitions.

---

78    Korsgaard, *The Sources of Normativity*, 93.

79    For a different way of stating this objection to Korsgaard, see Wallace, "Normativity, Commitment, and Instrumental Reason." Wallace draws a distinction between subjectively normative states and objectively normative states (85). Parfit's objection, then, is that Korsgaard's theory of normativity is actually a theory only of subjective normative states.

80    Velleman, "A Brief Introduction to Kantian Ethics," 32.

81    Velleman, "A Brief Introduction to Kantian Ethics," 33. Compare Cohon, "The Roots of Reasons": "What I do not see is how commanding myself, or even willing, makes a reason of a consideration … that was not a reason before" (68).

## The Categorical Imperative

Brandishing her Kantian credentials, Korsgaard locates the source of normative authority in the categorical imperative. Her argument, following Kant, begins from the premise that the will is autonomous. According to Korsgaard, the will "must have its *own* law or principle."[82] This principle cannot be imposed on the will from outside, for then the will would not be autonomous. Moreover, if the will's principle is to serve as the basis of normativity—that is, if this principle is to halt the regress of justification—it cannot be what Korsgaard calls an *external* principle or law.[83] It must not be a law that requires further justification. It must, in other words, be immune to questions about whether we have reason to follow it. For, as we saw in the introduction, if we could coherently ask what reasons we have to follow this basic principle, we would have to find an even more fundamental principle in order to identify such reasons. That is, we would need some independent standard of justification in order to demonstrate that our basic principle is justified. The principle of the will must, therefore, be an *internal* principle, a principle or law that we can derive from the very nature of autonomous willing.

What sort of law can the will make for itself? Korsgaard's answer is that if nothing but the will determines what this law must be, then "*all that it has to be is a law*."[84] The principle of the will cannot be a more specific or substantive law, for the will does not yet have a reason to choose such a law. The will requires *its own principle* in order to determine what does and does not count as a reason. Hence, the only way that the will can make a law for itself is to make a perfectly formal and general law. The resultant principle is simple: choose a principle (and act on it). And this, Korsgaard observes, is nothing more than the basic formulation of Kant's categorical imperative.[85] The categorical imperative, then, is the authoritative norm governing action. It is not an external norm, since we cannot raise normative doubts about it. If

---

82    Korsgaard, *The Sources of Normativity*, 98.

83    For more on her distinction between *internal* and *external* norms, see Korsgaard, "The Normativity of Instrumental Reason," 249.

84    Korsgaard, *The Sources of Normativity*, 98.

85    For Korsgaard's most complete presentation of the argument in this paragraph, see Korsgaard, "Morality and Freedom," 163–7.

we ask why we should act according to a law or principle, the answer is that we have no choice: choosing a principle is *constitutive* of action.

As Korsgaard acknowledges, though, the categorical imperative is also a purely *formal* principle.[86] It prescribes nothing about the substance or content of our principles of action. This presents a problem for Korsgaard's view: an agent satisfies action's standard of correctness merely by acting. "To be rational," Korsgaard notes, "*just is* to be autonomous."[87] The same goes for morality: "the moral law *just is* the law of an autonomous will."[88] And it therefore appears to be impossible to distinguish the categories of autonomous action and correct (justified, rational) action. Put another way, there can be no autonomous but incorrect (unjustified, irrational, immoral) actions. As autonomous agents, we cannot help but act rationally. For us, the norm governing action is inviolable. When we act for a reason we cannot help but act for a good one. Yet, as we saw above, the possibility of error—of wrong or bad or irrational action—is central to our concept of normativity, a concept Korsgaard clearly shares:

> There is no normativity if you cannot be wrong.[89]

> How can there be rational action, in any sense, if there is no irrational action? How can there be an imperative which no one ever actually violates?[90]

> How can you be guided by a principle when anything you do counts as following it?[91]

And so if the categorical imperative is inviolable, it cannot constitute the source of normativity—even by Korsgaard's own lights.

Elsewhere, of course, Korsgaard denies that her theory has this implication: "Certainly I am not saying that reflective endorsement—I mean the bare *fact* of reflective endorsement—is enough to make an action right."[92] She repeatedly insists that she *can* distinguish between

---

86     See Korsgaard, *The Sources of Normativity*, 99–100, 103.

87     Korsgaard, "The Normativity of Instrumental Reason," 219, emphasis in original.

88     Korsgaard, "The Normativity of Instrumental Reason," 249, emphasis in original. Compare Wallace, "The Normativity of the Will," 77.

89     Korsgaard, *The Sources of Normativity*, 161.

90     Korsgaard, "The Normativity of Instrumental Reason," 228.

91     Korsgaard, "The Normativity of Instrumental Reason," 229.

92     Korsgaard, *The Sources of Normativity*, 161, emphasis in original. In her reply to Cohen's Mafioso example ("Reason, Humanity, and the Moral Law," 183–4), however, Korsgaard seems to acknowledge the problem once again (*The Sources of Normativity*, 256–8).

correct actions and merely autonomous actions. I shall not, however, discuss the details of her various attempts to draw this distinction. Instead I shall argue that any such attempt is bound to fail, given Korsgaard's account of the sources of normativity.

In order for Korsgaard to make room for the possibility of practical error, she must provide some additional standard of correctness for action—a norm that not all autonomous actions meet. This norm cannot follow directly from the categorical imperative, for if it did, then it too would be inviolable. It must, in other words, be an external norm. As we saw above, however, we can always ask what reason we have to adhere to an external norm. And so in order to demonstrate that this new, external standard is justified, we must appeal to a more fundamental norm—a standard that is not itself susceptible to normative doubt. The categorical imperative is, according to Korsgaard, the only such standard. But the categorical imperative cannot possibly justify an external norm for action. It is purely formal; it demands only that we act in accordance with *some* law or principle. Under the categorical imperative, *any* autonomous action is justified; and thus we can never derive from the categorical imperative a normative standard for action according to which only some actions are correct.[93]

Since autonomy is supposed to be the source of normativity and obligation, any action-governing norm that is not linked to autonomy (via the categorical imperative) will neither generate proper normative reasons for action nor yield genuine obligations, unless an agent happens to make that norm his principle of action. Now Korsgaard faces a dilemma. On the one hand, if she presents a substantive rational standard that autonomous agents are capable of violating, then she cannot demonstrate that this principle is authoritative. It will be authoritative only for those agents who choose to make it their principle of action. On the other hand, any standard of rationality that derives its authority from the categorical imperative (as any normative standard must) will fail to provide a substantive norm that autonomous agents can actually violate. Korsgaard is stuck either with authoritative formality or unauthoritative substance. She can capture either the normativity or the substance of rationality (or morality), but not both.[94]

93   Compare Wallace, "Normativity and the Will," 78–9.

94   This might seem to pose a problem for normativity in general. If Korsgaard is correct that only an inviolable formal norm can serve as the source of normativity, and if I am right that a formal norm can never generate a substantive normative standard, then there is little hope for substantive normativity. Fortunately, as I shall argue in chapter 4, Korsgaard is not correct.

Korsgaard later denies that the categorical imperative is meant to encompass all of our reasons and obligations. In particular, she reports, it does not give us "the whole content of morality."[95] And so Korsgaard develops a second argument which purports to show that we are bound not only by the categorical imperative but also by *the moral law*, which resembles Kant's formula of humanity.[96] Crucially, the moral law as Korsgaard conceives it is *not* inviolable: we are all too capable of treating others merely as means to our own ends. If this second argument succeeds, then, Korsgaard will have identified a norm that is both substantive and internal (and hence authoritative). The problem I have raised will therefore have been solved.[97]

This second argument revolves around Korsgaard's claim that as agents we must value our own humanity. From there she reasons to the conclusion that we must value everyone else's humanity as well. Lo and behold: "Enlightenment morality is true."[98] I shall not consider the full argument here.[99] It involves a number of steps, each of which has already received plenty of criticism.[100] Rather I shall try to show that any such argument cannot succeed, at least given Korsgaard's meager premises.

Consider the form of the crucial intermediate conclusion: we *must* value our own humanity. The argument turns on the meaning of this "must." In what sense *must* we value our own humanity? What sort of necessity is at stake here? The necessity that attaches to the categorical imperative is fairly strong. When we act we *must* act on a principle; if we do not act on a principle then we do not act at all. Autonomous action without self-legislation is impossible. Notice that the strength of this necessity is what makes the categorical imperative a strong candidate for the source of normativity. It does not make sense to challenge the categorical imperative—or to ask why we should abide by this supposedly fundamental

95    Korsgaard, *The Sources of Normativity*, 99.

96    See Kant, *Groundwork of the Metaphysics of Morals*, 4:429, here and hereafter cited by the volume and then page number of the Germany Academy edition.

97    I am indebted to Ernesto Garcia for making this point on Korsgaard's behalf and to Sharon Street for helpful comments about Korsgaard's argument.

98    Korsgaard, *The Sources of Normativity*, 123.

99    See Korsgaard, *The Sources of Normativity*, 100–5, 120–5.

100   See, for instance, the four commentaries included in *The Sources of Normativity* (especially Cohen, "Reason, Humanity, and the Moral Law"; and Nagel, "Universality and the Reflective Self"), as well as Gibbard, "Morality as Consistency in Living"; Smith, "Search for the Source;" Cohon, "The Roots of Reasons," 75–85; and Bratman, review of *The Sources of Normativity*.

norm—*because acting on a principle just is acting autonomously*. The necessity of self-legislation is what underlies the categorical imperative's resistance to calls for further justification.

Return now to the intermediate conclusion that is supposed to establish the moral law: we *must* value our own humanity. If the necessity here is as strong as it is in the categorical imperative, then the moral law will also be inviolable. And so there must be a weaker sort of necessity at work here. This is where Korsgaard introduces the idea of "rational necessity."[101] We are *rationally compelled* to value our own humanity, much as we are rationally compelled not to believe both *p* and not-*p*. If the necessity at work in the moral law is only rational necessity, then this law is not inviolable, for we flout rational necessity whenever we knowingly act against our better judgment. But rational necessity is just normative necessity. So it turns out that the crucial intermediate conclusion is that we *ought* to value our own humanity. And now Korsgaard must explain what underwrites the authority of that 'ought'. What sort of reason do we have to value our own humanity? It could not be an epistemic reason, since valuing—for Korsgaard—is not a cognitive or theoretical activity. Recall that Korsgaard is an autonomist internalist, and so the source of obligation is supposed to reside in pure *practical* reason. A reason to value must be a practical reason. But if the 'ought' in the moral law is a practical 'ought', then we can ask *why* we ought to value our own humanity. I can acknowledge that the moral law demands it, but ask: *why should I follow the moral law?*

At this point Korsgaard has only two options, at least that I can see. First, she can acknowledge that the moral law is external, in which case she must either give up her claim that it is authoritative or give up her internalism. Second, she can try to argue that the moral law is an internal law because it follows from the categorical imperative, which is itself internal. The second is clearly Korsgaard's preferred alternative. If anything, she is even more averse to the "bluff and brow-beating" of the externalist approach than Williams is. But now the objection I pressed above comes back into the play. The moral law is substantive, and so it cannot follow from the merely formal categorical imperative. The categorical imperative simply directs us to act: Do something! Act for a reason! A more substantive norm will never follow from a standard that can be satisfied merely by acting. What makes this even more apparent is Korsgaard's claim that the categorical imperative is meant to be not merely formal but also *analytic*. The conclusion that autonomous action requires us to choose some principle under

101    See "Internalism and the Sources of Normativity," 59, as well as note 75 above.

which to act is meant to follow *from the very concept* of deliberative, autonomous agency.[102] Yet the moral law and the intermediate conclusion that we must value our own humanity are both clearly *synthetic*. How would we derive a synthetic conclusion about what we ought to value from this sort of conceptual analysis? Analytic premises cannot by themselves yield a synthetic conclusion. We need to add a synthetic premise in order to make the sort of implication Korsgaard wants even possible. But her account has no additional premise to offer. It is the categorical imperative—and the categorical imperative alone—that serves as the basis of normative authority:

> The reflective structure of human consciousness requires that you identify yourself with some law or principle which will govern your choices. It requires you to be a law to yourself. *And that is the source of normativity*.[103]

If the categorical imperative is the source of normative authority, then the moral law cannot be both substantive and normative. Nothing substantive follows from the categorical imperative, and so no substantive laws are authoritative.

## Lessons Learned

Autonomist internalists start off with a step in the right direction when they reject the hydraulic conception of agency underlying empirical naturalist internalism in favor of a more sophisticated account, one that does justice to the various forms practical reasoning can take. This move enables them to capture the full practical force of judgments about reasons. If reasons just are the considerations that survive autonomous reflection and deliberation, then their practical significance is unavoidable.

As I have tried to demonstrate, though, autonomist internalists follow this new direction a little too zealously and run into trouble when they attempt to translate their account of the practical significance of reasons into a theory of reasons' normative *authority*. Autonomists like Korsgaard hope to ground normativity in the nature of autonomous willing. But there is a snag, for although autonomy requires that we rule ourselves authoritatively, it also requires that we rule ourselves *absolutely*: when we are acting autonomously, what we say goes. There is no

---

102    In "Morality and Freedom," Korsgaard approvingly quotes Kant's *Groundwork*: "if freedom of the will is presupposed, morality together with its principle follows from it by the mere analysis of its concept" (161; quoting Kant, *Groundwork of the Metaphysics of Morals*, 4:447).

103    Korsgaard, *The Sources of Normativity*, 104, emphasis added.

transgressing our self-legislation, for as soon as we decide to act differently, we thereby overturn the old law in favor of a new one. Whatever we do, we are following our own instructions when we do it. We cannot help but follow the law of our autonomous will.

This problem stems in part from the fact that there is no recognizable difference between acting for a good reason and acting for bad one, at least from the perspective of autonomously willing agents. The first-person perspective does not and cannot discriminate between correct and incorrect actions. Put another way, from the agent's point of view, *all* actions seem correct precisely in virtue of being autonomous. What this shows is that autonomist internalism—whether it relies on Korsgaard's categorical imperative or some other norm derived from the nature of autonomy—cannot capture genuine authority while treating normativity as a purely first-person phenomenon. Now, this does not demonstrate that the autonomist internalists are wrong to situate normativity within autonomous agency. Instead what it shows is that they need more than a merely first-person characterization of autonomy. Autonomist internalism will be able to ground genuine normative authority only if it can accommodate the fact that our first-person take on our reasons at least sometimes contradicts the way things actually are.[104]

Although empirical naturalists steer clear of this particular problem, they drive too far in the opposite direction. As Darwall observers, they too "situate normativity within agency."[105] But what separates them from the autonomist internalists are their radically different accounts of the nature of agency and action for a reason. The empirical naturalist conception of agency does not allow for pure practical reasoning, and so—as we have seen—it reduces practical reasoning to a form of theoretical reasoning. Above I suggested that this is fatal to the empirical naturalist project. Whatever practical authority our epistemic reasons possess must ultimately derive from our motives, which—according to the empirical naturalists—are actually running the show. But there is nothing authoritative about motivational force. Action for a reason is no more inherently authoritative on the empirical naturalist view than it is on the autonomist view. Both sorts of internalist seem to assume that an account of normative authority will simply fall out of their respective accounts of governance. Yet, as we have seen, there is nothing particularly authoritative about wanting something, or even wanting something for its own sake. Nor is there anything especially authoritative about autonomously taking something to be a reason.

104    Compare Velleman, "A Brief Introduction to Kantian Ethics," 32–3.
105    Darwall, "Internalism and Agency," 169.

With all of the internalists' emphasis on governance, they seem to have misplaced the authority, just as the externalists in chapter 2 predicted.

Where does that leave our search for the foundations of ethics? Externalism runs afoul of the *practicality* of reasons, whereas internalism appears to be inconsistent with the *authority* of reasons. Is there a way to capture both? I believe there is. In chapter 4 I shall present and defend a theory of reasons for action that does justice to both their practicality and their authority. By way of preview, let me make one final observation about internalism.

Empirical naturalist internalism is not as dead as it seems. The open question objections I rehearsed above are fatal not to the approach in general but rather only to the specific views I considered. This is because any empirical naturalist internalism will only be as plausible as the conception of agency it employs. The real problem with Williams and Brandt—what ultimately lends the open question objections bite—is that they are working with an instrumentalist conception of agency. Williams and Brandt hold that our motives exert genuinely authoritative force only because they believe that practical deliberation is ultimately a matter of pursuing goals and satisfying desires that, at least from the practical perspective, must be taken as given.[106] And it is no wonder that such an impoverished conception of agency and practical reasoning would yield such a barren theory of reasons for action.

