

## Sympathy for the Devil

*Kieran Setiya*

We tend to want what we perceive as being in some way good, to choose what seems worthy of choice, and to act in ways we think we can justify, at least to some extent.<sup>1</sup> The question for action theory is not whether to accept or deny these platitudes about human agency, but how to interpret and explain them. Are they contingent generalizations, principles of human nature, or something more? Does a connection between agency and evaluation belong to the essence of intentional action, and thus to any adequate account of the capacity to act for reasons?

In what follows, I will argue that it does not. What I say expands upon, and qualifies, a previous discussion of the topic, the gist of which appears in section 2.<sup>2</sup> My aim, however, is not to respond to objections or to resume a polemic. Much of the argument below is devoted to understanding the different ways in which the exercise of rational agency might depend on judgments or appearances of the good, the logical relations among such claims, and the scope for views that concede some truth in what might seem to be their opposites. This approach will tend to blur some contrasts and bring into relief the ones that matter most.

Begin with a distinction and a first concession. There are two contrasting directions of dependence that might be thought to hold between desire, or intentional action, and the good. On the one hand, there are forms of “motivational internalism” whose shape is roughly this:

If A believes that  $x$  is good, or perceives it as being good, then she desires it.

On the other hand, there is the doctrine that we act intentionally “under the guise of the good”:

If A is doing  $\phi$  for reasons, or doing it intentionally, she sees some good in doing it.

These formulations are deliberately crude. The point of introducing them is to set aside the question of motivational internalism, and to focus on the guise of the good. Let it be a necessary truth that judgments or appearances of the good have some defeasible influence on action. It

would be a further claim that attitudes of this kind are involved whenever we exercise the capacity to act for reasons. One can accept motivational internalism, even in the most unguarded form, while denying that we act intentionally under the guise of the good.

Our topic can be further refined. It is the guise of the good, understood as a claim about rational agents, as such: about what it is to act for reasons or to act intentionally. It is not sufficient for the truth of this claim that, as it happens, we tend to act in ways we see as being in some way good, plausible though that is. Nor would it be sufficient if this were a fact of human nature, of the sort to be described in section 4. My conjecture is that part of the appeal of the guise of the good, in its action-theoretic form, derives from being conflated with doctrines of this other kind. At any rate, what I reject is the claim that it belongs to rational agency, in the abstract, to be exercised under the guise of the good, so that the need for a positive evaluation of action can be derived from the bare idea of its being intentional or being done for reasons.

Objections to the guise of the good often take the form of examples, cases of perversity or depression in which it is argued that someone acts in a way that she regards as bad without qualification, or simply finds indifferent.<sup>3</sup> Without elaboration, however, this strategy is bound to fail. The description of the examples is controversial, and advocates of the guise of the good will find ingenious ways in which to make sense of them.<sup>4</sup> What is more, taken by themselves, the examples leave untouched the *grounds* for accepting the guise of the good. How does our reflection on agency push us toward it? And where does it go wrong? Without answers to these questions, the examination of cases, however subtle and psychologically perceptive, will be unsatisfying.<sup>5</sup> The point cuts both ways. If descriptions of spiteful action do not refute the guise of the good, it is not established by Anscombe's (1963: 70–71) well-known remarks about wanting a saucer of mud. She doubts that anyone could have this desire unless he can say what makes its object desirable. But the example is inconclusive. To begin with, there is a problem of indeterminacy: It is not clear what the object of the man's desire is meant to be. When we have a desire, what we want is to act in a certain way or for something to be the case; desire is for an action or an outcome. The ascription of desire for an object is always elliptical. As soon as we fill this gap, Anscombe's desire begins to seem possible. Someone may want to own a saucer of mud, to have one in his hand even for a moment, for one to exist, without seeing any good in any of these things. Apart from a more theoretical investigation of rational agency, we won't know how seriously to take such appearances—that Anscombe's example is possible, or that it isn't. For the most part, then, my arguments will turn on rather abstract principles of reasons-explanation. (Even when they do rely on examples of acting for reasons, as in section 3, the examples will be used to make a systematic point.) Despite my title, nothing will rest on thoughts about the proper interpretation of Milton's Satan.<sup>6</sup>

The structure of the chapter is as follows. In section 1, I argue that the guise of the good is fundamentally a claim about *reasons*, not about *desires*. In section 2, I argue against a simple version of the guise of the good. It may be true that, in acting for reasons, one must know one's reason for acting, but what one knows is a reason that explains what one is doing, not a reason that purports to justify it. In the remaining sections, I examine more sophisticated versions of the guise of the good. It might be held, for instance, that the kind of explanation one gives of what one is doing, in doing it for a reason, has normative or evaluative content, and thus involves a qualified but positive evaluation of one's action. That proposal is the target of section 3. Finally, it might be held that the preceding arguments mistake the kind of generality the guise of the good is meant to have, wrongly treating it as a claim about every instance of intentional action. A different sort of generality can be employed in specifying the nature of something, as when we say that cats have four legs or human beings have thirty-two teeth, allowing for car accidents and British dentistry.<sup>7</sup> Once we have such "generic" essentialism in view, the guise of the good may seem to be revived. I end by considering this move in section 4, arguing that, while the guise of the good may be a principle of human nature, an account of rational agency cannot be given in generic terms. This result sheds incidental light on the foundations of practical reason.

## 1. DESIRING THE GOOD

On the face of it, there are two quite different ways in which the guise of the good could turn out to be true, one resting on the concept of a reason, the other on the concept of desire:

(R) If someone acts intentionally in doing  $\phi$ , she is doing it for a reason, and reasons must be seen as good.

(D) If someone acts intentionally in doing  $\phi$ , she is acting on a desire, and desires represent their objects as good.

Many hold views in the vicinity of (R). They follow Davidson (1963: 6) in "defining an intentional action as one done for a reason" and think of acting for a reason as acting on a ground one takes to *be* a reason—that is, at least a partial or *pro tanto* justification—for what one is doing.<sup>8</sup>

The basic case [of acting for a reason] must be that in which  $A$   $\phi$ 's, not because he believes only that there is some reason or other for him to  $\phi$ , because he believes of some determinate consideration that it constitutes a reason for him to  $\phi$ . (Williams 1979: 107)

I cannot act for reasons if I do not care about doing what's justified or (as I would prefer to put it) what makes sense. (Velleman 1992a: 121)

Both choice and decision are subject to rules of rational constraint, the most important of which is that one can only choose or decide for a reason, i.e. for what one takes to be a good reason for the option chosen. (Raz 1997: 8)

It is perhaps less common these days to find echoes of (D). But it seems more or less trivial to say that people want to do whatever it is they do intentionally, and desires have been conceived as appearances of the good.

Desire is a kind of perception. One who wants it to be the case that  $p$  perceives something that makes it seem to that person as if it would be good were it to be the case that  $p$ , and seem so in a way that is characteristic of perception. (Stampe 1987: 359)

Desiring is, in my view, simply *identified* with conceiving something to be good from a certain perspective. (Tenenbaum 2007: 14)

It is a picture of desire that forms the classical source of the guise of the good, in Plato's *Republic*—"Every soul pursues the good and does whatever it does for its sake" (505e)<sup>9</sup>—and in the moral psychology of Aristotle's *De Anima*: "It is always the object of desire which produces movement, [and] this is either good or the apparent good" (433a27–29).<sup>10</sup> These formulations became a dogma of scholastic philosophy, cited with approval by Kant in the *Critique of Practical Reason*: We desire only what we conceive to be good; we avoid only what we conceive to be bad.<sup>11</sup>

Despite these precedents, the guise of the good is best understood as a claim about reasons, and only derivatively a claim about desire. The argument for this turns on some modest connections among desire, reasons, and the good. First: When the object of desire, an action or outcome, is good, there is always some respect in which it is good, which is a reason to perform or to pursue it. This need not be heard as a reductive claim, an analysis of "good" for actions and outcomes on which a good thing for  $A$  to do just is something there is reason for her to do, and a good state of affairs is one there is reason to bring about.<sup>12</sup> What we need is something weaker, that there are "good-making" features of good actions and outcomes that count as reasons of the appropriate kinds, whatever the constitutive story turns out to be.

Second: If desires represent their objects as good, they represent them as being good in some respect—say, in being  $F$ —and the fact that the object is  $F$  is a reason why the agent wants to perform or pursue it. This proposition has two parts. To begin with, it excludes the conception of desire on which it depicts its object as being good but leaves us wholly in the dark when we ask what is good about it, as though this were a matter for guesswork or speculation. That seems absurd. Once we accept that desiring something is conceiving it as good, we should think of desires as presenting the specific appeal of their objects, what it is about them that seems to make them good.<sup>13</sup> This is reflected in the first quotation about desire, above: "One who wants it to be the case that  $p$  perceives *something that makes it seem* to that person as if it would be good were it to be the case that  $p$ " (Stampe 1987: 359, my emphasis). The proposition further implies that the respects in which desire represents its object as being good are reasons why the agent wants it. If sleeping late seems good to me

in being restful, and this constitutes a desire to sleep late, my reason for wanting to sleep late is that it will be restful. If learning history seems good to me as a source of knowledge, and this constitutes a desire, my reason for wanting to learn history is that it is a source of knowledge. In the limiting case, what I want may be something that seems good for its own sake.<sup>14</sup> Why do I want to know things? What seems good about that? Perhaps no more than its being knowledge that I'll have: That is what makes the state of affairs good in which I know things, and my reason for wanting to bring it about. Once we allow for reasons of this kind, we can see that whenever a desire presents its object as being good in being *F*, its being *F* is an answer to the question "Why do you want it?"

