

Moral Motivation

Author(s): David O. Brink

Source: *Ethics*, Oct., 1997, Vol. 108, No. 1 (Oct., 1997), pp. 4-32

Published by: The University of Chicago Press

Stable URL: <https://www.jstor.org/stable/2382087>

## REFERENCES

Linked references are available on JSTOR for this article:

[https://www.jstor.org/stable/2382087?seq=1&cid=pdf-reference#references\\_tab\\_contents](https://www.jstor.org/stable/2382087?seq=1&cid=pdf-reference#references_tab_contents)

You may need to log in to JSTOR to access the linked references.

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



The University of Chicago Press is collaborating with JSTOR to digitize, preserve and extend access to *Ethics*

JSTOR

## SYMPOSIUM ON MICHAEL SMITH'S *THE MORAL PROBLEM*

### Moral Motivation\*

*David O. Brink*

One of the principal objectives of Michael Smith's clearheaded, vigorously argued, and rewarding book, *The Moral Problem*, is the reconciliation of morality's *intellectual* and *practical* dimensions. This reconciliation is no easy matter. The problem that occupies most of Smith's attention concerns moral motivation. Moral judgments are typically motivationally efficacious. If we think that motivation involves pro-attitudes, such as desires, we may conclude from the motivational or "dynamic" aspects of morality that moral judgments express noncognitive attitudes, rather than beliefs. But this noncognitive conclusion may seem to miss intellectual aspects of morality, which cognitivism captures. To avoid it, it may seem that we need to reject the idea that moral judgment has some internal connection with motivation. But this may seem to abandon the practical dimension of morality. We could understand moral motivation in some new way that does not involve pro-attitudes, but this may seem hard to square with familiar assumptions about the nature of intentional action.

Smith discusses the structure of this problem and usefully examines a number of traditional and contemporary views in ethical theory as responses, express or tacit, to this problem. In the wake of criticisms of alternative solutions, he develops his own solution, which preserves a cognitivist (and broadly naturalistic) interpretation of the intellectual aspects of moral inquiry and defends an internalist interpretation of the practical aspects of morality by treating it as a conceptual truth

\* This is a discussion of Michael Smith, *The Moral Problem* (Oxford: Blackwell, 1994); it is a revised version of my contribution to a symposium on Smith's book held at the 1996 American Philosophical Association (APA) Pacific Division meetings. Unless otherwise noted, parenthetical references are to pages in this book. Thanks to Michael Smith and Steve Yalowitz for comments on an earlier draft. A substantial portion of this material descends from a larger work in progress, tentatively titled "Objectivity, Motivation, and Authority in Ethics."

that moral requirements are requirements of practical reason. On this view, it is a conceptual truth that to think that morality requires a course of conduct is to think that that action is what one would desire to do if one was fully rational. If so, moral judgments will be motivationally efficacious just insofar as we are practically rational. In this way, Smith claims, we can reconcile intellectual and practical aspects of morality without resorting to extravagant or ad hoc assumptions about human motivation.

Though there is much to admire in Smith's discussion, I have some reservations. Some of my reservations are metaphilosophical. I am pretty sure we disagree about the extent to which it is useful to think of moral, metaethical, and (more generally) philosophical theorizing as conceptual analysis. He is an advocate of this perspective; I am not. I suspect that these metaphilosophical differences influence our rather different views about the most plausible form of ethical naturalism: I favor a metaphysical naturalism, whereas Smith rejects metaphysical naturalism in favor of a "network analysis" of moral terms. However, I would like to focus on Smith's claims about moral motivation, touching on our metaphilosophical differences only where these seem directly relevant to issues about moral motivation. As I will explain, I favor an externalist interpretation of morality's practical dimensions and am not convinced by Smith's reservations about externalism. Nonetheless, we are in substantial agreement about a number of other issues about moral motivation; concerning moral motivation, there is probably more about which we agree than disagree. I found it useful to try to understand and assess his account of the moral problem and its solution, because we've been thinking about moral motivation and related issues in ethical theory in similar or related ways. As a result, it will be useful, and certainly easier for me, to understand and assess Smith's discussion in terms of my own thinking about many of the same issues.