This prompts Darwall to conclude pessimistically:

> Once the empirical naturalist conception of reason is accepted, any accounts of normativity it can ground must be frankly *revisionist*. Neither instrumental nor calm reflective deliberation theories can capture the idea of a 'dictate of reason,' nor should they attempt to. For the empirical naturalist, reason cannot strictly *dictate* or counsel anything; it is not inherently practical.[107]

This conclusions seems correct, at least regarding the empirical naturalist internalists Darwall actually discusses, namely, Hobbes, Locke, and Cumberland. Neither an instrumental account of agency (defended by Hobbes, Williams, and Brandt, among others) nor a calm reflective deliberation theory of agency (exemplified by Cumberland) leaves room for genuine practical reasoning. However, I believe it would rash to regard these failures as a sign that *any* empirical naturalist account of normativity will similarly fail to accommodate genuinely practical reasoning. I think that an empirical naturalist internalism informed by the virtues

---

106    See Brandt, *A Theory of the Good and the Right*, ch. 5.

107    Darwall, *The British Moralists and the Internal 'Ought'*, 320, emphasis in original.

and drawbacks of both externalism and autonomist internalism can capture the practicality *and* the authority of reasons. All such a theory requires is a *naturalist conception of autonomous agency*—a third-person account of agency that nevertheless does justice both to the fact that we are autonomous agents and to the distinctive phenomenology of action for a reason.

An empirical naturalist internalism along these lines will still close normative questions that seem wide open to us. Yet the naturalist conception of autonomy will enable us to explain *why* these questions *seem* so open: it is only from the practical standpoint of an autonomous agent that the questions are really closed. Moreover, it is only the practical standpoint that matters when it comes to ethics. If the naturalist account of autonomous agency is a convincing one, then this *autonomist empirical naturalist internalism* will have all of the resources it needs to meet the open question challenge, identify genuinely authoritative norms, and thereby establish the foundations of ethics.

# CHAPTER 4

## CONSTITUTIVISM

> Those who make a practice of comparing human actions are never so perplexed
> as when they try to see them as a whole and in the same light; for they commonly
> contradict each other so strangely that it seems impossible that they have come from
> the same shop.
>
> —Montaigne, "Of the Inconsistency of Our Actions"

The opposition between externalism and internalism presents us with an unpalatable choice. The shortcomings of externalism are evident and severe. It relies on the deeply unsatisfying idea of an authority that lacks the capacity to govern. Moreover, and perhaps more seriously, it presupposes a conception of normativity that divorces normative questions from practical questions—that is, from questions about what to do. Unfortunately, the internalist alternatives are equally unappetizing. The purported internal connection between reasons and motives (or between reasons and the will) leaves no room for those reasons to be genuinely authoritative. The project of finding the foundations of ethics therefore faces a dilemma, and on neither horn do reasons possess the sort of authority that can put a stop to the regress of justification.[1]

These two views apply competing pressures to any account of the normative authority of reasons for action. Although neither view can be the whole story when it comes to normativity, each places a constraint on what that story must look like. Internalists insist that reasons can be entirely external to our motives and commitments only on pain of being utterly irrelevant to practical deliberation. But externalists warn us about the dangers of associating reasons and motives too closely. If reasons arise from motives or the will, it is difficult to see how we can act contrary to our reasons. The norms of practical reason become inviolable. We are therefore left with a puzzle: normative reasons for action must be internal—but not so internal that they

---

1    Compare Bond, *Reason and Value*: "either objective reasons exist but lack the power to motivate, which seems absurd, or reasons are not objective but merely a function of an agent's actual desires" (7).

lose their standing as authoritative. In a recent paper, David Enoch succinctly describes where our project stands: "the normative, though somehow essentially tied to us and our desires (or at least our will), can nevertheless not be as arbitrary as our (usual) motivations."[2] Our challenge is to arrive at a conception of reasons for action that can do justice to both aspects of their normativity: the authority emphasized by the externalists, and the governance or guidance central to the internalist approach.

I suspect that one problem with the views we have considered thus far is that they initially tackle the issue with only one of these two facets of normativity in mind. Externalists are almost exclusively concerned with finding a robust and independent authority—one that can ground genuine correctness conditions for action. Any account of normative governance is merely tacked on as an afterthought. Internalists make the opposite mistake. They are so focused on the psychological force of reasons that they all but ignore questions about their authority. But one of the lessons of the previous two chapters is that authority and the capacity to govern are not two independent features of reasons. Rather, authority and governance are two sides of the same normative coin. And so as we go in search of a more satisfactory account of reasons for action, we should be on the lookout for reasons that govern in virtue of their authority and that are authoritative partly in virtue of their capacity to govern. In other words, rather than study these two aspects of normativity separately, we should endeavor to develop a view in which they arise together.

One promising avenue of investigation centers on the phenomenon of action for a reason. Here we have normative guidance at work—authority and governance working hand in hand. But it is not immediately clear how an inquiry into the nature of normative guidance can ever by itself provide all of the materials for a theory of practical normativity. After all, action for a reason is not always action for a *good* reason. We act for bad reasons as well as good ones, and normative guidance presumably works the same way in both cases. The difference between action for good reasons and action for bad ones lies not in the way in which we are guided or governed by our reasons, but rather in the *authority* of those reasons. Good reasons are authoritative; bad reasons are not. Since not all instances of normative guidance involve the presence of genuine authority, it is not clear how an approach that focuses solely on normative guidance and action for a reason can produce a satisfying conception of normativity. It seems

2    Enoch, "Agency, Shmagency," 178.

that any such approach will inevitably suffer from an inability to capture the authority of reasons—the same inability that afflicts internalism.

In the face of these worries, Peter Railton, Christine Korsgaard, and David Velleman have proposed a way forward.[3] Working independently, these three have suggested that one way to find the elusive authority and thereby halt the regress of practical justification is to locate a norm or standard of correctness that is somehow *constitutive* of action for a reason. Following others, I shall call this the *constitutivist* strategy. Constitutivism is the topic of this final chapter. I shall argue that in light of the failures of internalism and externalism exposed in the previous two chapters, constitutivism is our best and indeed our only hope for providing a foundation for ethics—and for normativity more generally.

### The Constitutivist Approach

To demonstrate how the constitutivist strategy is supposed to stop the regress of justification in ethics, it may be helpful to look first—as Railton and Velleman do—at the case of normative epistemology, where a similarly problematic regress looms. As we shall soon see, constitutivism seems like a natural response to the regress of epistemic justification. And if there are sufficient parallels between ethics and epistemology, we may be able to adopt a similar approach to the regress of practical justification.

Recall that the regress of justification arises in the practical case because any norm or standard to which we might appeal in order to justify our actions itself requires justification. The authority of every rule seems to depend on the authority of some other rule standing behind or under it. A similar problem arises for the justification of beliefs. Any norm with respect to which our beliefs are warranted must itself be warranted or supported by some other norm, and the ensuing regress threatens to undermine any attempt to provide a foundation for normative epistemology.

This is where constitutivism enters the picture. Railton and Velleman both suggest that belief is governed by a constitutive norm, and that such a norm can put a stop to the

---

3    See especially Railton, "In Search of Nonsubjective Reasons" and "On the Hypothetical and Non-hypothetical"; Korsgaard, "Self-Constitution"; and Velleman, "The Possibility of Practical Reason" and the introduction to *The Possibility of Practical Reason*.

     Although Railton, Velleman, and Korsgaard were the trailblazers, others have since followed along. See, for example, Rosati, "Agency and the Open Question Argument"; and O'Hagen, "Practical Identity and the Constitution of Agency" and "Belief, Normativity and the Constitution of Agency."

regress of epistemic justification. Now, generally speaking, to say that some kind of thing has a constitutive norm is to say that part of what makes something a member of that kind is its being subject to that norm.[4] Being a member of that kind at least partly consists in falling under the norm in question. An attitude or activity's constitutive norm must therefore somehow be built into the very nature or concept of the attitude or activity in question. Now it just so happens that belief seems to have such a norm: part of what makes an attitude a belief is that it is correct if and only if it is true. Call this *the truth norm*. To say that the truth norm is belief's constitutive standard of correctness is to say that being a belief at least partly consists in being subject to this standard.[5]

How can this fact about belief help us halt the regress of epistemic justification? The regress is driven by the worry that a norm can be justified only if it is supported by some other justified norm. But, the constitutivist argument goes, the authority of a *constitutive* norm does not depend on the authority of other norms in this way. Its authority derives from the nature or concept of the attitude or activity it governs. Why should we believe the truth? Because that is just what beliefs are: attitudes that are governed by the truth norm! We can answer this demand for justification without appealing to another norm. Instead we defer to the nature of belief, which settles the matter conclusively. For if some attitude is not governed by this norm, then it cannot be a belief. Any challenge to the authority of the truth norm over belief thus betrays some misunderstanding of the nature or concept of belief. Fully comprehending the nature or notion of belief requires accepting that beliefs are correct if and only if they are true. And so once we comprehend the nature of belief, the question of whether to believe what is true ceases to be an open question. The truth norm puts a stop to the regress of epistemic justification, and it can therefore provide a foundation for normative epistemology.

Whether and how the truth norm is actually the constitutive standard of correctness for belief is not my immediate concern. I have discussed the case of belief only to illustrate how the constitutivist strategy is supposed to work—how it is supposed to establish a foundation

---

4   As Korsgaard puts it, constitutive norms are "standards that apply to a thing simply in virtue of its being the kind of thing that it is" ("Self-Constitution," 1.21, here and hereafter cited by lecture number and then page number).

5   This is often stated metaphorically as the claim that belief constitutively "aims at" the truth. See Williams, "Deciding to Believe"; and Velleman, "The Aim of Belief." I shall return to the notion of *constitutive aims* below.

for epistemic normativity. What interests me is the practical case. Railton, Velleman, and Korsgaard all propose that we adopt (or at least consider) a constitutivist strategy for establishing the foundations of *ethics*.[6] Their hope is that a constitutive norm of *action*—a norm or standard of correctness somehow built into the nature or concept of action—can halt the regress of practical justification just as the truth norm promises to arrest the regress of epistemic justification. If there is a constitutive norm for action, then it would not make sense to question whether one should act in accordance with it, since action just is (at least in part) behavior that falls under this standard. Whether one should act in accordance with this norm would be an open question only to someone who does not fully grasp the nature or concept of action. As with belief's constitutive standard, the authority of such a norm would derive not from any other norm, but from the very nature or concept of the activity or attitude it governs, in this case, action.

As Velleman has observed, this is a call for us to seek the foundations of ethics in the philosophy of action.[7] It requires us to investigate the conditions of action and agency in the hope of revealing that all actions have something crucial in common, namely, that they are all governed by a single norm or standard. What the constitutivist must find, in other words, is a norm one must be subject to in order to count as an agent—a norm whose authority is part of the price of admission into agency.[8] For this strategy to succeed, its proponents must show that Montaigne—as quoted at the beginning of this chapter—is mistaken: although actions may indeed "contradict each other so strangely" that they make the practice of comparing them "perplexed," they all "come from the same shop" nonetheless.

But this is only the first phase of the constitutivist project. After showing that action actually has a constitutive norm and identifying what this norm is, the constitutivists must then demonstrate that the foundations provided by such a standard are strong enough to support a genuinely normative ethics. I shall consider these two phases in turn.

---

6    Unlike the others, Railton has serious misgivings about how much this strategy can accomplish. See "On the Hypothetical and Non-hypothetical," 309–19.

7    Velleman, "Deciding How to Decide," 233.

8    Compare Railton, "In Search of Nonsubjective Reasons," 130.

## Korsgaard's Constitutive Standards

There are, broadly speaking, two different ways in which one might try to find a constitutive norm in the philosophy of action, reflecting two different ways in which a norm can be constitutive. The first approach—adopted by Christine Korsgaard in her Locke Lectures—sees constitutive norms as somewhat like Platonic or Aristotelian *forms*. The second—explored by David Velleman in a series of papers—attempts to derive constitutive norms from constitutive *aims*. Both approaches face obstacles; but as we shall soon see, only Velleman's focus on constitutive aims offers any hope of providing a foundation for ethics. Korsgaard's account founders, but it repays close examination nonetheless. The shoals on which she runs aground are treacherous, and any successful constitutivist account must steer well clear of them.

The easiest path into Korsgaard's view of constitutive norms begins with her discussion of houses.[9] Following Aristotle's theory of the form and matter of artifacts, Korsgaard contends that membership in a kind is determined by the candidate members' teleological organization. Belonging to a particular kind consists simply in having the appropriate teleological arrangement. Something counts as a house, for instance, if and only if it is organized so that it can serve as a habitable shelter. Aristotle called this the "form" of a house. The idea is that being a house just *consists in* being a habitable shelter; the latter is a *constitutive condition* of the former. This is where Korsgaard's talk of *norms* and *standards* enters the picture. She moves directly from the idea of a constitutive teleological condition to the notion of a constitutive norm. A thing's constitutive "teleological organization" simply "gives rise" to "normative standards," which Korsgaard calls "constitutive standards" or norms.[10] This transition from talk of forms to talk of norms is licensed, according to Korsgaard, by the way in which a kind's constitutive teleological organization can definitively answer normative questions and thereby halt a regress of justification. Why should I add a roof to the house I am building? Because then it will be able to serve as a habitable shelter. Why should my planned house be able to serve as a habitable shelter? Because that is simply what houses are. To ask why a house should be a habitable shelter is to betray a lack of understanding of the nature of houses. Such a question just "does not make sense."[11]

9     See Korsgaard, "Self-Constitution," 1.19–25.

10    Korsgaard, "Self-Constitution," 1.20–1.

11    Korsgaard, "Self-Constitution," 1.21.

Korsgaard suggests that we can draw a similar conclusion with regard to action. "Action," she claims, "is self-constitution."[12] It is activity whereby one constitutes oneself. This is what action consists in—what it is for activity to count as action. And from the fact that self-constitution is the constitutive condition (or form) of action it follows that it is also its constitutive standard (or norm). Just as it does not make sense to question why a house should be able to provide shelter, so it does not make sense to ask why an action should constitute its agent. That is just what actions do! Once we understand the nature of action as self-constitution, there is no longer any sensible question as to why one should act so as to constitute oneself. The nature of action settles the matter; no further justification is required—or even possible.

I have no interest in challenging the particulars of Korsgaard's philosophy of action. My concern here is not with her specific claim as to the constitutive conditions of agency (namely, that action is self-constitution), but rather with her more general view that this sort of constitutive standard can provide a foundation or source for practical normativity. Her argument for the normativity of action's constitutive norm mirrors her discussion of the constitutive standard of houses, and so I shall again start there.

The constitutive norm of houses may indeed stop the regress of justification when it comes to questions about houses, but Korsgaard is aware that this is not enough to show that we fall under the authority of this standard. After all, why should we care about houses? From the fact that houses are subject to a particular standard it does not follow that *we* are subject to it as well. This is just another external norm of the sort we discussed and eventually dismissed in chapter 2. And so the constitutive standard for houses does not ground or generate any normativity by itself. Taken in isolation it has no normative or practical force for us at all. It is nothing to us—normative for houses, not for people. If we adopt some house-related aim, though, the standard acquires some normative force. If, for example, I plan to build a house, the constitutive standard for houses can guide my execution of that plan. What began as an irrelevant, external norm is now internal and hence normative. It is only when we set for ourselves a house-related goal or project, then, that the constitutive norm of houses becomes normative *for us*. In the absence of such a goal, this standard has no authority over us at all.

12    Korsgaard, "Self-Constitution," 1.17.

If the normativity of the constitutive standard of houses is dependent on our having house-related aims, what about the constitutive norm of action? Korsgaard insists that this norm lies at the foundation of ethics; it is supposed to be the source of normative authority and of all genuine reasons for action. Now, this norm may halt the regress of justification when it comes to actions: it may indeed be pointless to ask why action should be an instance of self-constitution. If Korsgaard is correct, then that is just what action is. But by itself this standard has no more authority over *us* than the constitutive standard of houses has in isolation. Just as we can reasonably ask why we should care about houses, so can we wonder why we should bother with agency and action. This raises a pressing question. In order for the constitutive norm of action to be the source of genuine reasons for action, its normativity cannot depend on some contingent project or aim. What then makes the constitutive norm of action normative for us? Whence does its authority over us (or even its relevance) arise? To answer this question, Korsgaard falls back on the necessity of action. "Human beings are *condemned* to choice and action," she proclaims at the outset of the first lecture.[13] Agency is not some contingent aim we can simply take or leave. Rather, it is "our plight: the simple inexorable fact of the human condition."[14] And if we cannot escape action, then we cannot escape the binding force of the norm that constitutively governs it. It is, therefore, the necessity of action that brings us under the authority of action's constitutive standard. "The principles of practical reason are normative for us … simply because we *must* act."[15]

This brings us to the first problem with—or at least disadvantage of—Korsgaard's brand of constitutivism. Korsgaard treats the necessity of action as something approaching a truism, but it is hardly obvious that we are condemned to act, at least in the sense she intends. Buddhism is predicated on the possibility that we can function in the world without constituting ourselves.[16] Anyone who reaches a state of Buddhist enlightenment thereby falls outside the scope of reasons for action. Reasons become nothing to him, just as the constitutive standard for houses is nothing to a person with no house-related aims. The difficulty of

13    Korsgaard, "Self-Constitution," 1.1.

14    Korsgaard, "Self-Constitution," 1.2.

15    Korsgaard, "Self-Constitution," 1.17 (emphasis added). In class on September 15, 2004, Velleman helpfully labeled Korsgaard's approach the "Gotta Dance! theory of normativity."