The upshot is that, if desires represent their objects as good, they must be had for reasons that are seen as good, at least in being seen as respects in which the *object* of desire is good. In principle, one could concede this point, insist that desires represent their objects as good, and still reject the guise of the good as a general constraint on reasons for wanting. One would have to claim that, while we can want things for reasons that we do not see as good, it is a necessary truth that every desire is had for some reason that is conceived as good. But that position is unstable. If some of my reasons for wanting to be famous are not respects in which being famous seems like a good thing to me, why must I have some other reason through which it does? And if I do, what is to prevent me from realizing that this other reason is misconceived or false without relinquishing the desire for fame, something I now want only for reasons that are not respects in which fame seems good? Only if reasons for desire *must* be seen as good, in the corresponding sense, can we explain what blocks such possibilities. It follows that, if desires represent their objects as good, the following principle holds:

When someone wants to  $\phi$ , or wants it to be the case that  $p$ , they want it for a reason, and reasons for desire must be respects in which the object of desire is seen as good.<sup>15</sup>

Although this is not a claim about the evaluation of reasons, as such, it involves a version of the guise of the good applied to reasons for desire. They need not be seen as good reasons for wanting, perhaps, but they must be seen as respects in which it would be good to  $\phi$  or for it to be the case that  $p$ . Thus, even when it rests on an evaluative conception of desire, the guise of the good for intentional action can be addressed by investigating the nature of reasons. We lose nothing of substance and gain something in generality if we adopt that focus throughout.<sup>16</sup>

## 2. REASONS

In order to do so, we have to be much more careful about the logic of propositions that use concepts of reason, making explicit a distinction

that has been tacitly presupposed so far. Switching from desire to action, there is a contrast between claims of the following two kinds:

The fact that  $p$  is a reason for  $A$  to  $\phi$ .

$A$  is doing  $\phi$  on the ground that  $p$ ; that is his reason for doing it.

To say that there is a reason for  $A$  to  $\phi$  is to say that  $A$  would be *pro tanto* justified in doing it. The justification may not be decisive; it may be outweighed by other reasons. But there is something to be said for doing  $\phi$ , a consideration that counts in favor. Claims of this kind are normative or evaluative; they belong to ethics, broadly conceived. When the fact that  $p$  is a reason for  $A$  to  $\phi$ , we can just as well report that it is a *good* reason for  $A$  to  $\phi$ . In this sense, "bad reasons" are not reasons at all.

To say that  $A$  is doing  $\phi$  for a certain reason, on the other hand, is to give a distinctive kind of explanation. It is to state a ground on which he is acting and thus to account for that action, at least in part. There is disagreement about the connection between explanations of this sort and ones that appeal to psychological states like belief, intention, and desire—"motivating reasons" in the technical sense employed by Michael Smith (1987, 1994), or Davidson's (1963: 3–4) "*primary reason[s]*." For the most part, I will try to be agnostic about that.<sup>17</sup> But we can say, at least, that it is not sufficient for the truth of our second proposition that  $A$  is doing  $\phi$  because he believes that  $p$  and has some relevant desire. This comes out in Davidson's (1973: 79) well-known examples of "causal deviance," as when a nervous climber wants to be rid of his companion's weight and knows that he can manage this by dropping the rope. These attitudes make him anxious, with the result that he inadvertently drops the rope. He does not act for a reason in doing so, despite the causal role of his belief. Giving someone's reason for acting is not just citing a belief that is a cause of action; it implies that he is acting *on* that reason, which bears on what he is doing in a more intimate way.

Our topic is not the problem of causal deviance but the connection between explanations that give our reasons for doing things and reasons that justify doing them. Or, to return to section 1, it is the connection between agent's reasons for wanting things and reasons that would justify those desires by showing their objects to be good. What we can note at once is this: The propositions distinguished above differ in that the first is factive and the second is not; and neither of them entails the other. In *Practical Reality*, Jonathan Dancy gives examples in which someone acts for a reason that turns out to be false:

His reason for doing it was that it would increase his pension, but in fact he was quite wrong about that.

The ground on which he acted was that she had lied to him, though actually she had done nothing of the sort. (Dancy 2000: 132)

These descriptions are sometimes questioned, though they seem innocent enough to me. Accepting them as true is far less contentious than accepting

Dancy's further claim that they are instances of irreducibly non-factive explanation. As remarked above, we can afford to be agnostic about the reduction of such claims to propositions about the causality of psychological states. More importantly, even explanatory claims that use "because" to give an agent's reason—"He is doing it because  $p$ "—which arguably do entail the truth of that proposition, and perhaps that the agent knows it to be true, are consistent with its being no reason at all in the justifying sense.<sup>18</sup> We can fail to act for reasons that count in favor of doing something, as when we are ignorant of them, and we can act for reasons that don't. That the house is on fire may be no reason for me to flee when my wife and child are sleeping upstairs; still I can run outside because of it. And it is a notorious fact that wicked pleasures do not provide us with reasons to act; but we can act in pursuit of them nonetheless.

None of this conflicts with the simplest version of the guise of the good for reasons:

If A is doing  $\phi$  because  $p$ , or on the ground that  $p$ , he *believes* that the fact that  $p$  is a reason for him to  $\phi$ .

According to this principle, agents' reasons reveal the positive light in which an agent saw his action by giving a consideration that he took to justify it, perhaps not consciously, perhaps not adequately or sufficiently, but to some extent. This is the sort of claim endorsed by the authorities cited at the beginning of section 1. It could be adapted to reasons for desire, which are believed to be respects in which the object of desire is good: worth doing or worth pursuing. And each claim could be weakened by replacing beliefs with appearances or facts about how things seem.

The crucial argument against the guise of the good for reasons, in this simple form, turns on the nature of intentional action and thus of action done for reasons. What it requires is not a tendentious theory of action-explanation, but some observations about the subject matter of action theory that appear at the beginning of Anscombe's *Intention*.<sup>19</sup>

What distinguishes actions which are intentional from those which are not? The answer that I shall suggest is that they are actions to which a certain sense of the question "Why?" is given application; the sense is of course that in which the answer, if positive, gives a reason for acting. (Anscombe 1963: 9)

Apart from its deliberately conditional note—the question "Why?" is "not refused application because the answer to it says that there is *no* reason" (Anscombe 1963: 25)—this seems hard to deny. When someone acts for a reason, in the explanatory sense, the question "Why?" has application to what they are doing, and they count as doing it intentionally. But Anscombe has a stronger premise in mind, not just that the question *has* application, but that the agent *gives* it application in being able to answer it. She imagines the question "Why?" being put to the person who is doing  $\phi$  and the answer as constituted by his response. Hence her otherwise

puzzling doubts about whether we can count on agents to be honest.<sup>20</sup> We can avoid one source of difficulty here by crossing the gap between belief and its linguistic expression. The pivotal claim is that the answer to the question "Why?" understood as a request for reasons has to be something that the agent who is acting for that reason *believes*.

This picture of what is involved in acting for a reason is not innocuous, but it is, so far, relatively weak.<sup>21</sup> In the present context, any doubts can be set aside. Our topic is the guise of the good for reasons, and this doctrine is best conceived as an interpretation of, and therefore as committed to, the requirement of belief on answers to the question "Why?" It is the view that when this question is given application, the answer, if positive, is contained in the agent's beliefs about the reasons for acting as he is: If A is doing  $\phi$  on the ground that  $p$ , he believes that the fact that  $p$  is a reason for him to  $\phi$ .

The problem is that, in illuminating the source and structure of the guise of the good for reasons, this perspective casts it in a negative light. The answer to Anscombe's question is an *explanation* of what one is doing and why one is doing it, not—or not explicitly—a justification. Its form is:

I am doing  $\phi$  because  $p$ ,

or,

I am doing  $\phi$  on the ground that  $p$ ; that is my reason for doing it,

not:

The fact that  $p$  is a reason for me to  $\phi$ .

That answering the question "Why?" is, in the first instance, giving an explanation comes out in Anscombe's (1963: 11) further observation that this question "is refused application by the answer: 'I was not aware I was doing that.'"

"Why do you humiliate him by telling that awful story?"

"Does he mind? I thought he'd be amused. I didn't mean to humiliate him—I wasn't doing that intentionally."

Switching again from the linguistic to the psychological mode, we can say that knowing *why* I am doing something—having an answer to Anscombe's question—is a way of knowing *that* I am doing it. Again, the belief that corresponds to the question is the belief that I am doing  $\phi$  because  $p$  or on the ground that  $p$ , the truth of which entails that I am doing it, not that the fact that  $p$  is a reason for me to  $\phi$ , which implies nothing of the sort. Knowing that a fact is a reason to  $\phi$  is not a way of knowing what one is doing. If there is a connection between answering the question "Why?" and believing that one's answer describes a normative reason for one's action, this connection is indirect.