## I. MORAL MOTIVATION

One apparent tension between intellectual and practical aspects of morality involves debates between cognitivists and noncognitivists. Cognitivists interpret moral judgments as expressing cognitive attitudes, such as belief, rather than noncognitive attitudes, such as desire. Internalists believe that moral judgments necessarily engage the will and motivate. It is a common view that motivation involves pro-attitudes, such as desires, and that no belief entails any particular desire. But these assumptions are in tension, as we can see if we try to see them as forming a puzzle about moral motivation.<sup>1</sup>

1. My formulation of the motivational puzzle and Smith's are similar to that of David McNaughton, *Moral Vision* (Oxford: Blackwell, 1988), p. 23.

1. Moral judgments express beliefs.
2. Moral judgments entail motivation.
3. Motivation involves a desire or pro-attitude.
4. There is no necessary connection between any belief and any desire or pro-attitude.

Assumption (1) expresses a cognitivist view of ethics. Moral judgments appear to express the appraiser's beliefs about the moral properties of persons, actions, and institutions. Assumption (2) expresses the internalist thesis that motivation is an essential part of moral judgment.<sup>2</sup> This is not the thesis that people necessarily act according to their moral judgments; on this view, though moral judgments necessarily motivate, this motivation may be overridden by other counter-vailing motivations. Assumption (3) expresses the familiar idea that motivation involves a desire or pro-attitude on the agent's part. And assumption (4) expresses the familiar idea that beliefs and pro-attitudes, such as desires, are independent mental states such that beliefs do not require any particular pro-attitude.

Each of these claims can seem plausible; at least, each has seemed plausible to a number of people. But whatever their individual appeal, not all four claims can be true.<sup>3</sup> To avoid inconsistency, we must reject at least one element of the puzzle. Noncognitivists (e.g., A. J. Ayer, C. L. Stevenson, R. M. Hare, and Allan Gibbard) appeal to (2)–(4) and, as a result, reject (1); they conclude that moral judgments express noncognitive attitudes, rather than beliefs. Some cognitivists (e.g., Philippa Foot and I) appeal to (1), (3), and (4) and, as a result, reject (2). On this view, it is possible to make moral judgments without being motivated to act. Because it rejects the idea that motivation is internal to moral judgment, this view might be called “externalism” about moral motivation. Other cognitivists accept internalism; they must reject either (3) or (4). We might call this view “rationalism” about moral motivation. Some rationalists (e.g., Thomas Nagel and John McDowell) think that recognition of moral duties can be intrinsically motivational without the benefit of a pro-attitude. Another form of rationalism concedes that moral motivation requires pro-attitudes but insists that

2. This is the view Darwall calls judgment internalism; see Stephen Darwall, *Impartial Reason* (Ithaca, N.Y.: Cornell University Press, 1983), p. 54.

3. We can demonstrate this symbolically, as follows:

1.  $\Box (J \equiv B)$
2.  $\Box (J \rightarrow M)$
3.  $\Box (M \rightarrow D)$
4.  $\Diamond (B \ \& \ \neg D)$
5.  $\Box (J \rightarrow D)$  [2, 3]
6.  $\Diamond (J \ \& \ \neg D)$  [1, 4]
7.  $\neg \Box (J \rightarrow D)$  [6]
8.  $\Box (J \rightarrow D) \ \& \ \neg \Box (J \rightarrow D)$ . [5, 7]

normative beliefs entail a pro-attitude toward the action in question. With suitable qualifications, which I will explain later, I think that Smith's own solution is best interpreted as this second kind of rationalism.

It might be worth noticing a couple of differences between my way of formulating the puzzle about moral motivation and Smith's formulation of his moral problem. He discusses three claims (p. 12).