16    Of course Buddhists believe that self-constituting action is not only avoidable but also undesirable.

demonstrating the necessity of self-constituting action strikes me as a weakness in Korsgaard's constitutivist approach, but I do not want to explore it further here, for there is a second and far more serious problem to consider.[17]

Korsgaard refers explicitly to the Aristotelian notion of form, but the constitutive norms she envisions also function like Platonic forms.[18] For both Plato and Aristotle, membership in a category or kind requires conformity to that category's form—or what Korsgaard calls its constitutive norm. In the case of action, membership comes in degrees, and the degree of membership is proportionate to the degree of conformity. "An action that is less successful at constituting its agent is to that extent less of an action."[19] But this conception of action's constitutive standard leaves no room for the possibility of *bad* or *incorrect* actions—actions that run afoul of the standard. Activity counts as action only to the extent that it conforms to the self-constitution standard. Thus if something fails to meet that standard, it is not a bad or incorrect action. *It is not an action at all.* To the extent that activity fails to meet action's constitutive norm, it simply falls outside of the category that this norm delineates.

This problem is not specific to the particular constitutive norm Korsgaard identifies (namely, self-constitution). It arises instead from her general conception of constitutive standards as Platonic or Aristotelian forms. On this conception, "every object and activity is defined by certain standards that are *both constitutive of it and normative for it.*"[20] There is no room for bad actions because what determines whether something is an action is *the very same standard* as what determines whether something is a *good* action. The same goes for houses— and for anything else governed by this sort of constitutive standard. If what makes something a good house is just what makes it a house, then the worse something is as a house, the less of a house it is.

Korsgaard, of course, denies that this conclusion follows. She offers the following recipe to distinguish good actions from bad ones: "What makes actions good or bad is how well they constitute you. People are more or less successful at constituting their identities, and a good

17    Compare Enoch, "Agency, Shmagency," 187–92.

18    Thanks to David Velleman for suggesting this comparison.

19    Korsgaard, "Self-Constitution," 1.17.

20    Korsgaard, "Self-Constitution," 1.24 (emphasis added).

action is one that does this well."[21] This sounds like a promising solution to the problem only because it trades on the ambiguity of the crucial word "well." On one understanding of this word, to constitute oneself well is just to do so successfully. But this could not be the basis of an account of bad or incorrect action, since to constitute oneself unsuccessfully is not to constitute oneself at all (and hence not to act at all). On the second understanding of "well," to constitute oneself well is to do so in accordance with some additional, independent norm or standard of evaluation. This *does* allow for the possibility of bad action, but now it is the additional norm—and not the constitutive standard—pulling all of the normative weight. Moreover, since the additional norm is not constitutive of action, it cannot halt the regress of justification and serve as a foundation for ethics. It stands in as much need of justification as any other external standard for action.

As long as Korsgaard holds on to her conception of constitutive norms, there can be no solution to this problem. Either the standard of good action is also the constitutive standard of action in general (in which case bad action is impossible), or we judge the goodness or badness of actions according to an independent standard (in which case bad action is possible but only because we have appealed to a norm that is not constitutive of action and must therefore itself be justified).

This is no minor failing. As I noted above in chapter 3, the possibility of error is close to an axiom of metaethics and is central even to Korsgaard's notion of normativity.[22] To see why, consider our path to this point. The motivation behind constitutivism is the desire to find norms whose authority cannot sensibly be questioned. The hope is that these norms can ground legitimate and authoritative criticism of agents who ignore their reasons for action. But, as Peter Railton has observed, constitutive accounts like Korsgaard's have the "unintended effect of pulling the claws of the very criticism" they are supposed to ground.[23] To the extent that her version of action's constitutive standard licenses criticism, it also ensures there will be no agents worth criticizing. Anyone who does not meet the standard fails to count as an agent and hence is not a proper target of rational criticism. Korsgaard could at this point simply

21    Korsgaard, "Self-Constitution," 1.17. See also "Self-Constitution," 5.14–24.

22    See page 136 in chapter 3 above.

23    Railton, "On the Hypothetical and Non-hypothetical," 309.

dismiss such a person as "not one of us," but—as Railton rightly insists—"this is xenophobia, not criticism."[24]

Korsgaard's constitutivism founders on her conception of constitutive norms. The Platonic and Aristotelian model does not allow for the possibility of error, and thus it cannot provide the basis for a constitutivist account of the foundations of normativity. Railton contends that "a similar problem confronts *all* constitutive arguments."[25] However, I shall presently show that there is a brand of constitutivism that successfully steers clear of the problems that sink Korsgaard's account. Whether it can avoid disaster altogether is another matter, though.

## Velleman's Constitutive Aims

David Velleman defends a version of constitutivism based on his view that action has its own constitutive aim. Action, according to Velleman, constitutively aims at integrative self-knowledge. This is a teleological claim, and from it Velleman argues to the conclusion that self-knowledge is also the constitutive *norm* or *standard of correctness* for action.[26] Call this *the self-knowledge norm*. Even without presenting the details or evaluating the cogency of Velleman's account, we can see how his approach makes room for the possibility of incorrect or irrational action.

On Velleman's view, the constitutive *norm* of action is different from the constitutive *condition* from which it is derived. Autonomous action is activity driven by the aim to know what one is doing. This is what constitutes action. Notice, however, that the constitutive condition of action does not require that we actually *attain* self-knowledge. Rather, it is merely the presence of the *aim* that constitutes action (and hence agency). But whereas activity counts

24    Railton, "On the Hypothetical and Non-hypothetical," 311.

25    Railton, "On the Hypothetical and Non-hypothetical," 310.

26    Velleman first presented his theory of autonomous agency in *Practical Reflection*, although there his constitutivist account of reasons for action exists only in embryonic form. Over the next decade Velleman refined his view in a series of papers, the most important of which are reprinted in *The Possibility of Practical Reason*. Here his constitutivism is more explicit, particularly in "The Guise of the Good," "The Possibility of Practical Reason," "Deciding How to Decide," and the introduction to *The Possibility of Practical Reason*. In recent years Velleman has continued to defend and refine his theory of agency and the attendant constitutivist account of reasons. See especially "From Self Psychology to Moral Philosophy," "Motivation by Ideal," "The Self as Narrator," and "The Centered Self," all in *Self to Self*; as well as "Précis of *The Possibility of Practical Reason*" and "Replies to Discussion of *The Possibility of Practical Reason*."

as action if it merely aims at self-knowledge, action is *correct* only to the extent that we *achieve* self-knowledge. In other words, action consists merely in *aiming* at self-knowledge, but correct action requires actually *attaining* such knowledge. This is a crucial departure from Korsgaard, according to whom constitutive conditions are identical to constitutive norms. Velleman opens a space between action's constitutive condition and its constitutive standard of correctness, and it is here that the possibility of incorrect action resides.

The problem with Korsgaard's conception of action's constitutive standard is that she thinks about action as though it were an artifact—hence her comparison to houses. Velleman treats action as a process or activity, and what distinguishes one activity from another is its *aim*. Take the example of soccer. As Velleman observes, "kicking a ball around on a field does not amount to playing the game of soccer unless one is trying to kick the ball into the net more times than one's opponent."[27] Outscoring one's opponent by kicking the ball into the net is the constitutive aim of soccer. It is what distinguishes soccer from other activities. If I am not at least half-heartedly pursuing this aim, I am simply not playing soccer.

Now, since on Velleman's view we derive constitutive standards from constitutive aims, there is no problem explaining how those standards are normative for us—how, that is, they bind us. The constitutive standard of soccer is normative for someone playing soccer because *he already has the relevant aim*: he is trying to score more goals than his opponent. Consequently, Velleman has no need to appeal to the necessity of action in order to account for the normative force or binding nature of action's constitutive standard. To the extent that I am an agent, I aim at integrative self-knowledge, and thus considerations in light of which I would know what I am doing already have a foothold among my motives. Action's constitutive norm binds us not because we are somehow condemned to act, but rather because as agents we are already aiming to comply with it—that is, to know what we are doing.[28]

The immediate lesson here is that not all constitutivist accounts are created equal. Despite Railton's claim to the contrary, identifying insurmountable obstacles for one version of constitutivism does not undermine the cogency of the constitutivist program in general.

27    Velleman, "Replies to Discussion," 283.

28    As some of Velleman's critics have observed, there remains the question of what reasons we have to be agents rather than nonagents. See, for example, Enoch, "Agency, Shmagency," 177–80, 185–90. Compare Shah, "How Truth Governs Belief," 458–9; and Hussain, "The Guise of a Reason," 273. I shall return to this objection below.

However, we are not yet in the clear, for the way in which Velleman avoids the problem of incorrect action raises an entirely different question. As we saw above, Velleman leaves room for incorrect action by opening a space between action's constitutive condition (what constitutes action) and its constitutive norm (what constitutes correctness for action). But if there is space between them, how do we derive conclusions about the latter from premises only about the former? Velleman's proposed solution relies on his claim that action's constitutive condition is actually a constitutive *aim*. But this only pushes the question to a new level. How can we derive normative conclusions about action's constitutive standard of correctness directly from teleological premises about action's constitutive aim? In order to derive genuinely normative conclusions about our reasons for action—in order to provide a foundation for ethics— Velleman must be able to bridge the gap between aims and norms. As we shall soon see, this is not as easy as it sounds.

Before we proceed any further, though, I want to introduce some of the details of Velleman's theory of agency. We shall be in a much better position to evaluate his brand of constitutivism once we have a clearer idea of just what he thinks the constitutive aim of action is and why he thinks that action has such an aim in the first place.[29]

Velleman's claim that action constitutively aims at self-knowledge is the lynchpin of his general account of action and agency. He develops this account in response to a failure of the standard, Davidsonian theory of action. According to Velleman, defenders of the standard theory—in their haste to distinguish between action and mere happenings or bodily movements—fail to notice the crucial distinction between fully autonomous action and mere motivated activity. The standard belief-desire model, Velleman argues, turns out to be a model only of the latter. It successfully delineates the category of purposive, motivated activity, but this category includes a number of less than fully autonomous actions: reflexive reactions, unconscious behavior, Freudian slips of the tongue, and so forth.

Certain features of autonomous action therefore fall beyond the explanatory reach of the standard theory. The standard theory cannot explain why we sometimes feel like spectators rather than the agents of our own behavior—why we only sometimes feel in control of our

---

29    What follows is a brief summary of Velleman's theory of agency. For a more complete but still succinct presentation, see Velleman's introduction to *The Possibility of Practical Reason* or "Précis of *The Possibility of Practical Reason*."

behavior.[30] Nor can it explain how autonomous action is accompanied (or even preceded) by direct, nonobservational knowledge of what we are doing (or going to do).[31] Of course we do come to know about mere motivated behavior as or after we act, but this knowledge is different from our knowledge of our genuinely autonomous actions. In a case of mere motivated activity—say, a Freudian slip of the tongue—my knowledge is perceptual. I discover what I am saying much as a bystander might: by hearing the (presumably wrong) words come out of my mouth, or perhaps by seeing the startled expressions on the faces of my audience. This knowledge is passive or receptive: I do not know what I am saying until I say it. Autonomous action is quite the opposite. When I act autonomously, I know what I am going to do before I do it, and my knowledge somehow guides or directs my action.

Velleman's account of the possibility of directive knowledge is rather involved, but the basic idea is both straightforward and familiar. Suppose that I strongly desire to drink some water. As a human being, I am self-conscious, and I soon become aware that I desire water. I know that I tend to act in order to satisfy such desires, and this leads me to believe that I am going to act on my desire for water: I believe that I shall go into the kitchen and pour myself some water. How does this belief guide my action? This is where action's constitutive aim enters the picture. I have, Velleman argues, a sub-agential aim or motive towards self-knowledge; I aim to know what I am doing—to integrate my own behavior into my general understanding of the world. Put another way, the constitutive aim of action is simply to do what *makes sense*—to act in such a way that I can explain my behavior.[32] This aim inhibits me from acting until I determine how I am inclined to act. What is more, once I realize what I am inclined to do, my

---

30    This is, Velleman acknowledges, a matter of degree. Actions range along a spectrum from mere motivated activity (to which we are merely spectators) to fully autonomous actions (of which we are in complete control), and there may be relatively few actions that are *fully* autonomous. See Velleman, introduction to *The Possibility of Practical Reason*, 10.

31    This sort of "knowledge without observation" is the central focus of G. E. M. Anscombe's pioneering work, *Intention*.

32    See Velleman, introduction to *The Possibility of Practical Reason*, 26, for more on his somewhat technical use of the phrase "makes sense." It is worth mentioning that when Velleman first formulated a view of action's constitutive aim, that aim was simply autonomy. (See "The Possibility of Practical Reason," 193; and "Deciding How to Decide," 241.) He quickly abandoned this position in favor of his current one, though, when it was pointed out to him that were autonomy the constitutive norm of action, all actions would count as correct. See Clark, "Velleman's Autonomism." Velleman credited Korsgaard for "publicly daring" him to express this mistaken view. Korsgaard's role is not surprising, since—as we have seen—her view suffers from the same flaw.

motive towards self-knowledge reinforces my other motives for doing it. To return to the above example, my desire to drink some water motivates me to go to the kitchen. As soon as I accept that I am about to go the kitchen, the aim to know what I am doing reinforces the motivational efficacy of my original desire. The aim of self-knowledge provides further inclination to do what I believe I am going to do (and what I am already motivated to do). My belief that I am going to act out of my desire for water—combined with my aim of knowing what I am doing—augments the force of my original motive and thereby guides or directs my action.[33]

This sketch of how the constitutive aim of action fits into his larger theory of agency hopefully provides a glimpse of where Velleman's constitutivist project stands. His philosophy of action builds up to the conclusion that autonomous action has a constitutive aim.[34] His constitutivist metaethic digs down to reveal the need for a constitutive standard of correctness. The challenge is getting these two projects to meet. Velleman thinks that they do meet. As I mentioned above, he claims that self-knowledge is also the constitutive *norm* of action. But how does he manage to pull this normative rabbit out of a merely teleological hat? How, in other words, can we derive norms from aims?

At times Velleman makes this trick seem easier than it really is by substituting talk of *success* for talk of *correctness*. His claim then becomes that we can derive criteria of success

---

33    In a more recent, unpublished work—"How We Get Along," hereafter cited by lecture number and then page number—Velleman contends that we have independent grounds for expecting action to have a constitutive aim: "The very concept of a reason for acting should lead us to expect that practical reasoning has a distinctive point or purpose, motivation toward which is constitutive of agency" (2.3). According to Velleman, action must have its own constitutive criterion of correctness if there are ever to be any reasons for action *as such*—that is, "fully constituted reasons for acting, not reasons that trail off in an endless regress" (2.4). "Now," Velleman continues, "once we have identified a substantive criterion of correctness for actions as such, we will find that a creature who is not in some way aiming to meet that criterion is not in the business of acting for reasons" (2.5). Here Velleman works backwards from a normative supposition to a teleological conclusion in the philosophy of action. From the premise that action has a constitutive norm he concludes that it must therefore have a constitutive aim as well. But this conclusion follows only if Velleman can bridge the aforementioned gap between constitutive aims and constitutive norms, and that remains to be seen. Moreover, Velleman explicitly acknowledges elsewhere that not all activities with constitutive standards of correctness also have constitutive aims. "It is a criterion of success in the activity of contributing to a book symposium that one address the views actually expressed in the book. But one need not aim at this goal in order to engage in the activity" ("Replies to Discussion," 283).

34    For some recent objections to this aspect of Velleman's project, see Setiya, "Explaining Action"; Dancy, "Discussion—On Knowing What One Is Doing"; and Mele, "Velleman on Action and Agency."

from constitutive aims.[35] And this seems quite reasonable. From the fact that my aim is to throw this crumpled up bit of paper into the trash bin it seems to follow straightaway that I will succeed if the paper ends up in the bin and fail if (as usually happens) it winds up on the floor. But this entailment is *too* easy, and that should make us suspicious. Criteria of success follow so straightforwardly from aims because "success" is—or at least can be—a teleological term. On this understanding of the word, however, success conditions follow *analytically* from aims. Success is just defined as the achievement of one's aims, in which case the fact that Velleman can derive constitutive criteria of success from constitutive aims is hardly worthy of much excitement: teleology entails teleology. What he is looking to derive is something *normative*. The whole point of his constitutivist program is to find nonnormative foundations for the norms of practical reason. What makes Velleman's derivation of success conditions from aims seem significant is his frequent transition from talk of success to talk of *correctness*.[36] Correctness *is* a normative notion, and a constitutive standard of correctness is precisely what we are after. But in the sense in which success conditions follow straightforwardly from aims, correctness conditions do not follow as well. Or at least they do not *obviously* follow.

Then again, maybe there is a sense in which aims *do* give rise to fully normative standards of correctness. I can aim at something by adopting it as an explicit goal. I can make it one of my ends—treat it as something worth pursuing or to be pursued. Perhaps this sort of aim generates genuinely normative correctness conditions.[37] After all, something along these lines lies behind what is often called the principle of instrumental reason: it is the normative status of my ends that gives me reasons to take the relevant means.