The assumption that in doing  $\phi$  intentionally one must believe that one is doing it may have to be qualified. There are cases in which that

condition appears to fail, as when I am doing  $\phi$  by doing something else intentionally as a means to it, and I am not sure that it is getting done.<sup>22</sup> But the present argument survives. What it requires is a claim about sufficiency, not necessity. Whether or not one must have an answer to the question "Why?" in order to be doing  $\phi$  intentionally, it is sufficient to answer this question that one has a belief of the form, "I am doing  $\phi$  because  $p$ ," in the sense of "because" on which this entails that I am doing  $\phi$  on the ground that  $p$  and therefore acting for a reason. The object of belief here is a proposition about the explanation of action. That it is cast in the first person cannot alter its logical powers. It does not follow from the fact that I am doing  $\phi$  because  $p$  that the fact that  $p$  is a reason for me to  $\phi$ , any more than it follows when those propositions are about someone else. The truth of my answer to the question "Why?" is thus consistent with the absence of any justification for what I am doing, and so I can give that answer without believing, or being required in consistency to believe, that I have such justification. Having a story about what justifies one's action goes beyond what is involved in having an answer to the question "Why?" More generally:

- (1) It is sufficient to answer the question "Why?" that one has a belief of the form, "I am doing  $\phi$  because  $p$ ," in the sense of "because" that gives an agent's reason.
- (2) That I am doing  $\phi$  because  $p$ , in this sense, is consistent with the fact that  $p$  not being a reason for me to  $\phi$ .
- (3) If one proposition is consistent with the negation of another, it is possible to believe the first without believing the second.

So:

- (4) It is possible to believe that I am doing  $\phi$  because  $p$ , and thus to answer the question "Why?" without believing that the fact that  $p$  is a reason for me to  $\phi$ .

We can illustrate this in cases of "silencing," where a consideration that would otherwise be a reason to act has no force whatsoever.<sup>23</sup> McDowell (1979: 56) may be wrong to say, in general, that when courage calls for action, "the risk to life and limb [should not be] seen as any reason for removing [oneself]." But there are surely occasions on which that is right, as for instance the one hinted at before, in which I discover that the house is on fire while my wife and child are sleeping upstairs.

"Why are you running outside in your underwear?"

"Because the house is on fire!"

"What about your family? Won't they be trapped by the flames? In a circumstance like this, the fact that the house is on fire is a reason to rush upstairs and rescue them, not to look after your own safety while they burn!"

"You're right. I can't justify my action at all: the danger is not a reason for me to flee; but it is the reason for which I am doing so."

The point of this example is not to carry the weight of the argument, but to clarify its conclusion. Insofar as it is motivated by the idea that in acting for a reason one can answer the question "Why?" the guise of the good is misconceived. The answer to that question is not a proposition about what justifies one's action, but about its explanation.

What else could be the source of the guise of the good for reasons? What could account for the alleged necessity that, in doing  $\phi$  because  $p$  one must believe that the fact that  $p$  is a reason to  $\phi$ ? Once we admit the possibility of someone who meets Anscombe's condition on acting for a reason without conforming to this demand, it is hard to see why any further belief should be required. So long as he is doing  $\phi$  and believes that he is doing  $\phi$  because  $p$ , where this is an explanation that purports to cite his reason, and so long as there is the right sort of connection between the two, a person is acting on the ground that  $p$ . The last resort for the defender of the guise of the good, in application to reasons, is to insist that the right sort of connection must be one that invokes the relevant evaluative belief. Perhaps the problem of causal deviance is solved, in part, by the belief that one's reason does something to justify one's action. But this is hopeless. If there can be the wrong sort of connection between the belief that one is doing  $\phi$  because  $p$  and one's doing it, there can be the wrong sort of connection between doing  $\phi$  and a belief about its justification. Problems about the right connection between attitude and action, as in cases of causal deviance, cannot be solved by adding more beliefs, whose relationship to what one is doing will be equally problematic.

It follows that, although we have not tried to find sufficient conditions of acting for a reason that define it in other terms, we have found sufficient *cognitive* conditions for acting on the ground that  $p$ . One need only believe that one is doing  $\phi$  because  $p$ , so long as there is the right sort of connection between one's action in doing  $\phi$  and this belief. (Doing  $\phi$  because  $p$  may require, in addition, that one's belief amount to knowledge.) Since the right connection need not involve the belief that one's reason for doing  $\phi$  is a good reason for doing it, one need not have that further belief. There is nothing in the cognitive conditions of acting because  $p$ , or on the ground that  $p$ , that could account for its necessary presence. The guise of the good for reasons, at least in the simple form considered so far, is false.<sup>24</sup>

Given the argument of section 1, the same point holds against evaluative conceptions of desire. Recall that, in order to represent their objects as good, desires must be had for reasons, and reasons for desire must be respects in which the object of desire is seen as good. The problem is that, if I am running outside because the house is on fire without believing that this fact provides a reason for flight, I also *want* to run outside for just that reason. No further evaluative belief need be involved. In finding sufficient cognitive conditions for *acting* on the ground that  $p$ , we have also found sufficient cognitive conditions for *desiring* on the ground that  $p$ . Since these conditions do not involve the belief that there is good reason

to flee, or that it would be good to do so, we have found an instance of nonevaluative desire.

Nor would it help the proponent of the guise of the good for intentional action to deny that these are real desires, or to restrict the evaluative conception to a special kind of affective state.<sup>25</sup> For it is crucial to his view that, in acting intentionally, one always acts on a desire. In any case, why should things be different for desires that do not issue in action? Here, too, we can answer the question "Why?" by giving an explanation of our desire that purports to cite our reason "I want to  $\phi$  because  $p$ "—and it is sufficient to count as wanting for a reason that this belief has the right sort of connection with that desire.

"Why do you want to run outside?"

"Because the house is on fire!"

"What about your family? Won't they be trapped by the flames?"

"You're right. The danger is not a reason for me to flee, though it is my reason for wanting to. I should resist temptation and attempt a rescue. I'm going upstairs."

No doubt we sometimes believe that an action would be right or good, or that an outcome is worth pursuing, and want it for the corresponding reasons. But that is not a condition of having a reason for one's desire.

Finally, although the argument so far has focused on evaluative beliefs, nothing changes when we allow for versions of the guise of the good that deal in appearances or how things seem. There are sufficient cognitive conditions for acting or desiring on the ground that  $p$  that do not include beliefs about justifying reasons or about respects in which the object of desire is good; nor do they involve appearances to that effect. Such presentations are not required for us to answer the question "Why?" and they would not help to explain the right connection between beliefs that answer that question and intentional action or desire: If beliefs can figure in the wrong sort of connection, or contribute to causal deviance, so can psychological states in which things seem to be a certain way.

Where do these conclusions leave the guise of the good? They show that one can act for a reason without believing that there is a reason that counts in favor of what one is doing, or any respect in which it is a good thing to do, and without either of these propositions seeming to be true.<sup>26</sup> It does not follow that "reason" is implausibly ambiguous or that there is no connection between the capacity to act for reasons and the capacity to govern one's action by one's conception of what the reasons are. For one thing, the senses of "reason"—justifying and explanatory—are very closely related. A justifying or *good* reason is, roughly speaking, a good thing to have as one's reason for acting; it sets a standard for what one's reasons ought to be.<sup>27</sup> For another, the capacity for evaluative control of one's action depends upon the capacity to know what one is doing and why. If I have no idea what my reasons are, I am no position to stop myself from acting on considerations that are not,

as I believe, good reasons to act. In fact, if I am to put my conception of reasons into practice, I had better know what my reasons are *spontaneously*, without observation or inference. Otherwise, the best I can manage is to aim at acting for good reasons, attend to whether I am doing so, and try to correct them afterward if I am not: a bizarre form of *post hoc* self-management. A line of dependence therefore runs from the capacity for rational self-governance to the kind of self-knowledge that Anscombe associates with acting for reasons.<sup>28</sup> It is even consistent with the failure of the guise of the good as a claim about the answer to the question "Why?" that the capacity to act for reasons depends on the capacity to evaluate them as good or bad, and so to entertain thoughts about the justification of action. We would need a story about why this dependence holds, but there is no principled obstacle to giving one, at least so far.

Nor is there, as yet, a decisive refutation of the guise of the good for intentional action, as opposed to some of the grounds on which it might be held. It may still be true that in acting for reasons one must see one's action in a positive light. What we have is pressure to interpret and defend this doctrine in another way. The skeptical argument of this section can therefore be regarded ecumenically. It helps to indicate the proper *form* of the guise of the good, as a putative constraint on rational agents. The mistake is to think that seeing one's action as good involves a belief—or an appearance; the nature of the attitude is immaterial—whose content goes beyond the proposition that one is acting on the ground that  $p$ . If the guise of the good is to apply to rational agency, as such, it must apply because beliefs of this kind are essential to its exercise, and because the proposition one thus believes already contains the positive light in which one's action is cast, even though it does not entail that the reason for which one is acting is a reason to act in that way, or that it is good to do so, in fact. Explanations of action in terms of reasons must involve some weaker affirmation, so that believing an explanation of that kind amounts to seeing one's action in a positive light. Only if this condition is met will the beliefs about one's action involved in doing it intentionally, or in doing it for reasons, vindicate the spirit of the guise of the good.