- 1s. Moral judgments of the form "It is right that I  $\phi$ " express a subject's beliefs about an objective matter of fact, a fact about what is right for her to do.
- 2s. If someone judges that it is right that she  $\phi$ s then, *ceteris paribus*, she is motivated to  $\phi$ .
- 3s. An agent is motivated to act in a certain way just in case she has an appropriate desire and a means-end belief, where belief and desire are, in Hume's terms, distinct existences.

One difference in our formulations is that, whereas my puzzle consists of a quartet of claims, Smith's problem consists of a trio. But this difference is trivial if, as I believe, his third claim combines the two I distinguish as (3) and (4). If, as we both believe, it is useful to see disparate views as different reactions to an underlying puzzle about motivation, each of which rejects a familiar claim in order to preserve others, then there may be some utility to seeing the puzzle as a quartet. Because the two forms of rationalism that result from denying the claims I distinguish as (3) and (4) are importantly different, it is helpful to keep these claims distinct.

A more significant difference is that Smith's second claim is weaker than mine. Whereas my version of internalism asserts that moral judgment entails motivation, his requires only that *ceteris paribus* moral judgment is accompanied by motivation. Only in this weakened form is Smith himself an internalist. Both of our versions of internalism involve defeasible commitments. My version asserts a defeasible connection between moral judgment and action: people need not act according to their moral judgments; though moral judgments necessarily motivate, this motivation may be overridden by other counter-vailing motivations. But it does assert a necessary connection between moral judgment and motivation. By contrast, Smith's version of internalism asserts a defeasible connection between moral judgment and motivation. This opens up the possibility that we are not in disagreement when he defends (weak) internalism and I reject (strong) internalism. However, as we shall see, there is an important residual disagreement, after this partly verbal disagreement is cleared away (Secs. IV–VI).

It might be worth saying why I formulate the internalism requirement as I do. Though it is unlikely that any interpretation of internalism (or, for that matter, other elements of the puzzle) will be faithful

to everything every party to debates about moral motivation has claimed, our interpretations should be guided at least in part by what central figures in the debate have claimed, including what they have claimed about the mutual relations among cognitivism, internalism, and desire-based views of motivation. First, I think that many parties to debates about moral motivation have in fact accepted and relied on my stronger formulation of internalism.<sup>4</sup> This certainly seems true of the noncognitivists. Second, only the stronger version of internalism makes the puzzle genuinely inconsistent; because Smith's formulation of the problem employs the weaker version of internalism, his triad is not genuinely inconsistent. This will become clearer when I discuss Smith's own solution (Sec. IV); basically, the idea is that he can square internalism with (1), (3), and (4) ([1s] and [3s]) only because he understands it in this weaker way. One consequence of this is that it requires us to interpret the familiar arguments for rejecting one element of the puzzle on the strength of the others as invalid. This gives us some reason to interpret the elements of the puzzle so as to make them jointly inconsistent and to make possible valid arguments for noncognitivism, externalism, and rationalism, provided this does not distort the way people have understood the constitutive elements. For these reasons, my discussion will be premised on my formulation of the puzzle; I will comment on how our different formulations of internalism affect the argument, where this is appropriate.

## II. NONCOGNITIVISM

The noncognitivist appeals to (2)–(4) to reject the cognitivist claim in (1). If moral judgments entail motivation, motivation involves pro-

4. Compare "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" (Richard Price, *The Principal Questions of Morals*, ed. D. D. Raphael [Oxford: Clarendon Press, 1974], p. 186). "'Goodness' must have, so to speak, a magnetism. A person who recognizes X to be 'good' must *ipso facto* acquire a stronger tendency to act in its favor than he otherwise would have had" (C. L. Stevenson, "The Emotive Meaning of Ethical Terms," reprinted in his *Facts and Values* [New Haven, Conn.: Yale University Press, 1963], pp. 10–31, p. 13). "To think that you ought to do something is to be motivated to do it. To think that it would be wrong to do something is to be motivated not to do it" (Gilbert Harman, *The Nature of Morality* [New York: Oxford University Press, 1977], p. 33). "But it is also held that just knowing them [objective values] or 'seeing' them will not merely tell men what to do but will ensure that they do it, overruling any contrary inclinations" (J. L. Mackie, *Ethics: Inventing Right and Wrong* [New York: Penguin, 1977], p. 23). "It seems to be a conceptual truth that to regard something as good is to feel a pull towards promoting or choosing it, or towards wanting other people to feel the pull towards promoting or choosing it" (Simon Blackburn, *Spreading the Word* [New York: Oxford University Press, 1984], p. 188). Though Blackburn uses this interpretation of internalism to motivate his expressivist view, he goes on to mention some reasons for thinking this version of internalism too strong (*Spreading*