There are two important observations to make here. The first is that although we are free to adopt self-knowledge as an explicit goal, this sort of aim cannot be what is constitutive of action. One must already be an agent in order to make something one's end. But, as Velleman notes, an aim can be constitutive of agency only if it is something *nonagents* can have, since

35 See, for example, Velleman, introduction to *The Possibility of Practical Reason*, 26; "The Possibility of Practical Reason," 176–82; "Deciding How to Decide," 230–1, 241; "On the Aim of Belief," 246; and "Replies to Discussion," 283, 287–8.

36 See, for example, the way in which Velleman shifts back and forth between talk of success and talk of correctness at "The Possibility of Practical Reason," 177–9; and again at "Deciding How to Decide," 231–4. In "On the Aim of Belief," Velleman treats "successful" and "correct" as synonymous: "And a belief can be justified by indications of its truth because being true is what would make it successful or correct" (246).

37 Compare Shah and Velleman, "Doxastic Deliberation," 498.

the constitutive aim of action is what you add to a nonagent in order to get an agent.[38] The aim that constitutes agency *drives* practical reasoning, and thus it cannot be just another end that is represented *in* reasoning as something desirable or to be pursued. One must already be reasoning practically before one can represent something as desirable, in which case one is already an agent.

A different metaphor may help. If action has a constitutive aim, then this aim is what gives rise to the theater of practical reasoning. It creates the venue in which our practical deliberations are played out. Explicit ends are merely the actors that populate those deliberations. Now, to say that an explicit aim is constitutive of action is to say that one of the actors also somehow constitutes or gives rise to the theater in which he appears. And this simply does not make sense. The theater is a precondition of the performance, just as the constitutive aim is a precondition of explicit aims. Without action's constitutive aim, there would be no theater in which an agent's ends could perform. The suggestion that action's constitutive aim is actually an explicit aim therefore borders on incoherence.

The second observation is that *no* norm-generating explicit goal could play the foundational role in any constitutivist account of the sources of normativity. Taking something as my end gives me a reason to pursue the relevant means only if the end—or my adopting it as an end—has some independent normative standing.[39] Velleman's constitutivism may ultimately be able to explain the normative standing of explicit aims. He cannot, however, simply presuppose this standing in order to derive a constitutive norm from an explicit aim. This would involve establishing the authority of action's constitutive norm by appeal to a different, independently authoritative norm. But this vitiates the essence of the constitutivist project, since the constitutive norm is supposed to be the foundation of *all* practical normative authority. It is supposed to be the point at which justification comes to an end, and thus it cannot rely on some other norm for justification. What would justify this other norm? Not action's constitutive standard, since then the justification would be viciously circular. Instead we would have to appeal to yet another norm, in which case the regress of justification would be off and running again.

---

38    Velleman, "The Possibility of Practical Reason," 191. Compare "What Happens When Someone Acts," 53; and "Replies," 293, where Velleman claims that action's constitutive aim just *is* the agent, functionally speaking (143).

39    This is one of the lessons of Korsgaard's "The Normativity of Instrumental Reason."

The point of the foregoing is not that Velleman's constitutivist strategy of grounding ethics in action's constitutive aim is necessarily doomed, but rather that the move from constitutive aims to constitutive norms—that is, the move from the philosophy of action to the foundations of ethics—is not as easy as he sometimes makes it seem. But I am not ready to give up hope. In the next section I shall consider some less straightforward ways in which the transition from constitutive aim to constitutive norm might be effected.

### Reductions and Open Questions

In *The Possibility of Practical Reason*, Velleman presents his constitutivist project as a reductionist one. His reductionism is most explicit in the parts of the book devoted to topics well within the boundaries of the philosophy of action. Many of the essays are devoted to showing how we can account for difficult notions like *autonomy* and *agent causation* in purely naturalistic terms.[40] But of course Velleman's reductionist hopes extend beyond the boundaries of the philosophy of action. He wants to provide a naturalistic account not only of action for a reason but also of *reasons for action*. That is, he hopes to trace practical normativity to a nonnormative source by reducing practical normativity to something merely psychological and teleological.[41] Velleman's reductionist ambitions actually extend even further, since he wants to provide a similarly reductionist constitutivist account of reasons for belief and epistemic normativity as well.

If reduction is on the agenda, then it is clear how we could move from constitutive aims to constitutive norms. This would not actually be much of a move at all, since *constitutive aim of action* and *constitutive norm of action* would be two concepts that pick out the very same thing (much as the concepts *water* and *$H_2O$* pick out the same stuff). But of course this reductionist proposal does not solve the problem of the previous section. It merely requires us to restate it in different terms. Instead of asking how we can *derive* constitutive norms *from* constitutive aims, we must now ask what entitles us to *reduce* constitutive norms *to* constitutive aims. That

40    In *Practical Reflection*, Velleman acknowledges that his naturalistic accounts of agency and autonomy are hypothetical in the following sense. These concepts have descriptive content, and in explicating them we may discover that nothing in the world actually answers to the descriptions in question (5–6). Velleman has since become a bit more ambitious. In "From Self Psychology to Moral Philosophy," for example, he discusses a number of empirical psychological studies that show that genuine agency may be instantiated in the world after all. Compare Velleman, "Replies to Discussion," 282–3.

41    See Velleman, "Replies to Discussion," 287.

self-knowledge is the constitutive norm of action is a (putative) normative fact. That agents constitutively aim at self-knowledge is a (putative) teleological and psychological fact. What licenses the reduction of the former to the latter?

Like all reductionist projects in metaethics, this one faces a seemingly insurmountable objection in the form of G. E. Moore's famous (or perhaps infamous) open question argument.[42] Even if we accept that self-knowledge is the constitutive aim of action, it remains an open question whether the attainment of such knowledge is what makes actions correct. As Stephen Darwall puts it, "for anything characterized in nonnormative terms, we can sensibly ask whether there is any reason to choose or pursue it."[43] But this would not be an open, sensible question if facts about action's constitutive aim settled the matter of action's constitutive standard of correctness, and so there seems to be something amiss with Velleman's proposed reduction.

Actually, some version of the open question argument has been the objection confronting Velleman's constitutivist project all along. If we formulate the strategy as one of *deriving* norms *from* aims, then the problem is: Whence comes the normativity? How does self-knowledge— the constitutive aim of action—lend reasons their normative force? What does knowing what one is doing have to do with what one ought to do? What is authoritative about a sub-agential aim? If, however, we present the strategy as one of *reducing* norms *to* aims, the question becomes instead: Whither goes the normativity? How can a nonnormative constitutive aim capture the authority of a standard of correctness? It seems that by reducing the normative notion of correctness to the teleological notion of aim, Velleman has lost the very thing he set out to vindicate. He adopted the constitutivist approach in order to halt the regress of justification, but like the empirical naturalist internalists we discussed in chapter 3, he seems only to have changed the subject. He has not grounded the normative authority of practical reason; he has eliminated it.

Interestingly, these Moorean worries are less of an issue for Velleman's constitutivist approach to epistemic normativity. As Nishi Shah has effectively argued, if anything Velleman

---

42    See Hussain, "The Guise of a Reason," 269–70.

43    Darwall, "Moore, Normativity, and Intrinsic Value," 485. Darwall presents this as an argument against the teleological claim that action has a substantive constitutive aim. I am obviously appropriating it as an argument against reducing normative facts about action's standard of correctness to teleological facts about action's constitutive aim.

faces the opposite problem there: when it comes to belief, there do not seem to be any open questions.[44] The question whether to believe what is true is decidedly closed. Put another way, deliberations about what to believe always and immediately give way to questions about what is true.

> Within the perspective of first-personal doxastic deliberation, that is, deliberation about what to believe, one cannot separate the two questions…. One cannot settle on an answer to the question *whether to believe that p* without taking oneself to have answered the question *whether p is true*.[45]

Shah points out that Velleman's teleological account of belief's constitutive aim cannot explain this phenomenon, often called the *transparency* of doxastic deliberation. In *The Possibility of Practical Reason*, Velleman's account of the aim of belief is couched in terms of causal regulation: belief "is constitutively regulated by mechanisms designed to ensure that it is true."[46] However, Shah argues, this sort of teleological claim could explain transparency—that is, could explain the exclusive role of evidence in theoretical reasoning—only if the causal regulation of belief for truth were so strong that it *always* overpowered the regulative influence of other factors. In that case, an agent reasoning about something other than what is true would not count as reasoning about what to believe. Whatever attitude he was forming, it would not be a belief, since it was regulated for something other than truth. But strengthening the causal regulation for truth to this extent would rule out the possibility of beliefs that are not strongly regulated for truth. And we do sometimes have such beliefs. We believe blindly, as it were—in the face of overwhelming evidence. We engage in wishful thinking. Velleman, therefore, is caught on the horns of a dilemma. Either he cannot explain transparency, or he must rule out the possibility of wishful thinking and willful blindness. Neither option is especially appealing.

Shah's solution—which Velleman has since adopted—is to abandon the purely descriptive, teleological account of belief's constitutive aim in favor of the view that the aim is also built into the *concept* of belief.[47] According to Shah, that beliefs are correct if and only if

---

44    Shah, "How Truth Governs Belief," 460–5.

45    Shah, "How Truth Governs Belief," 447.

46    Velleman, introduction to *The Possibility of Practical Reason*, 17.

47    Shah, "How Truth Governs Belief," 465–75; and Shah and Velleman, "Doxastic Deliberation," 497–502. For a similar take on the concept of belief (albeit one more concerned with the problem of content and rule-following), see Boghossian, "The Normativity of Content." For a dissenting view, see Gibbard, "Truth and Correct Belief."

they are true is part of the very concept of belief. This is why the deliberative question *whether to believe that p* inevitably gives way to the question *whether p is true*. Belief, it turns out, is a normative concept, and the solution to the problem of the regress of epistemic justification is analytic. It is the concept of belief that halts the regress. Why should you believe what is true? If you fully grasped the concept of belief, you would not need to ask.[48]

I have followed this detour through the case of belief only to point out that Shah's proposed analytic solution to the regress of epistemic justification is not workable in the practical case. The sorts of questions that are closed for belief are wide open for action. The concept of action certainly does not encompass the idea that actions are correct only to the extent that their agents know what they are doing. This is evident from the fact that first-personal practical deliberation—that is, deliberation about what to do—does not always and immediately give way to deliberation about what would maximize integrative self-knowledge. Practical deliberation, in other words, is opaque, not transparent. That action has *some* standard of correctness may be one aspect of our concept of action, or more specifically, our concept of action *for a reason*.[49] But we cannot derive any *substantive* normative content from the mere concept of action.

In the face of these open question worries, Velleman has backed away from his earlier reductionist project.[50] In the next section I shall argue that this is a mistake. The constitutivist approach can convincingly show that the supposedly open questions are not really open at all, at least in the relevant sense. If Velleman is right that self-knowledge is the constitutive aim of action, then it is also action's constitutive standard of correctness, and considerations in light of which one would know what one is doing are also genuinely normative and authoritative reasons for action.

---

48     For some misgivings about analytic constitutivism, see Railton, "On the Hypothetical and Non-hypothetical," 309.

49     We may be able to describe this standard in *formal* terms, such as "acting rationally" or "acting in accordance with one's strongest reasons." But see Velleman, "The Possibility of Practical Reason," 176–8, for an explanation of why a formal standard of correctness cannot form the centerpiece of a constitutivist account of practical reason. For a different take on this question, see Wedgwood, "Practical Reasoning as Figuring Out What Is Best."

50     See Velleman, "Replies to Discussion," 294–7; and "The Centered Self," 282–3.

Shah's analytic strategy ensures that the question of whether to believe the truth is absolutely closed. Denying that belief is normatively governed by truth involves one in a contradiction. Question closed, and thus case closed. Absolute, analytic closure is not available when it comes to action, or so I have argued. Determining which action makes the most sense does not conceptually or analytically close the question of what one ought to do. *But this is not the only sort of closure there is.* If self-knowledge or doing what makes sense is action's constitutive aim, then figuring out what makes the most sense *does* close the question *from the practical point of view*, that is, *from the point of view of an agent who is trying to figure out what to do*. The act of practical deliberation is powered or driven by the aim to know what we are doing, and once we alight upon a course of action that would satisfy this aim, deliberation is effectively over. At this point the constitutive aim kicks in, and we act.

Let me elaborate. Think of practical deliberation as the attempt to solve a problem or puzzle. If we accept Velleman's account of action's constitutive aim, then the way we solve the problem is by identifying the action that would most contribute to our integrative self-knowledge. We are looking for the action that would give us the best explanatory grasp of our own behavior. Once we find it, the problem has been solved and deliberation is over. Thus, although the question of whether to do what would provide us with an explanatory grasp remains analytically open, it is practically closed: "There is no open question whether to be swayed by explanatorily relevant considerations—no open question *for an agent*, that is."[51]

Skeptics might still be wondering what all of this has to do with normativity—with our *reasons* for action and what we *ought* to do. According to Velleman, reasons are "considerations out of which we can fashion a description that would embody a knowledge of what we were doing, if we applied that description to ourselves in the way that would prompt us to behave accordingly."[52] In other words, reasons are considerations that provide an agent with an explanatory grasp of his own behavior. But, the skeptics persist, knowing what one is doing is a *theoretical* enterprise, not a *practical* one. How can an answer to this merely theoretical question also settle the normative question of what one has reason to do? After all, reasons

---

51    Velleman, "Replies to Discussion," 290 (emphasis added).

52    Velleman, introduction to *The Possibility of Practical Reason*, 27.

are supposed to involve more than explanation; they are also about *justification*. And yet the claim that practical deliberation is driven by the aim to do what makes sense is ultimately just a motivational claim. On Velleman's view, the constitutive aim of action is a type of motive, and so the purported force of reasons is merely the motivational force that results from triggering or engaging this motive. Velleman insists that it "thereby exerts a rational influence" as well, but now we have once again returned to our original objection: Whence comes the normativity of these reasons?[53] In Velleman's words, "what gives them rational authority as opposed to brute motivational force?"[54]

This is precisely the problem that derailed empirical naturalist internalism in chapter 3. These internalists tried to account for normativity by reducing it to motivational or psychological force. What they could not explain, though, is what makes motivation authoritative, and an account of normativity that eliminates the authority is not much of an account. Now, on Velleman's view, not all motivational force is authoritative; not all motives generate reasons. We can, in other words, distinguish between the force of a mere motive and the force of reason.[55] Reasons are considerations that engage not just any motive, but rather agency's constitutive motive or aim. This still leaves the basic problem intact, however. There may be a real distinction between run-of-the-mill motives and action's constitutive motive, but we do not yet have a justification for valorizing the latter as normative. How does the distinction between constitutive and nonconstitutive motives amount to the distinction between normative and merely motivational force?

The answer to this question lies in the practical conception of ethics that emerged back in chapter 2. According to this conception, the problem of practical normativity is first and foremost a practical problem; it arises when we are considering what to do. Were we not reflective creatures who can step back from and question our motives and inclinations, we

---

53    Velleman, "The Possibility of Practical Reason," 199.

54    Velleman, "Replies to Discussion," 293.

55    Of course, there are plenty of so-called internalists who claim to be able to distinguish between motivational force and rational force, but they do so by appealing to an independent standard of full rationality. Michael Smith might say, for example, that an addict's fully rational counterpart would advise the addict not to act on his urge, and therefore the urge's motivational influence is not an instance of rational influence. However, as we saw in chapter 1, philosophers who take this approach are actually externalists, since their view is grounded in an external standard of full rationality. This is not a problem for a Velleman, since the norm according to which an addict is irrational is internal on Velleman's view.

would have no use for normative concepts such as *reason* and *ought*, at least as applied to our behavior. And if the problem is ultimately practical, then so too is the solution. Recall Pamela Hieronymi's account of reasons. According to Hieronymi, we should think of reasons as *considerations that bear on a question*.[56] But then what is the question on which practical reasons or reasons for action bear? It cannot be the question of what I have most reason to do, since then we would be caught in an uninformative circle. No, the relevant question is simply: What to do? Or as Hamlet might have put it in one of his more philosophical moments: To $\phi$, or not to $\phi$—*that* is the question! To ask what I have *reason* to do is to look for considerations that bear on the question of *what to do*. Answers to the normative question are, therefore, also answers to the practical question. This is, after all, why we call reasons for action *practical* reasons. So, if we accept Velleman's account of action and agency, then the considerations that bear on the question of what to do are the considerations in light of which we know what we are doing. And that is why those considerations are valorized as reasons: they settle the practical question of what to do, and that is precisely the role reasons play.

Skeptics might still be worried that now we are simply back where we started at the beginning of this chapter. After all, we ended the previous chapter worried that the practical conception of ethics had led us to a dead end. Both brands of internalism we considered halted the regress of justification in ways that seemed to sap reasons of their authority. If, again following Hieronymi, reasons are considerations that bear on questions, the problem with internalism was that it yielded answers that were ultimately irrelevant to the normative question of what to do. What am I to do? The empirical naturalist internalists responded to this question by telling me what I am motivated to do—what I am likely to do or even what I am going to do. But that still leaves the practical question open. I know that I am inclined to eat too much dessert, but is that really the thing to do? I know that my addiction to Mahler is going to get the best of me, but is going to the concert instead of marking papers really the thing to do? My inclinations and desires do not *answer* the practical question so much as *bypass* it. The autonomist internalists, on the other hand, replied to the practical question with an empty formalism. What to do? *You decide!* But if I was ready to do that, I would not be asking what to do. *Do something!* But I want to know *what* to do, and it is hard to see how being told to do *something* is going to help. Neither empirical naturalist nor autonomist internalism can

---

56    Hieronymi, "The Wrong Kind of Reason." Compare Shah, "How Truth Governs Belief," 466–7.

really answer the practical question, and that is precisely what theories of practical normativity are supposed to do.