### 3. INTELLIGIBILITY

Taken critically, the conclusion of section 2 is that reasons for action and desire need not be seen as good. Taken constructively, its conclusion is that the guise of the good for intentional action must rest, in the end, on a claim about the normative or evaluative character of explanations that appeal to agents' reasons. It must take the following shape:

- (a) In doing  $\phi$  on the ground that  $p$ , one believes that one is acting for that reason, if not in doing  $\phi$  itself, then in taking further means.
- (b) Although the explanation, "A is doing  $\phi$  on the ground that  $p$ ," does not imply that the fact that  $p$  is a reason for A to  $\phi$ , it casts that action

in a positive light. To explain an action in this way is to accept some weaker proposition about the justification for doing  $\phi$  in the circumstance that  $A$  is in.

- (c) Thus, in acting on the ground that  $p$ , one accepts a weak proposition about the justification for one's action, or for taking the relevant means. In this modest sense, one sees some good in what one is doing.
- (d) Since acting intentionally is acting for a reason, it follows that we act intentionally under the guise of the good.

Although this argument could be disputed in several ways, it captures the most compelling source of the guise of the good. Some will balk at its initial premise, which is inspired by the passages from Anscombe already discussed. There are difficult questions there, but since I accept the premise, I am willing to set them aside.<sup>29</sup> Others will suggest, with Anscombe (1963: 25) or Hursthouse (1991), that we can act intentionally for no particular reason, rejecting (d). That objection leaves untouched the core idea that, when we do have reasons for acting, we see what we are doing as in some way good. Our principal focus should thus be on explaining action by giving reasons, the sort of explanation of what one is doing that answers the question "Why?" Do explanations of this kind involve a positive assessment of action, albeit one that is weaker than the claim that there is good reason to perform it?

The conception on which they do is sometimes expressed in terms of intelligibility:

[The] concepts of the propositional attitudes have their proper home in explanations of a special sort: explanations in which things are made intelligible by being revealed to be, or to approximate to being, as they rationally ought to be. This is to be contrasted with a style of explanation in which one makes things intelligible by representing their coming into being as a particular instance of how things generally tend to happen. (McDowell 1985: 328)<sup>30</sup>

If someone's reason for acting makes what he is doing intelligible by showing it to be *approximately* rational, it need not, in fact, be a reason for what he is doing, even when it is true. Still, the explanation casts his action in a positive light. To believe such an explanation is to believe a suitably weak proposition about the justification of action, in the circumstance its agent occupies. In explaining one's action in this way, one would conform to a modest version of the guise of the good.

Although it is offered by McDowell as an interpretation of Davidson on the "constitutive ideal of rationality," the claim that reasons-explanation is in this way normative goes beyond the existence of limits on the degree of irrationality consistent with thought. According to Davidson:

The semantic contents of attitudes and beliefs determine their relations to one another and to the world in ways that meet at least rough standards of consistency and correctness. Unless such standards are met to an adequate degree, nothing can count as being a belief, a pro-attitude, or an intention. (Davidson 1987: 114)

What this passage demands is rough conformity to standards of reason across the whole array of one's psychological states. It does not follow that, in each particular case, the explanation of belief by belief, or the motivation of intentional action, approximates to rationality. Within a profile of beliefs, desires and dispositions that is more or less rational, there may be room for individual episodes of thought that are thoroughly defective or confused. We can consistently hold that thought is subject to the constitutive ideal of rationality and that explanations that appeal to agents' reasons rely on dispositions that approximate to reason only in general, not in every case. This weak constraint does nothing to support the picture of explanation required for the guise of the good.

Should we then accept the further claim that explanation by reasons makes action intelligible by revealing it to be at least approximately rational? The reasons for which we act may not be reasons for acting in that way, not only because they can be false, but because it can be a failure of practical reason to be moved by them at all. Still, the suggestion runs, to explain what someone is doing by giving their reasons is to bring out the sense in their behavior by showing how it resembles or comes close to being an exercise of practical rationality, understood not as the mere capacity to act for reasons—the claim is not trivial—but as the excellence of that capacity. In acting for reasons, one's practical reasoning or practical thought must be approximately good.

Talk of "approximation" is unhelpfully vague, but even so, it is possible to frame an argument against the present account of reasons-explanation. The difficulties come out in the common understanding of acting from a trait of character as a matter of acting for distinctive sorts of reasons. As Rosalind Hursthouse (1999: 128) notes, the courageous person is moved by such thoughts as "I could probably save him if I climb up there," "No-one else will volunteer," "If we give in now, it will be hard to stand firm later"; the temperate person is moved by such thoughts as "I'm driving," "You need it more than I do," "The cheaper one will do the job"; and so on.<sup>31</sup> The question is: In explaining one's action in terms like these, how does one show it to be approximately rational?

The answer is not, or not always, by relating it as causal or constitutive means to an end supplied by a further desire, and thus to instrumental reason.<sup>32</sup> Take, for instance, generosity. Someone who is generous may be acting in character when she helps a stranger because he needs help. And then it will be correct to say that she *wants* to help him, and to explain the particular things she does for him by citing that desire. What need not fit the instrumental pattern is her account of why she wants to help the stranger in the first place, and thus why she is doing so. She is helping him because he needs help, of course, but how does her belief that he needs help present the act of helping him as the means by which to satisfy a prior desire? It would do so if she had a completely general desire to bring aid to those in need. But that is not what generosity implies. A generous person need not be in the business

of helping just anyone who needs help or want to do so in every case. Suppose she comes upon a thief who needs help making off with stolen goods?<sup>33</sup> Even if we are wary of the view that virtues can be expressed only in acting well, so that one displays no generosity in providing help when it is unjust to do so, one does not show a *lack* of generosity in refusing it: "Do we not think that someone not ready to act unjustly may yet be perfect in charity, the virtue having done its whole work in prompting him to do that the acts that are permissible?" (Foot 1978: 15).<sup>34</sup> If the generous person is also just, there will be a nuanced web of conditions in which she will not want to offer aid; if she is also temperate, the web becomes more intricate; if she is honest, loyal, brave, even more so. Acting from generosity need not be acting from an unqualified or unconditional desire to help.

This does not by itself prevent us from treating the potential complexity of generous motivation on the model of means to ends. In helping the stranger because he needs help, the generous person may be acting in a way that satisfies a highly qualified conditional desire, in light of her beliefs—a desire to help those in need *if* certain conditions are met and others are not. But since we are looking for an explanation that the generous person herself would give, her own account of what she is doing and why, this is not enough. For she need not be able to articulate the conditions under which she wants to help others and so explain her own behavior in helping the stranger as a way of satisfying a prior desire: "I want to help those in need if and only if *x*, *y*, and *z*. Those conditions were met in the present case. So I'm going to help." This is the truth in the moral-psychological doctrine of "uncodifiability" (McDowell 1979: 57–58): It is not a condition of virtue that one have the power to formulate one's practical reasoning as the deductive application of a general principle to the case at hand. It does not follow that there is no finite codification of practical reason, only that it need not figure as the explicit object of one's propositional attitudes in acting from a virtue of character. Instead, the character of a virtuous person partly consists in being disposed to act and desire for just these reasons in just these ways, and to know that she is doing so when she does, whether or not she is able to specify, in advance, how she will react to every case.

At this point, several moves could be made on behalf of the instrumental model, of which I consider three. First, if the generous person takes herself to be doing  $\phi$  on the ground that  $p$  and  $q$  and . . . , in this particular case, and so to be moved by the corresponding beliefs, won't she also accept a means-end account of her action, on which it is explained in part by the desire to  $\phi$  if  $p$  and  $q$  and . . . ? This account reveals her as approximating instrumental rationality. Reply: If being moved to  $\phi$  by the beliefs that  $p$  and  $q$  and . . . is in this way sufficient for desiring to  $\phi$  if  $p$  and  $q$  and . . . , the requirement of intelligibility as approximate rationality is trivialized. Whenever someone acts on *any* belief, they count as having a desire in light of which their motivation shows some degree of means-end rationality, and is therefore made intelligible. Approximate rationality no longer

constrains or limits the grounds on which we can act.<sup>35</sup> Second, can it be said that our generous person understands her motivation as instrumental by referring to a less determinate desire, like the desire to help those in need, other things being equal? Not necessarily. If conditions are sufficiently bad and it is rare for those in need to deserve help, justice may sharply circumscribe her desire. She wants to help sometimes, not always, not even for the most part. Nor could the proposal in question work for every virtue: What general desire would characterize the courageous person, or the temperate? It is a special feature of generosity that it can be roughly specified in terms of an end or goal. Finally, can we supplement the generous person's understanding of herself with the desire to do whatever it is generous to do and assume that she would explain her action in helping the stranger by appeal to this desire? Again, the answer is no. For she need not conceive what she is doing in just those terms. The point applies to other virtues, too: "A courageous person does not typically choose acts as being courageous, and it is a notorious truth that a modest person does not act under the title of modesty" (Williams 1985: 10).