attitudes, and there is no necessary connection between an appraiser's beliefs and pro-attitudes, then moral judgments must express the appraiser's noncognitive attitudes, rather than her beliefs.<sup>5</sup> Though noncognitivism represents one obvious solution to the puzzle about moral motivation, I, like Smith, believe that we should accept noncognitivism only if the other solutions prove unacceptable.

For one thing, cognitivism is the natural starting point for any domain whose discourse is descriptive. When a set of judgments are expressed in the declarative mood, involve singular reference to and quantification over various sorts of entities, and employ predicates as modifiers of noun and verb phrases, it is natural to treat them as making cognitively meaningful assertions, having truth conditions. People *come* to reject cognitivist views when they become convinced that there are insurmountable metaphysical, epistemological, or semantic objections to a cognitivist view.

This is true in metaethics as well. The syntax of moral discourse is descriptive. Moral judgments are expressed in the declarative mood and treat moral predicates as modifiers of noun and verb phrases. As such, they appear to express propositions ascribing moral properties to persons, actions, and institutions; acceptance of these propositions seems to take the form of belief. We ought to give up a cognitivist construal of moral discourse only if cognitivism can be shown to involve unacceptable commitments.

As a historical matter, cognitivism has functioned as the default metaethical view. Analytical ethical theory began this century in a cognitivist (realist) form, namely, *intuitionism*. Intuitionists—such as Sidgwick, Moore, Ross, Broad, and Prichard—all combined a belief that moral properties are nonnatural properties and a foundationalist epistemology, according to which moral knowledge must rest on self-evident truths, with the assumption that there are moral truths that we can discover or intuit. Noncognitivism developed as a reaction to what was thought of as the extravagant metaphysical and epistemolog-

---

*the Word*, pp. 188–89). Unfortunately, he does not explain how this argument for expressivism survives a weakening of the internalist premise.

5. See A. J. Ayer, *Language, Truth, and Logic*, 2d ed. (London: Gollancz, 1946), chap. 6; Stevenson, "The Emotive Meaning of Ethical Terms"; R. M. Hare, *The Language of Morals* (Oxford: Clarendon Press, 1952), chaps. 5–7; Blackburn, *Spreading the Word*, pp. 188–89; and Allan Gibbard, *Wise Choices, Apt Feelings* (Cambridge, Mass.: Harvard University Press, 1990), esp. chaps. 3–4. Compare David Hume: "Since morals, therefore, have an influence on the actions and affections, it follows, that they cannot be deriv'd from reason; and that because reason alone, as we have already prov'd, can never have any such influence. Morals excite passions, and produce or prevent actions. Reason of itself is utterly impotent in this particular. The rules of morality, therefore, are not conclusions of our reason" (*Treatise of Human Nature*, ed. P. H. Niddich [Oxford: Clarendon Press, 1978], bk. 3, pt. 1, sec. 1, p. 457).



ical commitments of intuitionism. Noncognitivists argued that we were better-off giving up the cognitive and realist commitments of moral discourse than accepting the metaphysical, epistemological, and semantic consequences of those commitments.

If there is a difference between cognitivism in ethics and in science it is, I think, only that the objections to cognitivism in ethics seem more obviously decisive to some. But that conclusion is perfectly consistent with assigning cognitivism the default position in metaethics. Moreover, recent attempts to defend cognitivism and realism in ethics without the metaphysical and epistemological baggage of intuitionism make it far from clear that cognitivist commitments in ethics have problematic metaphysical, epistemological, or semantic consequences or that there is any special problem being a cognitivist about ethics.