Constitutivism—or constitutivist internalism—is different: the authority of reasons does not stem from my supposed authority over myself (as the autonomist internalists suggested); nor does it reside in the coercive authority of my motives (as the empirical naturalist internalists proposed). Rather, the authority of reasons for action is grounded in the nature of agency—in the nature of practical deliberation and reasoning.[57]

Constitutivism identifies the considerations that answer the practical question—and that do so *definitively*. The constitutive aim of action sets the agenda for practical deliberation. On Velleman's view, that agenda is simple: to know what one is doing. Considerations in light of which an agent will know what he is doing are the signposts he must follow in order to complete that agenda. This is why those considerations are valorized as *reasons*. They do

57   Interestingly, Velleman's constitutivism is a nearly perfect synthesis of empirical naturalist and autonomist internalism. Velleman rejects the standard model of agency in part because it does not leave room for genuine practical reasoning—because it does not leave room for the will as distinct from our everyday desires and motives. As Darwall describes in *The British Moralists and the Internal 'Ought'*, a similar failing of empirical naturalist internalism inspired the first autonomist internalists to formulate a theory of pure practical reason—a theory according to which "reason is a creative faculty, with ends internal to it" (107). Velleman follows in the autonomists' footsteps, but he throws in a remarkable twist: he offers a naturalist reduction of agency and practical reasoning. According to this reductionist account, practical reasoning is reasoning about what makes sense, and the will is the aim to know what one is doing. Thus, unlike the empirical naturalists, Velleman believes that there *is* such a thing as practical reason. However, practical reason turns out to be a branch of theoretical reasoning. Thinking about what to do is just thinking about what makes sense; it is genuinely practical *but also* theoretical. As Velleman puts it, our autonomy is "a function of our intellect" (introduction to *The Possibility of Practical Reason*, 31).

To see how Velleman straddles the divide between empirical naturalist and autonomist internalism, consider Darwall's description of Cudworth's reaction to Hobbes:

> One account he found utterly inadequate was Hobbes's. He agreed with Hobbes's critique of faculty psychology that it is as absurd to suppose that there is a faculty of will that wills as it is to think, say, that the faculty of walking walks. It is *the person* who wills. In identifying will as 'the last appetite in deliberating,' however, Cudworth believed that Hobbes was insufficiently sensitive to his own insight. The will 'is no particular appetite,' Cudworth objects. It involves 'the soul comprehending its whole self.' (*British Moralists*, 132; see Darwall for the appropriate citations to Cudworth's works.)

For Velleman, however, the will is *both* an "appetite" *and* the "soul." Velleman's reasons for rejecting the standard theory mirror Cudworth's for rejecting Hobbes' theory: there is no room for the agent or the self to play a role in willing. Where Velleman departs from the autonomist line is in his reductionist account of agency: the agent, functionally speaking, *just is* the constitutive aim to know what one is doing. For Velleman, as for the autonomists, reason can be "purely practical by enabling a form of *normative judgment* that is intrinsically motivating" (Darwall, *British Moralists*, 19). But as for the empirical naturalists, this sort of reasoning is ultimately driven by a motive: action's constitutive aim.

more than merely engage an agent's motive; they also *justify* authoritatively. After all, to seek further justification for a reason is to ask *why* it bears on the question of what to do. And almost any answer simply invites the same question again: Why does *that* bear on the question of what to do? Why does the fact that I made a promise to my students bear on the question of whether to go to the concert or mark papers this evening? But if self-knowledge is the constitutive aim of action, then the regress of justification will halt there. Velleman does not merely identify considerations that are *relevant* to the practical question. These considerations *settle* the question *conclusively*. Whether to do what makes sense is not an open or live practical question. On the contrary, this is where practical reasoning comes to an end, and therein lies the authority of practical reasons.[58] Thus, although the constitutive aim of action *is* a motive, it is not *just* a motive. It is the engine of practical reasoning, and it determines what a definitive answer to the practical question must look like. Since what answers the practical question also answers the practical normative question, the constitutive aim of action is also action's constitutive *norm*.

Recall that in the introduction to this chapter, I suggested that we search for a view in which the two crucial aspects of normativity—authority and the capacity to govern—arise together. This is precisely what we have found. On the constitutivist theory I have been defending, reasons govern in virtue of their authority: the self-knowledge norm governs our actions *because* it sets the agenda for practical deliberation. Moreover, the self-knowledge norm is authoritative *because* of the way it governs us—because, in other words, it is action's constitutive *aim* as well as its constitutive norm. And so action's constitutive aim explains *both* the authority *and* the influence of practical reasons. This is why constitutivism can halt the regress of justification: it captures both facets of normativity at once.

---

58    Hussain also charges that Velleman cannot account for the *experience* or *awareness* of the normative force of reasons—an awareness that is supposed to accompany all action for a reason, according to Velleman. (See Hussain, "The Guise of a Reason," 268–9.) But there is an ambiguity in Hussain's objection. If the charge is given a *de dicto* reading, then it is true but hardly a threat to constitutivism. It is certainly the case that not all instances of rational influence will involve the agent's awareness of that influence *in those terms*. (If action for a reason were always accompanied by a *de dicto* awareness of normative force, then such action would be impossible for people who do not possess the relevant normative concepts. See Velleman, introduction to *The Possibility of Practical Reason*, 15, for a similar point.) If Hussain's charge is given a *de re* reading, then it is simply not true. The normative force of reasons just is the practical force of considerations in light of which we can make sense of our actions. And thus to the extent that we experience the latter, we also have a *de re* awareness of the former.

If the foregoing is correct, then Velleman was wrong to abandon his reductionist project. What led him away from reductionism was his fear that it could not bridge the gap between his philosophy of action and his constitutivist metaethic—between action's constitutive aim and its constitutive norm. But I have shown that the practical conception of ethics *can* effect this reduction. The practical concept *the thing to do* and the natural concept *the thing that makes sense* pick out the very same property, just as the concepts *water* and $H_2O$ pick out the same stuff. Of course there is no conceptual or analytic reduction here. This is why the questions *what to do* and *what ought I to do* do not inevitably and explicitly give way to the question *what makes sense*—why there is no transparency in the case of action. Rather, these questions are simply different ways of arriving at a single answer. We can ask in the practical mode—*What to do?*—or the theoretical mode—*What makes sense?* But in each case we are seeking the same thing: we are aiming at self-knowledge.[59]

## Agency, Shmagency

Near the beginning of this chapter I described two distinct phases of the constitutivist project. We now have at least a sketch of the first. Velleman has provided a sound method for finding action's constitutive norm: find action's constitutive aim. If the arguments of the previous sections are correct, then this approach yields a bona fide authoritative norm—and

---

59    I should mention yet another possible objection here. The version of constitutivism I am defending obviously requires that action actually have a constitutive aim. And a number of philosophers have doubts on this front. In "Because I Want It," Stephen Darwall—who seems to accept the constitutivist approach for belief—claims that this is one way in which the practical realm is different from the theoretical realm. "The aims of practical reason seem (practically) normative all the way down" (135). A similar conviction underlies Gibbard's expressivism in *Wise Choices, Apt Feelings* (chs. 1 and 9). See also Gibbard, "Knowing What to Do, Seeing What to Do," 224–8, where he suggests that practical reasoning is "planning" all the way down. Linda Radzik is another who assumes that practical reasoning is different from theoretical reasoning in this respect. See Radzik, "Incorrigible Norms," 645.

     It is not always clear what underlies these general doubts about practical deliberation having a constitutive aim, especially in light of all of Velleman's work on the subject. One of the upshots of this work is that action is not so different from belief after all, at least in this crucial respect. Recently Darwall has provided a somewhat more specific reason for doubting that action has a constitutive aim: "No evaluative or practical attitude has a 'substantive aim,' like truth, which can be understood independently of norms for the relevant attitude and which closes the normative question of what to value or choose" ("Moore, Normativity, and Intrinsic Value," 485). This is just the open question argument once again. (I suspect something similar lies behind Gibbard's skepticism.) If, as I argued above, the constitutive aim of action *can* close the question of what to do, at least from the practical point of view, then the open question argument provides no reason at all even to doubt that action has a constitutive aim.

one that can be transgressed, thereby avoiding the problem that plagues Korsgaard's Platonic or Aristotelian approach. But now we must turn to the second phase: demonstrating that the foundations provided by this constitutive norm are strong enough to support a genuinely normative ethics. Even if the skeptics of the previous section are willing to accept that constitutivism has earned the notions *reason* and *norm*, they will undoubtedly still be worried.

On the constitutivist view, the constitutive norm of action works like the rules of a game such as soccer or chess. The rules of action are binding on agents just as the rules of chess are binding on chessplayers. But, as David Enoch has recently noted, we need not be chessplayers. The ultimate normative authority of the rules of chess is contingent on whether we *should* be chessplayers.[60] If you do not have an independent reason to play chess, Enoch argues, then you have no reason to try to checkmate your opponent. (Similarly, if you have no independent reason to play soccer, you have no reason to try to kick the ball into your opponent's goal.) But if the normativity of action's constitutive norm is modeled on the rules of a game, then the same must go for action: if you have no independent reason to play the action game—that is, to be an agent—then you have no real reason to aim for self-knowledge. Why, Enoch asks, should he be an agent, when he can just be a shmagent instead?

> I am perfectly happy being a shmagent—a nonagent who is very similar to agents but who lacks the aim (constitutive of agency but not of shmagency).… I am perfectly happy performing shmactions—nonaction events that are very similar to actions but that lack the aim (constitutive of actions but not of shmactions).…[61]

Enoch's point is that the rules of a game are normatively relevant only if the game itself has some normative weight. If you are not pursuing the constitutive aim of action, the worst that can be said of you is that you are not an agent. But so what? "You are," Enoch insists, "perfectly justified in treating this information as normatively irrelevant."[62]

As a response to the constitutivist strategy I have been scouting here, Enoch's objection is confused.[63] In particular, Enoch misunderstands the way in which reasons are normative

---

60    Enoch, "Agency, Shmagency," 185.

61    Enoch, "Agency, Shmagency," 179. Here Enoch sounds a bit like Railton's recalcitrant student Gary in "On the Hypothetical and Non-hypothetical."

62    Enoch, "Agency, Shmagency," 186.

63    It may be a much stronger objection to Korsgaard's version of constitutivism, however. Since I have already left Korsgaard's view behind, I shall not explore this further here. But see Enoch, "Agency, Shmagency," 178–80, 187–92.

on Velleman's view. Return for a moment to the chess example. The reasons generated by the rules and constitutive aim of chess are chess-related reasons—reasons for chess players. I shall simply call them *chess reasons*. Now, according to constitutivism, the constitutive norms of chess halt the regress of chess reasons. It does not make sense to ask what chess reason you have to checkmate your opponent. The very nature of chess makes it so. Enoch's point, then, is this: from the fact that you have a chess reason to checkmate your opponent, it does not follow that you have concomitant practical reason or reason *for action*. This is perfectly true. But notice that it is nothing but a new version of the problematic regress of justification we first encountered in the introduction. The norms of morality tell us what moral reasons we have, but without some further justification we have no authoritative practical reason to heed these moral rules.

We can now restate Enoch's objection to constitutivism in terms that should reveal the confusion. The constitutive norm of action may stop the regress of practical reasons or reasons for action, but without some further justification we have no authoritative reason to heed reasons for action. Just as we need a practical reason to play chess in order to establish the ultimate authority of chess reasons, we need some other sort of reason to establish the ultimate authority of reasons for action. We need, in other words, "an *independent* reason to be agents."[64] But now the confusion should be clear, for a question immediately presents itself. What other sort of reason is there? The puzzle is how to understand Enoch's claim that we have no "independent" reason to be agents rather than shmagents. Independent of what? What modifier should we put before the word "reason" here? Recall that reasons only justify in relation to a norm or standard of correctness. But standards of correctness are always standards *for* some attitude or activity: belief, chess, action, soccer, and so forth.[65] In relation to what standard of correctness do Enoch's "independent" reasons justify? The norms of being? Grammatical considerations aside, being is not an activity; it is not something one can do correctly or incorrectly.

At times Enoch suggests that he is still talking about *practical* reasons—plain, old reasons for action. If so, then in one sense his charge is true but also pointless. If you are not already an agent, then of course you do not have any practical reason to be or become one. But this is

---

64    Enoch, "Agency, Shmagency," 192 (emphasis added).

65    For more on the different relata that participate in the normative relation of *being a reason*, see Skorupski, "The Ontology of Reasons," 113–5; and Skorupski, "Propositions about Reasons," 27–31.

because you are not in the market for practical reasons at all. To demand a practical reason for someone who is not already an agent to become one is therefore to make some sort of category mistake. It is akin to asking for an epistemic reason to enter the belief game, or a chess reason to play chess. What makes Enoch's objection initially plausible is his claim that becoming an agent is something a nonagent can *do*, and hence something a nonagent might demand a practical reason to do. But of course becoming an agent is not really something a nonagent can *do*—or at least it is not something he can do for a reason.[66] When I awake each morning, I return to agency, but this is not something I *do* for a reason. It would be ludicrous for you to demand a reason for me to wake up.[67] How could I wake up for a reason? When I am sleeping I am not yet an agent, and so strictly speaking *I* cannot *do* anything. I certainly cannot do anything for a reason.[68] Thus, if Enoch's complaint is that constitutivism cannot provide nonagents with a reason to become agents, it is surely correct, but it is also beside the point. Only agents can do things for reasons, and if you are already an agent, there is little point in becoming one.

Enoch might now object that since practical reasons are only applicable to agents, I have not stopped the regress of justification. You can still ask why you should be an agent rather than a shmagent. But this is a live question only if *you already are an agent*. The question *whether to be an agent* is just a specific version of the practical question *what to do*, and so the considerations that bear on the former are just the considerations that bear on the latter— namely, practical reasons. Doubtless you can entertain the question of whether to *remain* an agent: you can contemplate the possibility of either suspending or even terminating your agency, perhaps by taking drugs, running head first into a brick wall, or committing suicide. Yet to entertain these as live options is to adopt the practical standpoint—the standpoint of agency. And from that standpoint, you *are* authoritatively governed by agency's constitutive norm. Now, it should go without saying that there are cases in which the norms of action

66  Korsgaard disagrees. In her Locke Lectures she argues that agents constitute themselves whenever they act. See Korsgaard, "Self-Constitution," 1.15–7, 1.25–8.

67  Of course, it would not be ludicrous for you to demand a reason for *you to wake me up*. But that is a different matter entirely.

68  Here our language may be confusing. I say, "I awoke at sunrise." Yet waking is not something I do autonomously, and so I do not do it for any reason. Rather, waking is something that happens to me. Compare Velleman, introduction to *The Possibility of Practical Reason*, 7–10.

recommends a temporary suspension of agency: every evening I remind myself that I should stop working and go to sleep. There are even cases in which terminating one's agency is the thing to do—for instance, jumping on a grenade to save one's comrades, or taking a bullet to shield one's child. But this too is beside the point. On Velleman's constitutivist view, the authority of reasons for action was never based on agency always being the best thing (or even a good thing). The normative authority of practical reasons is not predicated on you having good reasons to remain an agent. It is predicated on the fact that *you are already an agent.*

Enoch supposes that by asking why he should be an agents he can extend or restart the regress of justification and thereby undercut the putative authority of agency's constitutive norm. But the only standpoint from which he can sensibly ask what reason he has to be an agent is one that is governed by the norm he is trying to undermine. The very question presupposes the authority of practical reasons; this is precisely how constitutivism halts the regress of justification. Return again to the example of chess. What makes the rules of chess authoritative for you is not that you *should* be a chessplayer. The normative authority of chess reasons does not hinge on whether chess is good. It hinges on *whether you are playing chess.* If you are, then you are subject to the rules of chess. If you are playing some other game, or not playing anything at all, then the rules of chess are not normative for you. The same goes for action. If you are not an agent, then the constitutive norm of action has no authority over you. But this does not restart the regress of justification, since nonagents or shmagents are in no position to demand practical justification. As soon as you ask why you should be an agent, you have readopted the practical perspective, and from that perspective justification comes to an end at the constitutive norm of agency. Put another way, Enoch does not undermine the authority of action's constitutive norm by asking whether he should continue being an agent any more than I subvert the rules of soccer by deciding to stop playing.[69]

69  Compare Railton's response to Gary in "On the Hypothetical and Non-hypothetical": "You already defer, in posing this question, to the very thing you seek to challenge. You must already see—and feel—the 'practical logic' of what you claim to find arbitrary or problematic" (317). Compare also Rosati, "Agency and the Open Question Argument," 522. In general, the problem with Enoch's paper is that he simply fails to notice the differences between Korsgaard's brand of constitutivism (which treats action like an artifact) and Velleman's (which treats action like an activity). Were the constitutive norm of action like the constitutive norm of houses, then the authority of practical reasons might very well depend on an "independent" reason to be agents. But as I have already pointed out, the version of constitutivism based on constitutive aims does not have this problem.