If the last three paragraphs are right, explanations that answer the question "Why?"—"I am showing him the way home because he needs help"—do not always make action intelligible by depicting it as the means to a prior end in light of the reason supplied. They need not show its motivation as approximating instrumental rationality. How else might they bring out the practical rationality of what someone is doing? In the case of acting from virtue, the answer may seem obvious. For it is plausible that the properly generous person sees what she is doing as the thing to do, and therefore acts under the guise of the good in the sense rejected (as a requirement on acting for reasons, as such) in section 2.<sup>36</sup> Her answer to the question "Why?" presents her as conforming to the principle of acting as one thinks one should. But this merely defers the search for intelligibility. We do not capture what is rational in someone's acting for a given reason merely by noting that she believes it to be a reason. The following dialogue is futile:

"She is drinking coffee because she loves Sophocles."<sup>37</sup>

"What? That makes no sense at all."

"Oh yes it does! She thinks it is a *reason* to drink coffee."

That she sees this consideration as a reason needs to be made intelligible, as approximately rational, no less than her being moved by it. In any case, we have already seen that it is not a general condition on acting for a reason that one regard it as a reason for what one is doing, let alone that one regard it as sufficient to establish that action as the thing to do. Even if the perfectly virtuous conceive their actions under the guise of sufficient reason, the partly or imperfectly virtuous need not; but they act for reasons nonetheless. On the view we are discussing, the explanations that give those reasons, which are in substance the same as those of the perfectly



virtuous person, must reveal approximate rationality even when they cannot be assimilated to the instrumental pattern and do not invoke beliefs about what there is reason to do.

There is, I think, only one way to vindicate this demand. If the reasons-explanations of the imperfectly generous person make her out to be approximately rational, even when she cannot cite a corresponding background desire or the belief that her reasons are good, they must do so "directly": simply because the disposition to be moved by these considerations, in the circumstance she takes herself to be in, is or approximates to being an expression of practical rationality. When we tried to assimilate such explanations to the instrumental pattern, we treated them as shorthand for more complete accounts of action that cite desires along with facts about what would satisfy them. These more complete accounts reveal approximate rationality because means-end efficiency is, or approximates to being, practically rational. The proposal at hand is that we can omit the reference to desire and take the relevant explanations to show approximate rationality because the disposition to be moved by the considerations they cite is, or approximates to being, a good disposition of practical thought. Practical rationality is at least partly constituted by dispositions that resemble those of the imperfectly virtuous person and in terms of which she explains what she is doing.

This conclusion is close to one McDowell explicitly accepts:

To explain an action we regard as virtuous, we typically formulate a more or less complex characterization of the action's circumstances as we take the agent to have conceived them. Why should it not be the case [ . . . ] that the agent's conception of the situation, properly understood, suffices to show us the favourable light in which his action appeared to him? (McDowell 1978: 80)

I have argued that this *must* suffice, all on its own, if reasons-explanation is to make action intelligible as approximately rational. It does not follow that, as McDowell also claims, we must "[take] a special view of the virtuous person's conception of the circumstances, according to which it cannot be shared by someone who sees no reason to act as the virtuous person does" (McDowell 1978: 80).<sup>38</sup> The sense in which his conception casts a positive light on what he is doing is not that it is impossible to accept it without concluding that there is reason to act as he does, but in depicting his motivation as more or less rational.

There are the makings of an argument here, from the possibility of explaining action by giving the sorts of reasons a generous or courageous or just person would give, and from the premise that reasons-explanations demonstrate at least approximate rationality, to conclusions about the content of practical reason that connect it with ethical virtue. The scope of that argument is unclear, and it is not my principal focus. For even if it is true that acting from a virtue of character is acting in a way that therefore counts as (approximately) rational, nothing similar can plausibly be

said about *vice*. The pursuit of interpretations on which we turn out always to be acting for something like good reasons tends to obscure the varieties of corruption and deformity to which our second natures are susceptible.

Some vices merely involve the unfettered pursuit of intelligible goals, as when one acts unjustly to benefit oneself. Here the explanation of what one is doing and why shows it to approximate to instrumental rationality. Like generosity, however, a defect of character need not work this way. It may issue in desires for particular actions so selectively and with such sensitivity to the details of the circumstance that they cannot be explained by the agent as directed at the means to an end. Nor does she have to regard her reasons as even *pro tanto* justifications for her action, given the argument of section 2. If not in displaying her conformity to the means-end pattern, then, or being offered as justification, how do the reasons of someone who acts in this way bring out her approximate rationality? Can we say, as we did with the virtues of character, that they do so "directly": by drawing on dispositions that are in their own right, or approximate to being, good dispositions of practical thought? That is hard to accept. Some defects of character are recognizable distortions of virtue, and their reasons mimic those of a decent person: "It's not my responsibility"; "I'll enjoy it more than he will"; "But they deserved it." Here, despite its obscurity, the claim of approximation gets some grip. In giving such reasons, one shows oneself to be in touch with the sorts of considerations that do provide reasons, if not in just this case. One's motivation can be seen as the flawed or imperfect exercise of a capacity to get things right. But other vices are more severe. As well as bitterness and spite ("It will ruin things for me" as a reason for doing it), there are pessimism and despair ("It's hard to achieve much in this world" as a reason for not even trying), bigotry and prejudice ("He's not one of us" or "It's a job for a woman" as reasons for disdain). The disposition to be moved by these considerations, in the situation one takes oneself to occupy, is not well conceived as a good disposition of practical thought in its own right, or even as resembling one. These are not the sorts of considerations that ordinarily help to justify action, and that merely fail to do so because the conditions are wrong. As reasons for acting, they are not just bad; they are awful. They do not even come close.

We can make this vivid by imagining an all-too-credible scenario. A certain community is viciously xenophobic, although their hatred of foreigners interacts with other putative virtues in complicated ways. They do not simply desire to harm others, but react to them with a nuanced array of violence, indifference, and contempt, depending on the circumstance of interaction. Perhaps the reflective members of the community have a story to tell about the value of all this. It helps to preserve their distinctive traditions and way of life. They see their xenophobic practice under the guise of the good, at least to some degree. The unreflective, however, do not. What they have is little more than an acquired tendency to act in

certain ways, on certain grounds. "He's not one of us," they say to themselves as they refuse to help or actively hurt an apparent outsider. The ways in which they do this are too complex for them to articulate for themselves as means to the satisfaction of a multiply conditional desire. Nor do they believe that their reasons for acting justify what they are doing. (The argument of section 2 ensures this possibility.) "Who cares whether it is right or wrong?" they ask. "This is what we do." That human nature is malleable enough to permit such corruption is surely plausible. What I need is something less: That deformations of this kind are consistent with rational agency. There is nothing in the nature of reasons to prevent the xenophobes from harming a stranger on grounds like these. When they do so, they can explain their action ("because he is not one of us") without reference to a further desire, to justifying reasons, or to dispositions that resemble those of good practical thought. Nor do things change when we drop the stage setting of reflective participants. That helps to give the story life, but it is not essential. Let the whole community be unreflective, at least in their prejudice, passed on by the contagion of habit. In this department, at least, they do not act under the guise of approximate rationality or the approximate good.

The argument of the last two paragraphs relies on claims about what is and is not a reason for acting, and about the sorts of dispositions that help to make up practical rationality. It would not be persuasive to someone who believes that xenophobia is or approximates to being good practical thought, even when it is thoroughly unreflective. But the principal claim is quite abstract: To hold that our account of what we are doing in acting for reasons must show it to be approximately rational, even when it does not conform to the instrumental pattern or involve the belief that we are acting for good reasons, is to lose the contrast between incomplete or imperfect possession of practical reason and habituation into forms of practical responsiveness that are simply misguided or wrong. This echoes the contrast, which belongs to common sense, between the failure to be virtuous and possession of a positive vice. Second nature can be shaped to incorporate dispositions that are not mere perversions of rationality, but actively depraved: basic tendencies to act and desire on irrelevant or unjustified grounds. In exercising dispositions of that kind, the vicious person need not conceive herself as acting for good reasons, or as aiming at the satisfaction of a general desire; she may explain what she is doing as the exercise of an acquired disposition that does not even approximate to practical rationality.

The demand for such approximation is sometimes expressed by saying that reasons-explanations "rationalize" action or desire, adding immediately that the term "rationalization" is being used in a technical not a colloquial sense. The contrast between imperfection and positive vice suggests that this is wrong on both counts. The doctrine of approximate rationality portrays the depraved and ill habituated as compulsive rationalizers, excusing their bad behavior with the semblance of good reasons. The truth

is less comforting. One can act for reasons that are wholly and irredeemably bad, and thus with knowledge of what one is doing and why that does not present it under the guise of the approximately good. Even in its modest form, which rests on the alleged normativity of reasons-explanation and its role in answering the question "Why?" the guise of the good for intentional action is false.