In any case, the noncognitivist owes us an account of why we should not take the descriptive aspects of moral discourse at face value. Moreover, there are a number of familiar semantic problems for noncognitivism, which I will briefly sketch.

One such problem involves the phenomena of unasserted contexts. The noncognitivist construes moral assertion as the expression of the appraiser's attitudes, rather than as a description of the way the world is. But, as Peter Geach observed, the noncognitivist proposal fails to account for the meaning of moral predicates in *unasserted* contexts, such as the antecedents of conditional statements.<sup>6</sup>

1. If it is *wrong* to murder innocent children, then it is wrong to pay someone else to murder innocent children.
2. It is wrong to murder innocent children.
3. Hence, it is wrong to pay someone else to murder innocent children.

Whereas the second premise asserts something about and perhaps expresses an attitude toward murdering innocent children, the first premise does neither (it asserts something about and perhaps expresses an attitude toward paying someone to murder innocent children conditional on the wrongness of murder). Because the first occurrence of the moral predicate "wrong" in premise (1) is in an unasserted context, traditional forms of noncognitivism seem to have no account of its meaning. This is bad enough. But the argument certainly seems valid, and for this to be true all four occurrences of the predicate must have the same meaning. But then the predicate "wrong" must mean the same thing in both asserted and unasserted contexts. Geach attributes the general point about the univocity of terms across asserted

6. See Peter Geach, "Assertion," *Philosophical Review* 74 (1965): 449–65.



and unasserted contexts to Frege. Whereas the cognitivist can accept Frege's point, the noncognitivist apparently cannot.<sup>7</sup>

Another problem for noncognitivism is how to explain the validity of moral arguments in ways that do not invoke truth. The problem can be illustrated with the argument involving unasserted contexts, but there are many valid arguments that involve only asserted contexts.

1. It is wrong to murder innocent children.
2. It is wrong to pay someone to murder innocent children.
3. Hence, it is wrong to murder innocent children, and it is wrong to pay someone to murder innocent children.

A natural explanation of why these and other argument forms are valid is that it would be inconsistent to assert the premises and deny the conclusion, and the natural explanation of inconsistency is truth-theoretic: inconsistent claims cannot all be true. But the classical noncognitivist denies that moral judgments are either true or false; a fortiori, moral judgments cannot be inconsistent (in this sense). Perhaps there is some sort of pragmatic objection to affirming the premises of these arguments while denying their conclusions that is not itself parasitic on the truth-theoretic account of validity. But this needs to be shown, and, in any case, the noncognitivist apparently lacks this straightforward and compelling account of what makes these arguments good.

Some seem to think that the noncognitivist can help herself to an account of truth by appealing to disquotational or redundancy accounts of truth or its ascription.<sup>8</sup> However, it is arguable that this conflates (a) disquotational or redundancy accounts of *ascriptions* of truth and (b) disquotational or redundancy accounts of *truth*. According to (a), to say that someone's moral judgment is true is to agree with that judgment.

"It is true that p" means the same as "p."

So, for example, an instance of this schema is:

"It is true that murder is wrong" means the same as "murder is wrong."

According to (b), truth for moral judgments involves no more than the following schema for moral sentences:

7. For interesting attempts to address this problem, see Blackburn, *Spreading the Word*, pp. 189–96, and "Attitudes and Contents," *Ethics* 98 (1988): 501–17; and Gibbard, pp. 92–99; cf. Bob Hale, "Can There Be a Logic of the Attitudes?" in *Reality, Representation, and Projection*, ed. John Haldane and Crispin Wright (Oxford: Oxford University Press, 1993), pp. 337–63.

8. See Stevenson, *Facts and Values*, pp. 214–20; cf. Paul Horwich, "Gibbard's Theory of Norms," *Philosophy and Public Affairs* 22 (1993): 67–78, esp. pp. 72–76.