Enoch's objection to constitutivism rests on the assumption that we can distance ourselves from action's constitutive aim but still be in a position to demand and respond to practical reasons. He reckons that we can step back from the constitutive aim of action and treat it as "normatively arbitrary."[70] But Enoch seems simply to misunderstand how this aim functions. No matter how far we step back from our desires and aims, if we are still looking for reasons for action then the constitutive aim of agency is driving our deliberations. As Velleman puts it,

> The agent cannot attain a perspective of fully detached reflection on the aim constitutive of action, because it is an aim that must be operative in order for him to reflect, in the first place…. There is no perspective in which the drive to understand is not behind the lens of introspection, even if also in front of it, because no mental processes that were not regulated by this drive would be sufficiently coherent to constitute a perspective at all.[71]

Even as we try to distance ourselves from the pursuit of self-knowledge, we continue to occupy the practical standpoint. This means that considerations of self-knowledge will continue to bear on whatever practical questions we may be asking ourselves; and as we saw above, this is just what makes such considerations normative! If we step back even further, however, and truly leave the constitutive aim of action behind, then we leave practical justification behind as well. If we are no longer asking practical questions, there will be no considerations that bear on such questions—and hence no reasons. We will have abandoned agency entirely, and reasons for action with it.[72]

There is surely *some* room for reflective distance or separation here, Enoch might insist. After all, when we are not actively trying to figure out what to do we can wonder whether self-knowledge is an aim worth pursuing. However, as I argued in chapter 2, this sort of idle normative speculation is not really normative at all. What makes something practically normative is that it bears on the question *what to do*. And so if we are not actively trying to decide what to do, we are not in the market for practical reasons or practical justification.

---

70    Enoch, "Agency, Shmagency," 178.

71    Velleman, "Replies," 292–3. See also Velleman, "The Centered Self," 282.

72    Another possibility is that we are simply unaware that self-knowledge is the constitutive aim of action. In our ignorance, we might then wonder from within the practical perspective why we should pursue this aim. But this possibility does not pose any threat to my constitutivist view, which proposes only a *synthetic* reduction of reasons for action. The crucial point is that whether we realize or not, considerations in light of which we know what we are doing will play their usual regulative and normative role so long as we occupy the practical standpoint.

Could Enoch be demanding something other than a *practical* reason to be an agent rather than a shmagent—some kind of "independent" reason? As I mentioned above, if he is looking for another kind of reason, then the obvious and appropriate question is: What kind? There are no other relevant kinds of "independent" reasons. Thus, either Enoch's question of what reason he has to be an agent is a practical question, in which case he is authoritatively governed by the norms of agency, or it is incoherent.

But perhaps I am forgetting something. After all, shmagency is supposed to be just like agency, according to Enoch. The only difference is that it does not have the same constitutive aim and hence is not governed by the same constitutive norm. So perhaps there are reasons for shmaction—shmactical rather than practical reasons. And perhaps we can ask of a shmagent whether he has any shmactical reason to become an agent instead. It should be obvious, though, that this is not really a coherent alternative. The constitutive aim of action is what makes action what it is. Something "just like" agency but without its constitutive aim would not actually be like agency at all. A shmagent would more closely resemble Harry Frankfurt's wanton: he would have "no identity apart from his desires" and would be "no different from an animal."[73] A shmagent would certainly not be the sort of creature who could do anything for reasons.

All of this suggests that the sort of normativity grounded by constitutivism is as robust and authoritative as we could want. We do not need a reason to opt for agency over shmagency in order to stop the regress of practical justification, and that is all that we started out to do. However, it is still too early to declare success even in that task. In the next section, I shall consider another objection to the constitutivist account of the foundations of ethics.

**Normativity Enough?**

The case for my reductionist constitutivism clearly rests on what I have been calling the practical conception of ethics—that is, the idea that reasons are normative in virtue of the fact that they bear on the practical question *what to do*. This appears to reduce normative force to something private and first-personal, and skeptics may reasonably balk at this idea. Normative force does not seem to be limited to the first-person perspective. Statements I address to myself about what reasons I have are not the only kind of normative statement. When I tell you that

73    Frankfurt, "Freedom of the Will and the Concept of a Person," 18.

you have a reason to go to law school, I am making a normative claim. This claim has normative force, at least if it is correct. The same goes for my third-person judgment that Caesar ought not to have crossed the Rubicon. Consider the three sorts of statement side by side:

> That I made a promise is a reason *for me* to skip the concert and mark papers instead.

> That you love the law is a reason *for you* to go to law school.

> That civil war will ensue is a reason *for Caesar* not to cross the Rubicon.

They certainly *seem* to be parallel. That is, they seem to be normative in precisely the same way. Can constitutivist internalism explain how?

Yes, and no. The three statements *are* parallel, although perhaps not in the way the skeptics suppose. One thing they all have in common is that each offers a consideration that bears on a particular practical problem or question. I am facing the question of what to do this evening, you are facing the question of what to do with the rest of your life, and Caesar was facing the question of whether to cast the die. Each is (or was) a pressing practical problem. But of course the statements are also different, in that the person offering the consideration is not always the person facing the practical problem. In this respect the second and third statements are different from the first. There is also a respect in which the second and third statements are different from each other: the second is addressed *to the person with the relevant practical problem*. This is a feature it shares with the first statement, not the third. So what should we make of all this?

According to internalism, the *statements* are not normative in the way that *reasons* are normative: the statements themselves are not considerations that bear on the question *what to do*. Rather, the statements are *about* such considerations, and so they are normative in their subject matter. In this respect all three are straightforwardly normative. But does their normativity extend beyond their subject matter? Do the statements themselves have any normative force? Many philosophers seem to think that they do, and that this is a problem for the view that normative force is entirely first-personal. The skeptics' worry is that if normativity is entirely first-personal, then only the first of the three statements has any normative force. But the most that actually follows from the practical conception of ethics is that only the first statement has normative force *for the speaker or judge*. This hardly entails that the other statements exert no force at all. The second is supposed to be normative *for you*. In saying

that your love of the law is a reason for you to go to law school, I am addressing *your* practical problem. I am offering a consideration that bears on *your* question *what to do*. In other words, the second statement has some force *for you*, because the consideration it mentions is normative *for you*. Does that mean it has no force for me, the speaker? It certainly entails that it has less force for me, since I am not trying to figure out whether to go to law school. But it may still have some normative force for me, given that general considerations about what I love are relevant to my decisions about what to do.

That the normative force of my second-person statement is primarily normative for you is actually an *advantage* of the internalist approach. To see this, consider another practically-minded theory of normativity: Allan Gibbard's expressivism. "To call a thing rational is to endorse it in some way," Gibbard observes at the outset of *Wise Choices, Apt Feelings*.[74] The challenge is to understand this endorsement—to "search for a sense of 'endorse' for which the dictum holds true."[75] The view Gibbard develops in response to this challenge is motivated by a sort of internalism different from the one I have been discussing. Gibbard accepts *judgment internalism*.[76] This is roughly the idea that normative force is a feature of normative *judgments* rather than normative *reasons*. The practicality of normativity is explained by appealing to the force of judgments employing normative terms.[77] Thus, Gibbard explains what it is *to be rational* by explaining what it is to *call something rational*: "to call something rational is to express one's acceptance of norms that permit it."[78]

On Gibbard's view, all three of my above statements have the same sort of normative force: they all express *my* acceptance of certain norms. But this gives rise to a puzzle. As Gibbard is well aware, expressivism has a difficult time explaining why I would ever bother to tell you that you have a reason to go to law school, and why you would ever bother to listen.[79] Suppose you come to my office worried about your future. You cannot decide whether to

---

74    Gibbard, *Wise Choices, Apt Feelings*, 6.

75    Gibbard, *Wise Choices, Apt Feelings*, 6.

76    See chapter 1, note 8 above.

77    Gibbard's judgment internalism is evident in the subtitle of his book: "A Theory of Normative Judgment." Internalism of the sort I have been discussing—existence internalism—is a theory of normative *reasons*, or a theory of normativity itself.

78    Gibbard, *Wise Choices, Apt Feelings*, 7.

79    See Gibbard, *Wise Choices, Apt Feelings*, chs. 9–10.

attend law school or become a horse jockey, and you ask for my advice. "You love the law," I reply, "and that is surely a reason for you to go to law school!" Now, if all I am doing here is expressing my own acceptance of norms that recommend doing what you love, then it is hard to see how my reply is anything but a *non sequitur*. You are seeking help with *your* practical problem, and I have merely expressed some norms that *I* accept and that would dictate a solution were it *my* problem. But if those are not your norms as well, how is that supposed to help? Of course I might be *assuming* they *are* your norms and that you have simply failed to see which course of action they prescribe. Yet that sort of concordance of norms frequently does not obtain. Alternatively, I might be *demanding* that you accept my norms, but on what basis do I make such a demand? Perhaps you accord me a certain sort of authority. After all, you came to me for advice. But even that is "puzzling" for expressivists, as Gibbard's acknowledges: "Should [you] ever take [my] acceptance of norms as reasons for accepting them, when [you] do not independently accept norms from which they follow?"[80] The answer is not obvious.

My point here is not that Gibbard's expressivism cannot solve this puzzle, but rather that on the constitutivist internalist view I am now defending *there is no puzzle*. You come to me for advice in the hope that I will be able to identify some considerations that bear on your practical problem but that you have so far missed. And I remind you that you love the law not to express *my* own norms but rather with an eye to commending your love for the law as something relevant to *your* practical deliberations. You come in search of considerations that are normative for you, and that is precisely what I provide. The fact that my statement about *your* reasons does not have the same normative force *for me* as my statement about *my* reasons is, therefore, an advantage of constitutivist internalism. My statement about your reasons lacks full normative force for me because it is *addressed to you* and has normative force *for you*.

This does not mean that my statement about your reasons lacks *all* normative force for me, since my practical deliberations are also governed by the same norm that governs your deliberations: the self-knowledge norm. There is a sense, then, in which I *am* deploying my own norm when I give advice to you, and even when I state that the prospect of civil war was a reason for Caesar not to cross the Rubicon. Of course Caesar is no longer around for me to address, and thus it seems that whatever normative force my claim about him has now is limited to the force it has for me and my audience. But, again, that is no weakness of the

80    Gibbard, *Wise Choices, Apt Feelings*, 176.

constitutivist view. When we talk idly about what Caesar should have done, we *are* somewhat removed from the normative force of our claims. We are not *entirely* removed, however, since what would have made sense for Caesar to do is also what it would make sense for us to do *were we in his sandals*. To the extent that we try to adopt Caesar's point of view, we confront the hypothetical practical question of whether to cross the Rubicon, and consequently we feel the force of action's constitutive norm.

But now the skeptics are ready with another objection. On this constitutivist internalist picture of normative force we are developing, reasons are only ever normative *for* someone—for me, for you, for Caesar. There is no normative authority *simpliciter*, no authority as such. Given the differences among agents, then, there are few if any considerations that count as objective reasons, or reasons for everyone. Moreover, even if there are such considerations, their objectivity or universality is merely contingent. By adopting the practical conception of normativity, we seem to have unwittingly reintroduced the subjectivism we struggled so hard to shed in chapter 1.

In one sense this objection is simply false. On the constitutivist view—or at least on Velleman's version thereof—there *are* considerations that are reasons for all agents, namely, considerations in light of which an agent knows what he is doing. In other words, the following schema is true for any agent $A$ and any action $\phi$: that by $\phi$-ing $A$ would know what he is doing is a reason for $A$ to $\phi$. The norm that prescribes self-knowledge is a substantive norm that governs all agents, and so strictly speaking the charge that all reasons are subjective is false. But this is too fast. The point of the objection is that what makes sense for one agent is not always what makes sense for another agent in the same circumstances; and so once we specify considerations in a bit more detail, none count as reasons for all agents. The norms of morality, for instance, are not normative for all agents, since there are undoubtedly some pretty horrific agents to whom, say, the suffering of others matters not at all. For these appalling agents, considerations of others' suffering will never bear on any of their practical problems, and thus they are not reasons for such agents.[81] We set out to find the foundations of ethics, the skeptics remind us, and surely those are supposed to include the foundations of objective morality.

81    Compare G. A. Cohen's discussion of the idealized Mafioso in "Reason, Humanity, and the Moral Law," 183–8.

The first thing to note in response is that there are at least two separate axes along which normative reasons can be objective. On the one hand, there is the axis that runs from subjective to universal. Are there substantive norms (beyond the constitutive norm of action) that authoritatively govern *all* agents, or is the authority of such norms always going to be dependent on or relative to *individual* agents' respective practical identities? The answer to this question depends on what sort of substantive norms we can derive from action's constitutive norm, and this in turn depends on what the constitutive norm of action is. I have adopted Velleman's account of this norm primarily for the purpose of demonstrating that *some such account* can stop the regress of justification and identify the source of normative authority. In that I think I have succeeded. Velleman is correct: the foundations of ethics are to be discovered in the philosophy of the action.[82] But the skeptics' point is well taken: Velleman's philosophy of action may not be able to support the normativity of objective morality. Perhaps future work on the nature of agency will produce a more substantive account—one that *can* ground a more robust normative ethic. But that is a project for future philosophers of action. The crucial point here is that there is a second axis along which reasons can be objective. This is the axis that runs from conventional authority to genuine or objective normative authority. And here the reasons we have identified are unquestionably on the objective side. Unlike the rules of soccer or chess, the authority of the constitutive norm of action does not stem from any arbitrary conventions. Nor does its authority depend on some other norm or norms for support. Rather, the constitutive norm of action derives its authority from the nature of action.[83] Thus, whether the reasons for action identified by constitutivism are subjective or universal—wherever, in other words, they fall on the first axis—their authority is as robust and objective as authority can get.

The skeptics will be disappointed nonetheless. They are ready to reject any account of the foundations of ethics that does not vindicate the universal authority of morality. They want to show that the laws of morality necessarily bind all rational agents. But their goal is not mine. I did not set out to discover the foundations of objective morality. My goal from the

---

82    I also think he is correct about the constitutive aim of action, and thus that the foundations he provides *really are* the foundations of ethics. But I have not defended his view of agency here.

83    This second axis of objectivity is similar to the one Thomas Nagel has in mind when he observes that even a purely instrumental theory of reasons for action can be objective. See Nagel, *The View from Nowhere*, 149–50. See also Darwall's distinction between the *thesis of universality* and the *thesis of objectivity* in *Impartial Reason*, 117–9. Compare Hampton, *The Authority of Reason*, 93–9.

beginning was to find the source of normative authority and thereby to put a stop to the regress of practical justification—or at least to identify a general method for finding such foundations. Objective morality would have been a nice bonus, but it was never a desideratum of success.[84] Any honest philosophical attempt to ground our pre-theoretical beliefs must leave open the possibility that some of those beliefs will turn out to be ungrounded. Besides, even though I have not shown that the laws of morality bind all agents, I have provided at least the outline of an explanation of how and why they bind *you and me*. We care about keeping our promises, preventing suffering, and helping others in need. We understand ourselves as responsible members of a moral community; we are committed to being moral agents. And this is precisely why the norms of morality bind us. Keeping our promises is what makes sense to us. When we help others in need we know what we are doing.

Following Velleman's constitutivist strategy, I have sketched a theory that can explain *why* moral considerations are genuine reasons for us, if not for everyone. If, like me, you are the sort of person who occasionally wonders and even worries about such things, then that, at least, is something.

84    Compare Dreier, "Humean Doubts about the Practical Justification of Morality," 98–9.

# BIBLIOGRAPHY

Altman, Andrew. "Breathing Life into a Dead Argument: G. E. Moore and the Open Question Argument." *Philosophical Studies* 117, no. 3 (February 2004): 395–408.

Anderson, Elizabeth. *Value in Ethics and Economics*. Cambridge, MA: Harvard University Press, 1992.

Anscombe, G. E. M. *Intention*. 2nd ed. Cambridge, MA: Harvard University Press, 2000.

———. "Modern Moral Philosophy." *Philosophy* 33, no. 124 (January 1958): 1–19.

Arkonovich, Stephen. "Defending Desire: Scanlon's Anti-Humeanism." *Philosophy and Phenomenological Research* 63, no. 3 (November 2001): 499–519.

Baier, Kurt. *The Moral Point of View: A Rational Basis of Ethics*. Ithaca: Cornell University Press, 1958.

Balguy, John. *The Foundations of Moral Goodness*. 1728–9. In Raphael, *British Moralists*.

Bennett, Jonathan. *Locke, Berkeley, Hume: Central Themes*. Oxford: Oxford University Press, 1971.

Bishop, John. *Natural Agency: An Essay on the Causal Theory of Action*. Cambridge: Cambridge University Press, 1989.

Blackburn, Simon. "Blackburn Reviews Dworkin." *Brown Electronic Article Review Service* (November 11, 1996), http://www.brown.edu/Departments/Philosophy/bears/9611blac.html.

Boghossian, Paul A. "The Normativity of Content." *Philosophical Issues* 13 (2003): 31–45.