#### 4. HUMAN NATURE

It is an assumption of the preceding arguments that the guise of the good takes the form of, or entails, a universal generalization:

Any possible instance of acting for a reason is an instance of acting under the guise of the good.

The task of section 2 was to show that it is not a condition of acting on the ground that  $p$  that one take it to be a reason for what one is doing or to indicate some respect in which one's action is a good one to perform. The task of section 3 was to show that, even if it is a condition of acting for a reason that one be able to supply an explanation of what one is doing that gives one's reason, this explanation may fail to cast one's action in the positive light of approximate rationality. It is possible to act for a reason in doing  $\phi$  without acting under the guise of the good.

This way of putting things prompts a final objection, which turns on the logical weakness of possibility claims. For there are forms of generality that permit exceptions, even as they seem to describe the essence or nature of what they generalize about. If the doctrine that we act under the guise of the good is intended as a nonuniversal generalization, the arguments above may seem to miss the point. Does the present discussion go wrong by presupposing an unduly simple view of the generality involved in the relevant essentialist claims?

Our question is inspired by a revived Aristotelianism about living things that finds its fullest expression in Michael Thompson's essay "The Representation of Life."<sup>39</sup> He draws attention to the way in which we state how a certain species of living things goes on:

Let us call the thoughts expressed in the field guide and in the nature documentary *natural-historical judgments*. We may take as their canonical expression sentences of the form 'The  $S$  is (or has, or does)  $F$ '—'The domestic cat has four legs, two eyes, two ears, and guts in its belly,' 'The Texas blue-bonnet harbors nitrogen-fixing microbes in certain nodes on its roots,' 'The yellow finch breeds in spring, attracting its mate with such and such song,' whatever. Such sentences I will call 'Aristotelian categoricals.' But our language of course permits the same judgments to be expressed in a number of other ways, for example, by ' $S$ 's are/have/do  $F$ ,' or 'It belongs to an  $S$  to be/have/do  $F$ ,' or ' $S$ 's characteristically (or typically) are/have/do  $F$ ,' and a hundred others. (Thompson 1995: 281)

What is involved here is a form of generality that is expressed by what linguists call “generic” sentences. Not all generics purport to capture the nature of a kind or species, but some do, and their doing so is consistent with their failing short of universal generality. That wolves hunt in packs is some sort of insight into what they are, even though this one or that one may go it alone.

No doubt more needs to be said about the metaphysics of this generic essentialism, about the prospects for its reductive treatment—which Thompson (1995: 284–88) contests—and about the scope of its application. Our interest is confined to its interaction with the guise of the good. Nothing in the argument so far conflicts with the truth of natural-historical judgments, or nature-expressing generics, that connect human action and desire with appearances of the good:

Human beings want what seems good to them; and they act for reasons they regard as good.

If generic essentialism makes sense, these propositions could be necessary truths of human nature even though some of us want to own saucers of mud and others act from spite or vanity or despair. It has been proposed, if only in passing, that this is the intended form of the guise of the good for desire. It is “an assertion about the ‘essence’ of desire, rather like the assertion that fish are vertebrates: either is consistent with the occurrence of freaks” (Stampe 1987: 366). Michael Stocker (2004: 319) considers, without endorsing, the related claim that it is “natural for any being to seek its good.” For all I have said, then, it may be a natural-historical fact about us, a necessary truth of human nature, that we act under the guise of the good.

What I have argued against is the rather different view that it belongs to rational agency, as such, to be exercised under the guise of the good. As at the end of section 2, the point can be stated ecumenically. There are two different ways to understand the doctrine that we act under the guise of the good: as a fragment of the natural history of human beings, limited to our particular way of acting for reasons and consistent with exceptions; and as a partial account of the abstract capacity to act for reasons. The first claim is perhaps defensible. At any rate, I don’t object to it. That human beings act under the guise of the good would help to justify the sense of aberration, though not impossibility, in cases where we don’t. What I have argued is that the guise of the good is not contained in the bare idea of intentional action or acting for a reason. To claim otherwise is to fall into misconceptions in action theory: of the kind of proposition that answers the question “Why?”—which gives an explanatory not a justifying reason—or of the nature and normativity of the explanation involved. Rejecting these misconceptions is not as radical as it may seem, since we can do so while accepting the analogous generic claims about humanity: We tend to want what we perceive as being in some way good, to choose what seems worthy of

choice, and to act in ways we think we can justify, at least to some extent.

This is a substantive concession, but a limited one. It allows for generic essentialism in application to human nature, but not the nature of rational agency, as such. It thus invites a further question. Why can’t we propose, in the mode of nonuniversal but nature-expressing generality, that rational agents, those capable of acting for reasons, act and desire for reasons they perceive as good? That would make the arguments of sections 2 and 3 irrelevant, directed at the wrong sort of proposition even about their abstract topic. What prevents this is a restriction on the kinds of things whose nature or essence is properly captured in generic terms. “Those capable of acting for reasons” do not form an appropriate target for the sort of generalization that permits exceptions even as it tells us what its subjects are. We say that the cat has four legs, and the human being thirty-two teeth, but not that “the rational agent” decides what to do in this way or that—unless we mean to imply that every one of them does so, of necessity, or to be reporting a merely accidental fact.

This seems evident to me, but it can also be supported by argument. When *F*s are by nature *G*, in the generic sense, but their being *G* depends on the circumstance in which they find themselves, there is a distinction to be drawn among such circumstances, between those that do and those that do not fit with the nature of the *F*. The conditions in which an *F* is not *G* are ones that prevent it from realizing its nature; they are excluded from its natural circumstance, which is itself to be specified with nonuniversal generalizations. Thompson illustrates this point with an example:

Now suppose I say, ‘Bobcats breed in spring’: it is again obvious that this isn’t going to happen in any particular case unless certain conditions are satisfied. Perhaps a special hormone must be released in late winter. And perhaps the hormone will not be released if the bobcat is too close to sea-level, or if it fails to pass through the shade of a certain sort of tall pine. But, now, to articulate *these* conditions is to advance one’s teaching about bobcats. It is not a reflection on the limited significance of one’s teaching. The thought that *certain hormones are released*, or that *they live at such and such altitudes and amid such and such vegetation*, is a thought of the same kind as the thought that *they breed in spring*. The field guide and the nature documentary assign an external environment to the intended life-form, after all, and in the same ‘voice’ they elsewhere employ in describing its bearers’ inner structure and operations. These conditions are thus ‘presupposed’ by the life-form itself, and not by the poor observing subject with his low-resolution lens. (Thompson 1995: 287)

All of this applies to the generalizations about human nature above, assuming that they are true. Human beings want what seems good to them; and they act for reasons they perceive as good. Still, some do not, as perhaps in conditions of bad upbringing or severe deprivation. But then it belongs to human nature not to be brought up in those ways or deprived of those things. Our natural environments are ones that foster the tendency to act and desire under the guise of the good. Or if they are not, the

guise of the good is not an essential truth about human nature; it is, at best, a contingent fact about humans hereabouts. The central point is this: When *F*s are by nature *G*, but it is possible for an *F* not to be *G*, there are further truths about the nature of the *F* that describe its natural circumstance, and this circumstance excludes the conditions that prevent an *F* from being *G*. Thus, if rational agents by nature act under the guise of the good, but some do not, it must be natural for them to inhabit conditions in which they are not prevented from coming to act under the guise of the good—as we may be prevented by corrupt habituation. But this is nonsense. There is no such thing as the natural environment of a rational agent, abstractly conceived, only for particular kinds of living thing. It follows that we cannot capture the essence of rational agency in generic terms. If it belongs to rational agents, as such, to act under the guise of the good, there cannot be exceptions.

Where does this leave our guiding question, about agency and evaluation? I have argued for the following principal claims:

If desires represent their objects as good, they must be had for reasons, and reasons for desire must be respects in which the object of desire is seen as good.

One can act and desire for reasons without regarding them as reasons for what one is doing, or as respects in which it is a good thing to do; thus one can act for reasons without regarding one's action as *pro tanto* justified.

Explanations of what one is doing in terms of reasons, of the kind one must accept in answering the question "Why?" need not reveal even approximate rationality.

It follows that the guise of the good does not apply to intentional action, as such; nor can it be rescued by generic essentialism.

It is consistent with these claims that the tendency to act and desire for reasons we regard as good is a natural-historical necessity of human life. The source of the guise of the good in action theory may be a familiar and tempting parochialism: the mistake of thinking that our characteristic form of agency shows us what agency essentially is.

There is a further moral to be drawn, in closing, from our brief examination of generic essentialism. When the nature of a kind can be specified in generic terms, as seems possible with species of living things, it is tempting to regard this specification as normative: It defines the good or healthy or well-functioning individual. As Thompson remarks:

We may implicitly define a certain very abstract category of 'natural defect' with the following simple-minded principle of inference: *From* "The *S* is *F* and 'This *S* is not *F*' to infer "This *S* is defective in that it is not *F*." (Thompson 1995: 295)

This is probably *too* abstract; it needs refinement.<sup>40</sup> But it is enough to encourage the hope that, at least sometimes and to some extent, generic essentialism provides a model for the derivation of norms from natures. One form of ethical rationalism applies this hope to the standards of practical

reason, which it purports to derive from the nature of rational agency, generically described.<sup>41</sup> If the argument above is right, however, this strategy cannot succeed. The nature of rational agency is not the sort of thing that can be stated in generic terms.