"p" is true iff p.

So, for example, an instance of this schema is:

"Murder is wrong" is true iff murder is wrong.

I'm not sure how plausible (a) is; it makes a claim about the semantic equivalence of two different judgments or utterances. The second judgment or utterance is a first-order judgment or utterance, whereas the first is a second-order judgment or utterance that takes the second as its object. Perhaps the ground for asserting the two is the same, but it doesn't follow that their meaning is the same. In fact, their meaning would seem to be different insofar as the first judgment is about a judgment whereas the second is not. Whereas (a) strikes me as implausible, I see no special reason the noncognitivist cannot endorse it. But (b) is quite different, and it is less clear how the noncognitivist can endorse it. In (b), the first half of the biconditional is a judgment or utterance, but the second half is neither. A disquotational account of truth for moral judgments seems to require moral states-of-affairs, or facts, or some such truth-maker for the second half of the biconditional. But the classical noncognitivist thinks that moral predicates are nonreferring and so cannot recognize moral truth-makers. In fact, this worry is related to the worry about unasserted contexts; (b) involves moral predicates occurring in unasserted contexts, namely, the right-to-left conditional.

I mention these problems for the noncognitivist not because I am sure that they are insurmountable, but to point to costs or, at least, obstacles the noncognitivist solution to the puzzle about moral motivation faces. Before we abandon the default cognitivist construal of moral discourse and tackle semantic problems associated with noncognitivism, we should examine other solutions to the puzzle about moral motivation.

### III. ONE KIND OF RATIONALISM: MORAL MOTIVATION BY BELIEF ALONE

The externalist solution may also seem to be a solution of last resort, because it may seem to deny the platitude that moral judgments are motivationally efficacious. For this reason, we might look seriously at rationalist theories of moral motivation, because they promise to represent moral judgments as intrinsically motivational without giving up cognitivism.

Some philosophers, such as Nagel and McDowell, maintain cognitivism and internalism about motivation by rejecting the assumption that motivation requires a desire or pro-attitude; they insist that purely cognitive states—beliefs—can motivate.<sup>9</sup> On this view, even if motiva-

9. See Thomas Nagel, *The Possibility of Altruism* (Princeton, N.J.: Princeton University Press, 1970), esp. pp. 29–30; John McDowell, "Are Moral Requirements Hypotheti-

tion often requires pro-attitudes, this is not universally true and, in particular, not true of normative motivation. Both Nagel and McDowell introduce a cognitive view about moral motivation by appeal to a cognitive view of prudential motivation. Prudential motivation can occur, on this view, when an agent recognizes that a course of action promotes her own interests. In explaining her behavior, we do not need to appeal to a current concern, on the agent's part, for her own future well-being. Similarly, on this view, moral motivation can occur when an agent sees that morality requires a particular course of action. If we insist on ascribing to the person who acts on her moral beliefs a desire to do the act in question, this ascription of pro-attitude is "merely consequential" on our interpreting her behavior as intentional; the action is produced by her moral beliefs and does not depend upon prior, independent conative states. McDowell takes this purely cognitive picture of motivation to be characteristic of the virtuous person's psychology; the virtuous person need only see what morality requires to be motivated to act.<sup>10</sup>

Like Smith, I do not find this form of rationalism compelling. Even if some desires or other pro-attitudes are ascribed merely consequentially, it is arguable that the motivation of all intentional action, including moral motivation, requires the existence of independent conative states or pro-attitudes. Consider a garden-variety case of non-moral motivation. I believe that it is raining. I want to go out, and I want to stay dry. I believe that an umbrella would keep me dry and that I have an umbrella in the closet. So I go to the closet to get my umbrella. Now it might be merely consequential to ascribe to me the proximate desire to get my umbrella from the closet; perhaps ascription of this desire just follows from the assumption that my action of getting my umbrella from the closet was intentional. I don't know. But it seems clear to me that the motivation for my action requires some less proximate and more ultimate desires—in particular, my desires to go out and to stay dry. If we count appeal to my belief that it is raining as explaining my action, this is only because we take my desires to go out and to stay dry for granted as part of my psychological background. If I didn't have these more ultimate desires, my beliefs