Bolton, Martha Brandt. "Locke on the Semantic and Epistemic Role of Simple Ideas of Sensation." *Pacific Philosophical Quarterly* 85, no. 3 (September 2004): 301–21.

Bond, E. J. *Reason and Value*. Cambridge: Cambridge University Press, 1983.

Boyd, Richard N. "How to Be a Moral Realist." In Sayre-McCord, *Essays on Moral Realism*.

Brady, Michael S. "How to Understand Internalism." *Philosophical Quarterly* 50, no. 198 (January 2000): 91–97.

Brandt, Richard B. "The Science of Man and Wide Reflective Equilibrium." *Ethics* 100, no. 2 (January 1990): 259–78.

———. *A Theory of the Good and the Right*. Oxford: Clarendon Press, 1979.

Bratman, Michael E. Review of *The Sources of Normativity*, by Christine M. Korsgaard. *Philosophy and Phenomenological Research* 58, no. 3 (September 1998): 699–709.

Brewer, Talbot. "The Real Problem with Internalism about Reasons." *Canadian Journal of Philosophy* 32, no. 4 (December 2002): 443–73.

Brink, David O. *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press, 1989.

Broad, C. D. *Five Types of Ethical Theory*. London: Routledge and Kegan Paul, 1930.

Broome, John. "Normative Requirements." In *Normativity*, edited by Jonathan Dancy. Cambridge: Blackwell Publishers, 2000.

———. "Reasons." In *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, edited by R. Jay Wallace, Philip Petit, Samuel Scheffler, and Michael Smith. Oxford: Clarendon Press, 2004.

Campbell, John. "A Simple View of Color." In Haldane and Wright, *Reality, Representation, and Projection*.

Chappell, Vere. "Comments." *Pacific Philosophical Quarterly* 85, no. 3 (September 2004): 338–55.

Chisholm, Roderick M. "The Descriptive Element in the Concept of Action." *Journal of Philosophy* 61, no. 20 (October 1964): 613–24.

Clark, Philip. "Velleman's Autonomism." *Ethics* 111, no. 3 (April 2001): 580–93.

Clarke, Samuel. *A Discourse Concerning the Unchangeable Obligations of Natural Religion and the Truth and Certainty of Christian Revelation*. 1705. In Raphael, *British Moralists*.

Cohen, G. A. "Reason, Humanity, and the Moral Law." In Korsgaard, *The Sources of Normativity*.

Cohon, Rachel. "Are External Reasons Impossible?" *Ethics* 96, no. 3 (April 1986): 545–56.

———. "Internalism about Reasons for Action." *Pacific Philosophical Quarterly* 74, no. 4 (December 1993): 265–88.

———. "The Roots of Reasons," *Philosophical Review* 109, no. 1 (January 2000): 63–85.

Collins, Arthur W. "The Psychological Reality of Reasons." *Ratio* 10, no. 2 (September 1997): 108–23.

Conee, Earl. "Comments On Bill Lycan's 'Moore Against The New Skeptics.'" *Philosophical Studies* 103, no. 1 (March 2001): 55–9.

Copp, David. "Against Internalism about Reasons—Gert's Rational Options." *Philosophy and Phenomenological Research* 62, no. 2 (March 2001): 455–61.

———. "The Ring of Gyges: Overridingness and the Unity of Reason." *Social Philosophy and Policy* 14, no. 1 (Winter 1997): 86–106.

Copp, David, ed. *The Oxford Handbook of Ethical Theory*. New York: Oxford University Press, 2006.

Cowley, Christopher. "A New Defence of Williams's Reasons-Internalism." *Philosophical Investigations* 28, no. 4 (October 2005): 346–68.

Cullity, Garrett. Review of *Thinking How to Live*, by Allan Gibbard. *Philosophical Quarterly* 57, no. 227 (April 2007): 308–11.

Cullity, Garrett and Berys Gaut. *Ethics and Practical Reason*. Oxford: Clarendon Press, 1997.

Dancy, Jonathan. "Discussion—On Knowing What One Is Doing." *Philosophical Studies* 21, no. 3 (January 2005): 237–47.

———. *Practical Reality*. Oxford: Oxford University Press, 2000.

———. "Why There Is Really No Such Thing as the Theory of Motivation." *Proceedings of the Aristotelian Society* 95 (1995): 1–18.

Daniels, Norman. "Wide Reflective Equilibrium and Theory Acceptance in Ethics." *Journal of Philosophy* 76, no. 5 (May 1979): 256–82.

Darwall, Stephen. "Autonomist Internalism and the Justification of Morals." *Noûs* 24, no. 2 (April 1990): 257–67.

———. "Because I Want It." *Social Philosophy and Policy* 18, no. 2 (Summer 2001): 129–153.

———. *The British Moralists and the Internal 'Ought': 1640–1740*. Cambridge: Cambridge University Press, 1995.

———. "Ethical Intuitionism and the Motivation Problem." In Stratton Lake, *Ethical Intuitionism*.

———. *Impartial Reason*. Ithaca, NY: Cornell University Press, 1983.

———. "Impartial Reason." In *Contemporary Ethics: Selected Readings*, edited by James P. Sterba. Englewood Cliffs, NJ: Prentice Hall, 1989.

———. "Internalism and Agency." *Philosophical Perspectives* 6 (1992): 155–174.

———. "Learning from Frankena." *Ethics* 107, no. 4 (July 1997): 685–705.

———. "Moore, Normativity, and Intrinsic Value." *Ethics* 113, no. 3 (April 2003): 468–89.

———. "Normativity and Projection in Hobbes's *Leviathan*." *Philosophical Review* 109, no. 3 (July 2000): 313–47.

———. "Reasons, Motives, and the Demands of Morality: An Introduction." In *Moral Discourse and Practice: Some Philosophical Approaches*, edited by Stephen Darwall, Allan Gibbard, and Peter Railton. Oxford: Oxford University Press, 1997.

———. "Reply to Terzis." *Canadian Journal of Philosophy* 18, no. 1 (March 1988): 115–24.

———. Review of *The Authority of Reason*, by Jean Hampton. *Philosophical Review* 109, no. 4 (October 2000): 583–6.

———. Review of *Ought, Reaons, and Morality*, by W. D. Falk. *Journal of Philosophy* 86, no. 4 (April 1989): 208–214

———. *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge, MA: Harvard University Press, 2006.

Darwall, Stephen, Allan Gibbard, and Peter Railton. "Toward *Fin de siècle* Ethics: Some Trends." *Philosophical Review* 101, no. 1 ( January 1992): 115–89.

Davidson, Donald. "Actions, Reasons, and Causes." In *Essays on Actions and Events*.

———. *Essays on Actions and Events*. Oxford: Clarendon Press, 1980.

———. "Freedom to Act." In *Essays on Actions and Events*.

———. "Mental Events." In *Essays on Actions and Events*.

———. "Psychology as Philosophy." In *Essays on Actions and Events*.

Dreier, James. "Humean Doubts about the Practical Justification of Morality." In Cullity and Gaut, *Ethics and Practical Reason*.

———. "Internalism and Speaker Relativism." *Ethics* 101, no. 1 (October 1990): 6–25.

———. "Meta-ethics and the Problem of Creeping Minimalism." *Philosophical Perspectives* 18 (2004): 23–44.

———. "Moral Relativism and Moral Nihilism." In Copp, *The Oxford Handbook of Ethical Theory*.

Dreier, James, ed. *Contemporary Debates in Moral Theory*. Oxford: Blackwell Publishing, 2006.

Dworkin, Ronald. "Objectivity and Truth: You'd Better Believe It." *Philosophy and Public Affairs* 25, no. 2 (Spring 1996): 87–139.

———. "Dworkin Replies to Critics." *Brown Electronic Article Review Service* (April 9, 1997), http://www.brown.edu/Departments/Philosophy/bears/9704dwor.html.

Enoch, David. "Agency, Shmagency: Why Normativity Won't Come from What Is Constitutive of Action." *Philosophical Review* 115, no. 2 (April 2006): 169–98.

———. "An Outline of an Argument for Robust Metanormative Realism." *Oxford Studies in Metaethics* 2 (2007): 21–50.

———. "Why Idealize?" *Ethics* 115, no. 4 ( July 2005): 759–87.

Ewing, A. C. *The Definition of Good*. New York: Macmillan, 1947.

Falk, W. D. "Action-Guiding Reasons." In *Ought, Reasons, and Morality*.

———. "Morality, Self, and Others. In *Ought, Reasons, and Morality*.

———. "'Ought' and Motivation." In *Ought, Reasons, and Morality*.

———. *Ought, Reasons, and Morality: The Collected Papers of W. D. Falk*. Ithaca: Cornell University Press, 1986.

Feldman, Fred. "The Open Question Argument: What It Isn't; and What It Is." *Philosophical Issues* 15 (2005): 22–43.

Field, Hartry H. *Science without Numbers: A Defense of Nominalism.* Princeton: Princeton University Press, 1980.

FitzPatrick, William J. "The Practical Turn in Ethical Theory: Korsgaard's Constructivism, Realism, and the Nature of Normativity." *Ethics* 115, no. 4 (July 2005): 651–91.

———. "Reasons, Value, and Particular Agents: Normative Relevance without Motivational Internalism." *Mind* 113, no. 450 (April 2004): 285–318.

Foot, Philippa. "Morality as a System of Hypothetical Imperatives." In *Virtues and Vices and Other Essays in Moral Philosophy.* Oxford: Clarendon Press, 2002.

Frankena, William K. "The Naturalistic Fallacy." In *Perspectives on Morality.*

———. "Obligation and Motivation in Recent Moral Philosophy." In *Perspectives on Morality.*

———. "Obligation and Value in the Ethics of G. E. Moore." In *Perspectives on Morality.*

———. "On Saying the Ethical Thing. In *Perspectives on Morality.*

———. *Perspectives on Morality: Essays of William K. Frankena,* edited by Kenneth E. Goodpaster. Notre Dame: University of Notre Dame Press, 1976.

Frankfurt, Harry G. "Freedom of the Will and the Concept of a Person." In *The Importance of What We Care About: Philosophical Essays.* Cambridge: Cambridge University Press, 1988.

Garner, Richard T. "On the Genuine Queerness of Moral Properties and Facts." *Australasian Journal of Philosophy* 68, no. 2 (June 1990): 137–46.

Gert, Bernard. *Morality: Its Nature and Justification.* New York: Oxford University Press, 1998.

Gert, Joshua. "Avoiding the Conditional Fallacy." *Philosophical Quarterly* 52, no. 206 (January 2002): 88–95.

———. "Skepticism about Practical Reasons Internalism." *Southern Journal of Philosophy* 39, no. 1 (Spring 2001): 59–77.

Gibbard, Allan. "Knowing What to Do, Seeing What to Do." In Stratton-Lake, *Ethical Intuitionism.*

———. "Morality as Consistency in Living: Korsgaard's Kantian Lectures. *Ethics* 110, no. 1 (October 1999): 140–164.

———. *Thinking How to Live.* Cambridge, MA: Harvard University Press, 2003.

———. "Truth and Correct Belief." *Philosophical Issues* 15 (2005): 338–50.

———. *Wise Choices, Apt Feelings: A Theory of Normative Judgment.* Cambridge, MA: Harvard University Press, 1990.

Gödel, Kurt. "A Remark about the Relationship between Relativity Theory and Idealistic Philosophy." In *Albert Einstein: Philosophy-Scientist,* edited by Paul A. Schilpp. La Salle, IL: Open Court, 1949.

Goldman, Alvin I. *A Theory of Human Action.* Englewood Cliffs, NJ: Prentice Hall, 1970.

Goodman, Michael and Robert Snyder, eds. *Contemporary Readings in Epistemology.* Englewood Cliffs, NJ: Prentice Hall, 1993.

Griffin, James. *Value Judgement: Improving Our Ethical Beliefs.* Oxford: Clarendon Press, 1996.

———. "Values: Reduction, Supervenience, and Explanation by Ascent." In *Reduction, Explanation, and Realism,* edited by David Charter and Kathleen Lennon. Oxford: Clarendon Press, 1992.

———. *Well-Being: Its Meaning, Measurement and Moral Importance.* Oxford: Clarendon Press, 1986.

Hacyan, Shahen. "On the Transcendental Ideality of Space and Time in Modern Physics." *Kant-Studien* 97, no. 3 (2006): 382–95.

Haldane, John and Crispin Wright, eds. *Reality, Representation, and Projection.* New York: Oxford University Press, 1993.

Hale, Bob and Crispin Wright. *The Reason's Proper Study: Essays towards a Neo-Fregean Philosophy of Mathematics.* Oxford: Oxford University Press, 2001.

Hampton, Jean E. *The Authority of Reason.* Cambridge: Cambridge University Press, 1998.

Hare, R. M. *Moral Thinking: Its Levels, Method, and Point.* Oxford: Clarendon Press, 1981.

———. "'Nothing Matters': Is 'the Annihilation of Values' Something That Could Happen?" In *Applications of Moral Philosophy.* Berkeley and Los Angeles: University of California Press, 1973.

Harman, Gilbert. "Moral Relativism Defended." In *Explaining Value and Other Essays in Moral Philosophy.* Oxford: Clarendon Press, 2000.

———. *The Nature of Morality: An Introduction to Ethics.* New York: Oxford University Press, 1977.

Hieronymi, Pamela. "The Wrong Kind of Reason." *Journal of Philosophy* 102, no. 9 (September 2005): 437–57.

Hubin, Donald C. "Hypothetical Motivation." *Noûs* 30, no. 1 (March 1996): 31–54.

Hume, David. *A Treatise of Human Nature,* edited by L. A. Selby-Bigge. 2nd ed. Oxford: Clarendon Press, 1978.

Hussain, Nadeem J. Z. "The Guise of a Reason." *Philosophical Studies* 121, no. 3 (December 2004): 263–75.

Hussain, Nadeem J. Z. and Nishi Shah. "Misunderstanding Metaethics: Korsgaard's Rejection of Realism." *Oxford Studies in Metaethics* 1 (2006): 265–94.

Johnson, Robert. "Internal Reasons and the Conditional Fallacy." *Philosophical Quarterly* 49, no. 194 (January 1999): 53–71.

———. "Internal Reasons: Reply to Brady, van Roojen, and Gert." *Philosophical Quarterly* 53, no. 213 (October 2003): 573–80.

Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. In *Practical Philosophy*, translated and edited by Mary J. Gregor. Cambridge: Cambridge University Press, 1996.

Kment, Boris. "Counterfactuals and the Analysis of Necessity." *Philosophical Perspectives* 20 (2006): 237–302.

Korsgaard, Christine M. *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press, 1996.

———. "Internalism and the Sources of Normativity." In *Constructions of Practical Reason: Interviews on Moral and Political Philosophy*, edited by Herlinde Pauer-Studer. Stanford: Stanford University Press, 2002.

———. "Morality and Freedom." In *Creating the Kingdom of Ends*.

———. "Normativity, Necessity, and the Synthetic *A Priori*: A Response to Derek Parfit." 2003. http://www.people.fas.harvard.edu/~korsgaar/Korsgaard.on.Parfit.pdf.

———. "The Normativity of Instrumental Reason." In Cullity and Gaut, *Ethics and Practical Reason*.

———. "Realism and Constructivism in Twentieth-Century Moral Philosophy." In *Philosophy in America at the Turn of the Century*. Charlottesville, VA: Philosophy Documentation Center, 2003.

———. "Self-Constitution: Agency, Identity, and Integrity." Locke Lectures, University of Oxford, 2002. http://www.people.fas.havard.edu/~korsgaar/#Locke%20Lectures.

———. "Skepticism about Practical Reason." In *Creating the Kingdom of Ends*.

———. *The Sources of Normativity*, edited by Onora O'Neill. Cambridge: Cambridge University Press, 1996.

———. "Two Distinctions in Goodness." In *Creating the Kingdom of Ends*.

Lavin, Doug. "Practical Reason and the Possibility of Error." *Ethics* 114, no. 3 (April 2004): 424–57.

Lehrer, Keith. "Why Not Skepticism?" In Goodman and Snyder, *Contemporary Readings in Epistemology*.

Lennon, Thomas M. "Through a Glass Darkly: More on Locke's Logic of Ideas." *Pacific Philosophical Quarterly* 85, no. 3 (September 2004): 322–37.

Lewis, David. "Dispositional Theories of Value." In *Papers in Ethics and Social Philosophy*. Cambridge: Cambridge University Press, 2000.

———. *On the Plurality of Worlds*. Oxford: Blackwell Publishing, 1986.

Loeb, Don. "Full-Information Theories of Individual Good." *Social Theory and Practice* 21, no. 1 (Spring 1995): 1–30.

Lycan, William G. "Moore Against the Skeptics," *Philosophical Studies* 103, no. 1 (March 2001): 35–53.

Mackie, J. L. "Counterfactuals and Causal Laws." In *Analytical Philosophy*, edited by Ronald J. Butler. New York: Barnes and Noble, 1962.

———. *Ethics: Inventing Right and Wrong*. London: Penguin Books, 1977.

———. *Truth, Probability and Paradox: Studies in Philosophical Logic*. Oxford: Clarendon Press, 1973.

McDowell, John. "Aesthetic Value, Objectivity, and the Fabric of the World." In *Mind, Value, and Reality*.