It does not follow from this alone that ethical rationalism is false. After all, it may take a different form, resting on essentialist claims that imply universal generalizations. But even here, our arguments support a provisional skepticism.<sup>42</sup> The ethical rationalist cannot rely on the doctrine that agency is exercised under the guise of the good or that it must conform to approximate rationality. The second restriction, in particular, is severe. If the form of explanation characteristic of rational agency does not invoke or draw upon the standards of practical reason, how could those standards be implicitly contained within it?

## Notes

1. For helpful discussion of this material, I am grateful to Robert Audi, Rachel Barney, Matt Boyle, Doug Lavin, Jessica Moss, Evgenia Mylonaki, Joseph Raz, Sebastian Rödl, Geoff Sayre-McCord, Michael Smith, Sergio Tenenbaum, Jennifer Whiting, to participants at conferences held in Toronto and Syracuse in the summer of 2007, and to Michael Smith's Seminar in Systematic Ethics at Princeton, fall 2008.

2. Setiya 2007, part 1.

3. A classic source is Stocker 1979. See also Stocker 2004: 324; Frankfurt 2004: 122ff.; Frankfurt's replies in *Contours of Agency*, ed. Buss and Overton, 2002: 87, 89, 160, 187, 223; and Setiya 2007: 36–38.

4. As, for instance, in Tenenbaum 2007.

5. The accusation is perhaps unfair to Stocker. His essay "Desiring the Bad" ends with a more general account of the connection between evaluations and motives on which "the controverting cases are not exceptions, aberrations, mere anomalies or mere counterexamples"; even in the positive case, where we do act in ways that we regard as good, or for reasons we take to justify what we are doing, the "connections between motivation and evaluation are mediated by arrays of structures of mood, interest, energy and the like" (Stocker 1979: 750–51).

When I consider people who see no hope for themselves or those they care for, who lack physical and spiritual energy, I am not at all surprised that—as political and anthropological data suggest—they may not seek even what little good they do perceive. Life may be too much for them. We, on the contrary, see the world as open to us, and more important, open for us. We can progress. We can make it. We see ourselves out there to be won. We have self-confidence and hope. Indeed, we have more than this: We have an optimistic certainty. We have energy. We know we are worthy. We know that, barring bad luck, our enterprise will be rewarded. And so on. Such an array of structures of mood, interest, energy, . . . makes it natural, almost inevitable, that we seek the (believed) good for ourselves or others. And it seems at least arguable that such an array must be posited to give an adequate account of how, at least according to our cultural ideal, motivation and evaluation are related in us. (Stocker 1979: 752)

6. Compare Anscombe 1963: 75; and Velleman 1992a: 118–19.
7. See Anscombe 1958: 38; and, more extensively, Moravcsik 1994, Thompson 1995, and Foot 2001.
8. Along with the authors quoted in the text, see Darwall 1983: 205; Bond 1983: 30–31; Velleman 1992b: 140–42; Korsgaard 1997: 221; Broome 1997; Wallace 1999; Dancy 2000: 97; Moran 2001: 124–28; and many others.
9. The translation is by G. A. Grube and C. D. C. Reeve (Cooper 1997: 1126).
10. Quoted from Hamlyn 1968: 69–70.
11. Kant 1788, Ak. 5:58–59. Kant's attitude to the "old formula of the schools" is, however, complicated. In saying that he cites it with approval, I do not mean that he would accept it in just its original sense.
12. For such accounts, see Thomson 1992: 107–13; and Scanlon 1998: 95–100.
13. For this point, see Johnston 2001.
14. What is it for something to be good "for its own sake"? The basic notion here is "for the sake of":  $x$  is valuable for the sake of  $y$  just in case the value of  $x$  is partly explained by its relation to the value of  $y$ . Perhaps  $x$  is a means to  $y$ , or approximates  $y$ , or bears some symbolic connection with  $y$ . For  $x$  to be valuable for its own sake is for it to have value that is not wholly explained by its relation to the value of other things. This is to be distinguished from having value as an end, not wholly as a means to other valuable things, which is a logically weaker property, and from being intrinsically valuable or valuable in virtue of one's intrinsic nature. If the value of  $x$  is explained by its relation to  $y$  though not through the value of  $y$ , then  $x$  is valuable for its own sake, but not intrinsically. And while value for the sake of something else is typically not intrinsic, it can be. Suppose, for instance, that  $x$  is valuable for the sake of one of its parts.
15. This conception of desire is found in Raz 1999b: 52–56.
16. The argument so far neither supports nor refutes the idea that desires represent their objects as good. It does suggest, however, that desires are not "appearances" of the good, except perhaps in the modest sense that would distinguish them from *beliefs*, so as to allow for illusions of value that we do not accept. In particular, we should be wary of the claim that desires are perceptual states (as in Stampe 1987). The fact that desires are typically had for reasons speaks against this. Although there are reasons why we perceive things as we do, and why they appear that way to us, we don't have reasons for perceiving that things are thus-and-so or for being appeared to as we are. Perceiving is not something we do for reasons, as we act and desire for reasons. Desires are in this respect quite unlike perceptual states.
17. I give a causal-psychological account of acting for a reason in Setiya 2007: 28–59.
18. This is apparently denied by Raz (1999a: 23), when he writes that "intentional action is action done for a reason; and [ . . . ] reasons are facts in virtue of which those actions are good in some respect and to some degree," and by Dancy (2000: 9): "to explain an action is [ . . . ] to show that it would have been [what there was most reason to do] if the agent's beliefs had been true." These claims are surely too strong, just as they stand: They are best read as exaggerations of the theory criticized in section 3.

19. The argument in the following text is a version of one I have proposed elsewhere—in part I of *Reasons without Rationalism*—stripped of its constructive ambitions. The objection to the guise of the good does not depend on giving a causal-psychological account of action, and the attempt to do so would be a distraction here.
20. See, for instance, Anscombe 1963: 11, 42–44, 48.
21. For similar claims about knowledge of reasons, see Milligan 1974: 187–88; Audi 1986: 82–85; Wallace 1999: 241; and Searle 2001: 16.
22. "A man may be making ten carbon copies as he writes, and this may be intentional; yet he may not know that he is; all he knows is that he is trying" (Davidson 1971: 50; see also Davidson 1978: 91–94). For further discussion, see *Reasons without Rationalism* (Setiya 2007: 25–26) and section I of "Practical Knowledge" (Setiya 2008).
23. See also Stocker 2004: 326–29 on "conditional or circumstantial goodness."
24. In making this argument, I ignore the suggestion that "practical knowledge"—the fact that we typically *know* and do not merely *believe* that we are doing  $\phi$  in doing it for reasons—is explained by the guise of the good. On the kind of account proposed by Wilson (2000: 12–16) and Moran (2001: 124–28), knowledge of what one is doing intentionally derives from practical judgment, a verdict about the reasons that bear on what to do. But this is the problematic, in two ways. First, what it explains is at most how the agent is "in a position to know" what she is doing, if she takes her action to be determined by the balance of reasons, not why she must believe that she is doing it. Second, accounts of this kind struggle to accommodate knowledge of intentional action against one's evaluative beliefs, or when one has sufficient reason for doing more than one thing. In "Practical Knowledge" (Setiya 2008), I give an account of knowing what one is doing that avoids these difficulties.
25. As, for instance, in Johnston 2001.
26. A difficult question can be raised about the scope of this result. Does it apply to reasons for belief, where we can equally distinguish propositions about justification and evidence from propositions about the grounds on which a given belief is held? That I believe that  $p$  on the ground that  $q$  is consistent with there being no evidence at all for the truth of my belief. Is it sufficient to answer the question, "Why do you believe that  $p$ ?" that I can cite a proposition of the former kind? In my view, the answer is no: Reasons attach to actions and beliefs in quite different ways. For instance, it is sufficient for believing that  $p$  partly on the ground that  $q$  that one believe that  $p$  and believe that the fact that  $q$  is evidence that  $p$ . There is no need for a causal relation between these beliefs, and no analogue of Davidson's argument in "Actions, Reasons and Causes" (1963: 9). Believing for a reason is independent of causality in a way that acting for a reason is not. This claim is no doubt controversial, and the issues are too complicated to examine here. I hope to address them more adequately in future work.
27. I try to articulate this connection precisely, through the idea of good practical thought, in the introduction to *Reasons without Rationalism* (Setiya 2007).
28. For a similar claim about reason and volition, see Frankfurt 2004: 120–21, esp. n. 2.