---

cal Imperatives?" *Proceedings of the Aristotelian Society, Supplement* 52 (1978): 13–29, esp. pp. 15–16, and "Virtue and Reason," *Monist* 62 (1979): 331–50; cf. Samuel Scheffler, *Human Morality* (New York: Oxford University Press, 1992), esp. pp. 86, 87, 90–92.

10. McDowell's attraction to a purely cognitive view of moral motivation is clearest in his earlier writings, especially "Are Moral Requirements Hypothetical Imperatives?" and "Virtue and Reason." This rather sharp contrast between cognitive and conative states gets challenged in his more recent work; see, e.g., John McDowell, "Values and Secondary Qualities," in *Morality and Objectivity*, ed. T. Honderich (London: Routledge & Kegan Paul, 1985).

about the rain and my umbrella would be impotent (or produce different results). Similarly with moral motivation. I believe that fairness requires me to keep my promise to you, even at significant personal cost to myself. I want to be fair. So I keep my promise to you, even at some cost to myself. Perhaps ascription to me of a proximate desire to keep my promise to you is merely consequential on interpreting my promise-keeping behavior as intentional. I don't know. But the motivation for my action does involve my more ultimate desire to be fair. If we count appeal to my beliefs about the requirements of fairness as explaining my action, it's only because we're taking my commitment to being fair for granted as part of my psychological background. If I did not have this more ultimate desire or commitment, my moral belief would lead nowhere (or elsewhere).

As Smith argues, one reason it seems implausible that beliefs might motivate without the appropriate desires or pro-attitudes is that intentional action seems to be the product of two kinds of intentional states—what we might call “representational” and “practical” states.<sup>11</sup> The difference between them, as Smith and others sometimes put it, consists in the “direction of fit” they bear to the world. On this view, the intentional states of an agent are representational insofar as she seeks to adjust them to conform to the world, whereas they are practical insofar as she seeks to adjust the world to conform to them. Cognitive states, such as beliefs, are representational, whereas pro-attitudes, such as desires, are practical. Intentional action, on this view, is the attempt to bring the world into line with one's practical states in ways that are constrained by one's representational states. If this general picture of intentional action is on the right track, a purely cognitive approach to moral motivation is implausible for quite general reasons. A belief that morality requires a certain action will be motivational only in conjunction with a desire or other practical commitment to being moral.<sup>12</sup>

So what do we say about the virtuous person, for whom moral belief appears sufficient to motivate? We can agree that moral beliefs are sufficient to motivate the virtuous person, but this is only because a virtuous person is someone with a certain well-developed psychological

11. Smith traces the view to Elizabeth Anscombe, *Intention* (Ithaca, N.Y.: Cornell University Press, 1957), p. 56. However, the view certainly has older origins. Something like it seems to be Aristotle's view in *De Anima* 3.3–9 and can be found in T. H. Green, *Prolegomena to Ethics* (New York: Crowell, 1969), secs. 86–87, 131–32, 136. Other contemporary discussions include Dennis Stampe, “The Authority of Desire,” *Philosophical Review* 96 (1987): 335–81, p. 355; I. L. Humberstone, “Direction of Fit,” *Mind* 101 (1992): 59–83; and David Velleman, “The Guise of the Good,” *Nous* 26 (1992): 3–26.

12. Whether or not these practical commitments have the particular functional profile of desires (whatever exactly it is), they have an abstract functional profile that is practical, and this makes them part of the larger class of pro-attitudes.

profile that is structured by various cognitive and conative states (cf. pp. 122–23). The virtuous person has beliefs about what justice, rights, mutual respect, friendship, and compassion require of agents and has standing aims to be just, to respect the rights and dignity of others, to be a true friend, and to display kindness. When the virtuous person adds to her background psychological states beliefs about what these moral categories require in particular circumstances, she may be motivated to act, but this will be in virtue of her cognitive and conative background and not simply because of her newly acquired cognitive states.