———. "Are External Reasons Possible?" In *Mind Value, and Reality*.

———. *Mind, Value, and Reality*. Cambridge, MA: Harvard University Press, 1998.

———. "Values and Secondary Qualities." In *Mind, Value, and Reality*.

McNaughton, David and Piers Rawling. "Naturalism and Normativity: Descriptivism, Normativity and the Metaphysics of Reasons." *Proceedings of the Aristotelian Society: Supplementary Volume* 77 (2003): 23–45.

Mele, Alfred R. *Autonomous Agents: From Self-Control to Autonomy*. Oxford: Oxford University Press, 2001.

———. "Velleman on Action and Agency." *Philosophical Studies* 21, no. 3 (January 2005): 249–61.

Miller, Richard. "Ways of Moral Learning." Philosophical Review 94, no. 4 (October 1985): 507–56.

Millgram, Elijah. "Was Hume a Humean?" *Hume Studies* 21, no. 1 (April 1995): 75–93.

———. "Williams' Argument against External Reasons." *Noûs* 30, no. 2 (June 1996): 197–220.

Moore, G. E. "The Conception of Intrinsic Value." In Moore, *Principia Ethica*.

———. *Principia Ethica*, edited by Thomas Baldwin. Rev. ed. Cambridge: Cambridge University Press, 1993.

———. "Proof of an External World." In *Philosophical Papers*. New York: Macmillan, 1959.

Nagel, Thomas. *The Possibility of Altruism*. Princeton, NJ: Princeton University Press, 1970.

———. *The View from Nowhere*. New York: Oxford University Press, 1986.

Newman, Lex. "Locke on Sensitive Knowledge and the Veil of Perception: Four Misconceptions." *Pacific Philosophical Quarterly* 85, no. 3 (September 2004): 273–300.

Nowell-Smith, P. H. *Ethics*. Oxford: Basil Blackwell, 1957.

Oddie, Graham. *Value, Reality, and Desire*. Oxford: Clarendon Press, 2005.

O'Hagen, Emer. "Belief, Normativity and the Constitution of Agency." *Philosophical Explorations* 8, no. 1 (March 2005): 39–52.

———. "Practical Identity and the Constitution of Agency." *Journal of Value Inquiry* 38, no. 1 (March 2004): 49–59.

Parfit, Derek. "Normativity." *Oxford Studies in Metaethics* 1 (2006): 325–80.

———. "Normativity, Naturalism, and Noncognitivism." Unpublished manuscript, 2007.

———. "Reasons and Motivation." *Proceedings of the Aristotelian Society: Supplementary Volume* 71 (1997): 99–130.

———. *Reasons and Persons*. Oxford: Clarendon Press, 1984.

Pettit, Philip and Michael Smith. "Backgrounding Desire." In *Mind, Morality, and Explanation: Selected Collaborations*, by Frank Jackson, Philip Pettit, and Michael Smith. Oxford: Clarendon Press, 2004.

Platts, Mark. "Moral Reality and the End of Desire." In Reference, Truth and Reality: *Essays on the Philosophy of Language*, edited by Mark Platts. London: Routledge & Kegan Paul, 1980.

Price, Richard. *A Review of the Principal Questions in Morals*. 1758, 3rd ed. in 1787. In Raphael, *British Moralists*.

Prichard, H. A. "Does Moral Philosophy Rest on a Mistake?" In *Moral Writings*.

———. "Duty and Interest." In *Moral Writings*.

———. "Manuscript on Morals." In *Moral Writings*.

———. *Moral Writings*, edited by Jim MacAdam. Oxford: Clarendon Press, 2002.

Proops, Ian. "Soames on the Metaphysics and Epistemology of Moore and Russell." *Philosophical Studies* 129, no. 3 (June 2006): 632–5.

Putnam, Hilary. *Reason, Truth and History*. Cambridge: Cambridge University Press, 1981.

Quinn, Warren. "Putting Rationality in Its Place." In *Morality and Action*. Cambridge: Cambridge University Press, 1993.

Radzik, Linda. "A Coherentist Theory of Normative Authority." *Journal of Ethics* 6, no. 1 (2002): 21–42.

———. "Incorrigible Norms: Foundationalist Theories of Normative Authority." *Southern Journal of Philosophy* 38, no. 4 (Winter 2000): 633–49.

———. "Justification and the Authority of Norms." *Journal of Value Inquiry* 34, no. 4 (December 2000): 451–61.

———. "A Normative Regress Problem." *American Philosophical Quarterly* 36, no. 1 (January 1999): 35–47.

Railton, Peter. "Facts and Values." In *Facts, Values, and Norms*.

———. *Facts, Values, and Norms: Essays toward a Morality of Consequence*. Cambridge: Cambridge University Press, 2003.

———. "In Search of Nonsubjective Reasons." In *Reason, Ethics, and Society: Themes from Kurt Baier, with His Responses*, edited by J. B. Schneewind. Chicago: Open Court, 1996.

———. "Moral Explanation and Moral Objectivity." *Philosophy and Phenomenological Research* 58, no. 1 (March 1998), 175–182.

———. "Moral Realism." In *Facts, Values, and Norms*.

———. "Naturalism and Prescriptivity." *Social Philosophy and Policy* 7, no. 1 (Winter 1989): 151–174.

———. "Normative Force and Normative Freedom: Hume and Kant, but Not Hume *Versus* Kant. In *Facts, Values, and Norms*.

———. "On the Hypothetical and Non-hypothetical in Reasoning about Belief and Action." In *Facts, Values, and Norms*.

———. "Reply to David Wiggins." In Haldane and Wright, *Reality, Representation, and Projection*.

———. "Some Questions about the Justification of Morality." *Philosophical Perspectives* 6 (1992): 27–53.

———. "Subject-ive and Objective." In *Truth in Ethics*, edited by Brad Hooker. Oxford: Blackwell Publishers, 1996.

———. "What the Non-Cognitivist Helps Us to See the Naturalist Must Help Us to Explain." In Haldane and Wright, *Reality, Representation, and Projection*.

Raphael, D. D., ed. *British Moralists: 1650–1800*. 2 vols. Oxford: Clarendon Press, 1969.

Rawls, John. *A Theory of Justice*. Rev. ed. Cambridge, MA: Belknap Press, 1999.

Raz, Joseph. "The Amoralist." In *Engaging Reason*.

———. *Engaging Reason: On the Theory of Value and Action*. Oxford: Oxford University Press, 1999.

———. "Explaining Normativity: On Rationality and the Justification of Reason." In *Engaging Reason*.

———. "Reason and the Will." In *Engaging Reason*.

Regan, Donald H. "How to Be a Moorean." *Ethics* 113, no. 3 (April 2003): 651–77.

Reid, Thomas. *Essays on the Active Powers of Man*. 1788. In Raphael, *British Moralists*.

Richards, David A. J. *A Theory of Reasons for Action*. Oxford: Clarendon Press, 1971.

Robertson, Theresa. "Internalism, (Super)fragile Reasons, and the Conditional Fallacy." *Philosophical Papers* 32, no. 2 (July 2003): 171–84.

Rogers, G. A. J. "Locke and the Objects of Perception." *Pacific Philosophical Quarterly* 85, no. 3 (September 2004): 245–54.

Roper, Andrew. "Towards an Eliminative Reduction of Possible Words." *Philosophical Quarterly* 32, no. 126 (January 1982): 45–59.

Rorty, Richard. *Philosophy and the Mirror of Nature*. Princeton: Princeton University Press, 1979.

Rosati, Connie. "Agency and the Open Question Argument." *Ethics* 113, no. 3 (April 2003): 490–527.

———. "Brandt's Notion of Therapeutic Agency." *Ethics* 110 no. 4 (July 2000): 780–811.

———. "Internalism and the Good for a Person." *Ethics* 106, no. 2 (January 1996): 297–326.

———. "Naturalism, Normativity, and the Open Question Argument." *Noûs* 29, no. 1 (March 1995): 46–70.

———. "Persons, Perspectives, and Full Information Accounts of the Good." *Ethics* 105, no. 2 (January 1995): 296–325.

Ross, W. D. *The Foundations of Ethics*. Oxford: Clarendon Press, 1939.

———. *The Right and the Good*, edited by Philip Stratton-Lake. Oxford: Clarendon Press, 2002.

Sayre-McCord, Geoffrey, ed. *Essays on Moral Realism*. Cornell University Press, 1988.

Scanlon, T. M. *What We Owe to Each Other*. Cambridge: MA: Belknap Press, 1998.

Schroeder, Mark. "Realism and Reduction: The Quest for Robustness." *Philosophers' Imprint* 5, no. 1 (February 2005). http://www.philosophersimprint.org/005001/.

Schueler, G. F. *Desire: Its Role in Practical Reason and the Explanation of Action*. Cambridge, MA: MIT Press, 1995.

———. *Reasons and Purposes: Human Rationality and the Teleological Explanation of Action*. Oxford: Clarendon Press, 2003.

Sellars, Wilfrid. "Empiricism and the Philosophy of Mind." In *Science, Perception and Reality*. London: Routledge & Kegan Paul, 1963.

Setiya, Kieran. "Explaining Action." *Philosophical Review* 112, no. 3 (July 2003): 339–93.

Shafer-Landau, Russ. *Moral Realism: A Defence*. Oxford: Clarendon Press, 2003.

Shah, Nishi. "How Truth Governs Belief." *Philosophical Review* 112, no. 4 (October 2003): 447–82.

Shah, Nishi and J. David Velleman. "Doxastic Deliberation." *Philosophical Review* 114, no. 4 (October 2005): 497–534.

Sidelle, Alan. *Necessity, Essence, and Individuation: A Defense of Conventionalism*. Ithaca, NY: Cornell University Press, 1989.

Sidgwick, Henry. "Some Fundamental Ethical Controversies." In *Essays on Ethics and Method*, edited by Marcus G. Singer. Oxford: Clarendon Press, 2000.

Skorupski, John. "The Ontology of Reasons." *Topoi* 21, nos. 1–2 ( January 2002): 113–24.

———. "Propositions about Reasons." *European Journal of Philosophy* 14, no. 1 (April 2006): 26–48.

Smith, Michael. *Ethics and the A Priori: Selected Essays on Moral Psychology and Meta-Ethics.* Cambridge: Cambridge University Press, 2004.

———. "Internalism's Wheel." In *Ethics and the A Priori.*

———. "Internal Reasons." In *Ethics and the A Priori.*

———. *The Moral Problem.* Oxford: Blackwell Publishers, 1994.

———. "Search for the Source." *Philosophical Quarterly* 49, no. 196 ( July 1999): 384–94.

Sobel, David. "Explanation, Internalism, and Reasons for Action." *Social Philosophy and Policy* 18, no. 2 (Summer 2001): 218–35.

———. "Full Information Accounts of Well-Being." *Ethics* 104, no. 4 ( July 1994): 784–810.

———. "Reply to Robertson." *Philosophical Papers* 32, no. 2 ( July 2003): 185–91.

Sosa, David. "Pathetic Ethics." In *Objectivity in Law and Morals*, edited by Brian Leiter. Cambridge: Cambridge University Press, 2001.

Stevenson, Charles L. "The Emotive Meaning of Ethical Terms." In *Facts and Values: Studies in Ethical Analysis.* New Haven, CT: Yale University Press, 1963.

Stocker, Michael. "Desiring the Bad: An Essay in Moral Psychology." *Journal of Philosophy* 76, no. 12 (December 1979): 738–53.

Stratton-Lake, Philip. Introduction to Ross, *The Right and the Good.*

———. Introduction to Stratton-Lake, *Ethical Intuitionism.*

———. "Why Externalism Is Not a Problem for Ethical Intuitionists." *Proceedings of the Aristotelian Society* 99 (1999): 77–90.

Stratton-Lake, Philip, ed. *Ethical Intuitionism: Re-evaluations.* Oxford: Clarendon Press, 2002.

Stratton-Lake, Philip and Brad Hooker. "Scanlon versus Moore on Goodness." In *Metaethics after Moore*, edited by Terry Horgan and Mark Timmons. Oxford: Clarendon Press, 2006.

Stroud, Barry. *The Significance of Philosophical Scepticism.* Oxford: Oxford University Press, 1984.

Sturgeon, Nicholas. "Ethical Naturalism." In Copp, *The Oxford Handbook of Ethical Theory.*

———. "Moore on Ethical Naturalism." *Ethics* 113, no. 3 (April 2003): 528–56.

———. "Moral Explanations." In Sayre-McCord, *Essays on Moral Realism.*

Svavarsdóttir, Sigrún. "How Do Moral Judgments Motivate?" In Dreier, *Contemporary Debates in Moral Theory.*

———. "Moral Cognitivism and Motivation." *Philosophical Review* 108, no. 2 (April 1999): 161–219.

Unger, Peter. "An Argument for Skepticism." In Goodman and Snyder, *Contemporary Readings in Epistemology*.

van Roojen, Mark. "Motivational Internalism: A Somewhat Less Idealized Account." *Philosophical Quarterly* 50, no. 199 (April 2000): 233–41.

Väyrynen, Pekka. "Resisting the Buck-Passing Account of Value." *Oxford Studies in Metaethics* 1 (2006): 295–324.

Velleman, J. David. "The Aim of Belief." In *The Possibility of Practical Reason*.

———. "Brandt's Definition of 'Good'." *Philosophical Review* 97, no. 3 (July 1998): 353–71.

———. "A Brief Introduction to Kantian Ethics." In *Self to Self*.

———. "The Centered Self." In *Self to Self*.

———. "Deciding How to Decide." In *The Possibility of Practical Reason*.

———. "From Self Psychology to Moral Philosophy." In *Self to Self*.

———. "The Guise of the Good." In *The Possibility of Practical Reason*.

———. "How We Get Along." Shearman Lectures, University College, London, 2007. http://homepages.nyu.edu/~dv26/Work/Shearman.pdf.

———. Introduction to *The Possibility of Practical Reason*.

———. "Motivation by Ideal." In *Self to Self*.

———. "Is Motivation Internal to Value?" In *The Possibility of Practical Reason*.

———. "The Possibility of Practical Reason." In *The Possibility of Practical Reason*.

———. *The Possibility of Practical Reason*. Oxford: Clarendon Press, 2000.

———. *Practical Reflection*. Princeton: Princeton University Press, 1989.

———. "Précis of *The Possibility of Practical Reason*." *Philosophical Studies* 121, no. 3 (December 2004): 225–38.

———. "Replies to Discussion of *The Possibility of Practical Reason*." *Philosophical Studies* 121, no. 3 (December 2004): 277–98.

———. "The Self as Narrator." In *Self to Self*.

———. *Self to Self: Selected Essays*. Cambridge: Cambridge University Press, 2006.

———. "What Happens When Someone Acts." In *The Possibility of Practical Reason*.

Wallace, R. Jay. "How to Argue about Practical Reason." In *Normativity and the Will*.

———. "Moral Motivation." In Dreier, *Contemporary Debates in Moral Theory*.

———. *Normativity and the Will: Selected Papers on Moral Psychology and Practical Reason.* Oxford: Clarendon Press, 2006.

———. "Normativity, Commitment, and Instrumental Reason." In *Normativity and the Will.*

———. "Practical Reason." In *Stanford Encyclopedia of Philosophy* (2003), http://plato.stanford.edu/entries/practical-reason/.

Warnock, G. J. *Contemporary Moral Philosophy.* London: Macmillan, 1967.

Wedgwood, Ralph. "Practical Reasoning as Figuring Out What Is Best: Against Constructivism." *Topoi* 21, nos. 1–2 (January 2002): 139–52.

Wiggins, David. "Reply to Peter Railton." In Haldane and Wright, *Reality, Representation, and Projection.*

———. "A Sensible Subjectivism?" In *Needs, Values, Truth.* 3rd ed. Oxford: Oxford University Press, 1998.

Williams, Bernard. "Deciding to Believe." In *Problems of the Self: Philosophical Papers, 1956–1972.* Cambridge: Cambridge University Press, 1973.

———. "Internal and External Reasons." In *Moral Luck: Philosophical Papers 1973–1980.* Cambridge: Cambridge University Press, 1981.

———. "Internal Reasons and the Obscurity of Blame." In *Making Sense of Humanity and Other Philosophical Papers: 1982–1993.* Cambridge: Cambridge University Press, 1995.

———. "Postscript: Some Further Notes on Internal and External Reasons." In *Varieties of Practical Reasoning,* edited by Elijah Millgram. Cambridge: MIT Press, 2001.

———. "Replies." In World, Mind, and Ethics: *Essays on the Ethical Philosophy of Bernard Williams,* edited by J. E. J. Altham and Ross Harrison. Cambridge: Cambridge University Press, 1995.

Wilson, George M. *The Intentionality of Human Action.* Rev. ed. Stanford: Stanford University Press, 1989.

Yaffe, Gideon. "Locke on Ideas of Substance and the Veil of Perception." *Pacific Philosophical Quarterly* 85, no. 3 (September 2004): 255–72.

Zangwill, Nick. "Zangwill Reviews Dworkin." *Brown Electronic Article Review Service* (December 2, 1996), http://www.brown.edu/Departments/Philosophy/bears/9612zang.html.