29. Again, see *Reasons without Rationalism* (Setiya 2007: 25–26) and section I of “Practical Knowledge” (Setiya 2008).
30. For related claims about the normative character of reasons-explanation, see Nagel 1970: 33–34; Korsgaard 1997: 221–22; Raz 1999a: 22–24; Dancy 2000: 9–10, 95–97, 106; Wedgwood 2006: 662; and Tenenbaum 2007: 9–17.
31. See also Williams 1995; Foot 2001: 12; and Setiya 2007: 71–74.
32. In the present context, we need not distinguish different versions of instrumental reason or the instrumental principle, some of which appeal to our intentions, others to mere desires. The argument relies on the general concept of being a means to an end.
33. The example is adapted from Herman 1981: 364–65.
34. For the claim that “a genuine virtue [must] produce nothing but right conduct,” see McDowell 1979: 52–53. This doctrine tends to support the unity of the virtues, though not without some further argument. The issues here are critically discussed in Watson 1984: 59–62, 67–69.
35. For related discussion, see *Reasons without Rationalism* (Setiya 2007: 101–6).
36. This seems to be accepted by McDowell (1978: 90; 1979: 51) in contexts where he rejects the assimilation of acting from virtue to acting from desire.
37. For this example, see Raz 1997: 8.
38. Unless, trivially, the conception is taken to include the fact that there is reason to act in that way, as, for instance, at McDowell 1978: 90; and in McDowell 1979; the more restrictive understanding in the text seems to operate at McDowell 1978: 87.
39. He is drawing on brief remarks by Anscombe (1958: 38); see also Moravcsik 1994 and Foot 2001.
40. As in Foot 2001: 29ff.
41. See Lavin 2004: 456n56 on “constitutive” accounts of practical reason:
- [We] lack a correct conception of the logical form of [...] claims describing the essence or nature of agency, the claims in virtue of which we are supposed to understand the force of “oughts” applying to particular agents. [...] A promising direction for the constitutivist to go, I think, is to resist the urge to assimilate [descriptions of essences] to universal generalizations and instead look toward generics to describe “the what it is” which is to serve to underwrite standards of assessment.
42. For more decisive resistance, see part 2 of *Reasons without Rationalism* (Setiya 2007).

## References

- Anscombe, G. E. M. 1958: Modern moral philosophy, repr. in G. E. M. Anscombe, *Ethics, Religion and Politics*. Oxford: Basil Blackwell, 1981: 26–42.
- . 1963: *Intention*, 2nd ed. Oxford: Basil Blackwell.
- Audi, R. 1986: Acting for reasons, repr. in A. Mele, ed., *The Philosophy of Action*. Oxford: Oxford University Press, 1997: 75–105.
- Bond, E. J. 1983: *Reason and Value*. Cambridge: Cambridge University Press.

- Broome, J. 1997: Reason and motivation, *Proceedings of the Aristotelian Society, Supplementary Volume* (71): 131–46.
- Buss, S., and Overton, L., eds. 2002: *Contours of Agency: Essays on Themes from Harry Frankfurt*. Cambridge, Mass.: MIT Press.
- Cooper, J. M., ed. 1997: *Plato: Complete Works*. Indianapolis: Hackett.
- Dancy, J. 2000: *Practical Reality*. Oxford: Oxford University Press.
- Darwall, S. 1983: *Impartial Reason*. Ithaca, N.Y.: Cornell University Press.
- Davidson, D. 1963: Actions, reasons and causes, repr. in D. Davidson, *Essays on Actions and Events*. Oxford: Oxford University Press, 1980: 3–19.
- . 1971: Agency, repr. in D. Davidson, *Essays on Actions and Events*. Oxford: Oxford University Press, 1980: 43–61.
- . 1978: Intending, repr. in D. Davidson, *Essays on Actions and Events*. Oxford: Oxford University Press, 1980: 83–102.
- . 1987: Problems in the explanation of action, repr. in D. Davidson, *Problems of Rationality*. Oxford: Oxford University Press, 2004: 101–16.
- Foot, P. 1978: Virtues and vices, in P. Foot, *Virtues and Vices*. Oxford: Basil Blackwell, 1978: 1–18.
- . 2001: *Natural Goodness*. Oxford: Oxford University Press.
- Frankfurt, H. 2004: Disengaging reason, in R. J. Wallace, P. Pettit, S. Scheffler, M. Smith, eds., *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*. Oxford: Oxford University Press, 2004: 119–28.
- Hamlyn, D. W. 1968: *Aristotle: De Anima, Books II and III*. Oxford: Oxford University Press.
- Herman, B. 1981: On the value of acting from the motive of duty, *Philosophical Review* 90: 359–82.
- Hursthouse, R. 1991: Arational actions, *Journal of Philosophy* 91: 57–68.
- . 1999: *On Virtue Ethics*. Oxford: Oxford University Press.
- Johnston, M. 2001: The authority of affect, *Philosophy and Phenomenological Research* 63: 181–214.
- Kant, I. 1788: *Critique of Practical Reason*, M. Gregor, trans., Cambridge: Cambridge University Press, 1997.
- Korsgaard, C. 1997: The normativity of instrumental reason, in G. Cullity and B. Gaut, eds., *Ethics and Practical Reason*. Oxford: Oxford University Press, 1997: 215–54.
- Lavin, D. 2004: Practical reason and the possibility of error. *Ethics* 114: 424–57.
- McDowell, J. 1978: Are moral requirements hypothetical imperatives? repr. in J. McDowell, *Mind, Value and Reality*. Cambridge, Mass.: Harvard University Press, 1998: 77–94.
- . 1979: Virtue and reason, repr. in J. McDowell, *Mind, Value and Reality*. Cambridge, Mass.: Harvard University Press, 1998: 50–73.
- . 1985: Functionalism and anomalous monism, repr. in J. McDowell, *Mind, Value and Reality*. Cambridge, Mass.: Harvard University Press, 1998: 325–40.
- Milligan, D. E. 1974: Reasons as explanation, *Mind* 83: 180–93.
- Moran, R. 2001: *Authority and Estrangement*. Princeton, N.J.: Princeton University Press.
- Moravcsik, J. 1994: Essences, powers and generic propositions, in T. Scaltsas, D. Charles, and M. L. Gill, eds., *Unity, Identity and Explanation in Aristotle's Metaphysics*. Oxford: Oxford University Press, 1994: 229–44.
- Nagel, T. 1970: *The Possibility of Altruism*. Princeton, N.J.: Princeton University Press.

- Raz, J. 1997: When we are ourselves: The active and the passive, repr. in J. Raz, *Engaging Reason*. Oxford: Oxford University Press, 1999: 5–21.
- . 1999a: Agency, reason and the good, in J. Raz, *Engaging Reason*. Oxford: Oxford University Press, 1999: 22–45.
- . 1999b: Incommensurability and agency, in J. Raz, *Engaging Reason*. Oxford: Oxford University Press, 1999: 46–66.
- Scanlon, T. M. 1998: *What We Owe to Each Other*. Cambridge, Mass.: Harvard University Press.
- Searle, J. 2001: *Rationality in Action*. Cambridge, Mass.: MIT Press.
- Setiya, K. 2007: *Reasons without Rationalism*. Princeton, N.J.: Princeton University Press.
- . 2008: Practical knowledge. *Ethics* 118: 388–409.
- Smith, M. 1987: The Humean theory of motivation. *Mind* 96: 36–61.
- . 1994: *The Moral Problem*. Cambridge: Cambridge University Press.
- Stampe, D. 1987: The authority of desire. *Philosophical Review* (96): 335–81.
- Stocker, M. 1979: Desiring the Bad. *Journal of Philosophy* 76: 738–53.
- . 2004: Raz on the intelligibility of bad acts, in R. J. Wallace, P. Pettit, S. Scheffler, M. Smith, eds., *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*. Oxford: Oxford University Press, 2004: 303–32.
- Tenenbaum, S. 2007: *Appearances of the Good: An Essay on the Nature of Practical Reason*. Cambridge: Cambridge University Press.
- Thompson, M. 1995: The representation of life, in R. Hursthouse, G. Lawrence, and W. Quinn, eds., *Virtues and Reasons: Philippa Foot and Moral Theory*. Oxford: Oxford University Press, 1995: 247–96.
- Thomson, J. J. 1992: On some ways in which a thing can be good. *Social Philosophy and Policy* 9, no. 2: 96–117.
- Velleman, J. D. 1992a: The guise of the good, repr. in J. D. Velleman, *The Possibility of Practical Reason*. Oxford: Oxford University Press, 2000: 99–122.
- . 1992b: What happens when someone acts? repr. in J. D. Velleman, *The Possibility of Practical Reason*. Oxford: Oxford University Press, 2000: 123–43.
- Wallace, R. J. 1999: Three conceptions of rational agency. *Ethical Theory and Moral Practice* 2: 217–42.
- Watson, G. 1984: Virtues in excess. *Philosophical Studies* 46: 57–74.
- Wedgwood, R. 2006: The normative force of reasoning. *Noûs* 40: 660–86.
- Williams, B. 1979: Internal and external reasons, repr. in B. Williams, *Moral Luck*. Cambridge: Cambridge University Press, 1981: 101–13.
- . 1985: *Ethics and the Limits of Philosophy*. Cambridge, Mass.: Harvard University Press.
- . 1995: Acting as the virtuous person acts, in R. Heinaman, ed., *Aristotle and Moral Realism*. London: University College London Press, 1995: 13–23.
- Wilson, G. 2000: Proximal practical foresight. *Philosophical Studies* 99: 3–19.