#### IV. ANOTHER KIND OF RATIONALISM: DESIRES THAT REFLECT NORMATIVE BELIEFS

A different form of rationalism denies that motivation is possible without a prior pro-attitude but insists that certain beliefs, in particular, normative beliefs, entail pro-attitudes. On this view, the belief that I have reason for action generates a desire to perform the action in question. Indeed, our account of the nature of representational and practical states explains why our pro-attitudes should be sensitive to our normative beliefs. Whereas the belief that the world should be a certain way is, for the cognitivist, a representational state, it is one that tends to bring forward a practical state to which the agent aims to make the world conform. Believing it is best that things be a certain way tends to produce a desire or pro-attitude to make things be that way.<sup>13</sup> On this version of rationalism, pro-attitudes are not merely consequentially ascribed because the action is intentional; the pro-attitudes are psychologically real prior to the action and play an ineliminable role in generating and, hence, explaining action. But these pro-attitudes are consequential or dependent on the belief that one should perform the action.

If this version of rationalism about normative motivation is to apply to moral motivation, two conditions must be met: (a) belief that an action is rationally authoritative—that is, belief that there is a reason to perform that action, such that failure to perform it is *pro tanto* irrational—must generate a pro-attitude toward that action, and

13. Insofar as intentional action aims at what is best, I think one can explain the way in which motivation is sensitive to normative belief without the aid of any analysis of practical reason and certainly without supposing, as Smith does, that it is a conceptual truth that to judge that one has a reason to do something is to judge that one would desire to do it if one were fully rational. However, I suspect that the contrast between our explanations is not as great as might first appear insofar as the analysis of judgments of practical reason, on which he makes his explanation depend, is nonreductive (it analyzes judgments about one's reasons for action into judgments about what one would desire *if one were fully rational*).

(b) belief that an action is morally required must involve belief that the action is rationally authoritative. Though no labels seem satisfactory, it will help to have labels for these claims. The first claim is a motivational internalist claim about judgments of practical reason; we might call it “normative internalism.” The second claim is a rationalist claim about the concept of moral judgment; we might call it “appraiser rationalism.”

Smith’s own solution to the motivational problem is similar to this rationalist solution. He accepts appraiser rationalism insofar as he construes moral judgments as judgments of practical reason. When I judge there to be (normative) reason to engage in a particular course of action, I am, on Smith’s view, judging that I would desire to engage in that course of action if I was fully rational, that is, deliberating correctly.<sup>14</sup> To avoid instrumentalist conclusions about practical reason, Smith can and does allow correct deliberation to include demands of explanatory coherence, as well as full empirical information and means-ends reasoning (pp. 155–61). He proposes to treat moral requirements as reasons having a particular content or substance, though he does not himself spell out what these contentful constraints are. So, on this view, to judge that a certain course of action is morally required is to judge that we would desire to act that way if we were fully rational, where the act in question “is an act of the appropriate substantive kind” (p. 184). It follows from this analysis that the belief that an action is morally required involves the belief that it is (*pro tanto*) rational or authoritative. Smith also accepts a version of normative internalism; he believes that *ceteris paribus* judgments about our normative reasons motivate. If beliefs about what we have normative reason to do are beliefs about what we would desire to do if we were fully rational, then judgments of practical reason rationally should affect one’s desires and will insofar as one is rational (p. 177). As Smith notes, this claim and appraiser rationalism imply a version of the internalist thesis about moral motivation (pp. 143–45).

However, as Smith recognizes, his claims do not vindicate the strong internalist claim, on which noncognitivist arguments rely, according to which moral judgment entails motivation. Just as he accepts a weaker version of the internalist thesis about moral motivation, so too he accepts a weaker version of normative internalism. Whereas

14. It might seem to be a problem for this view that rationality sometimes requires indirection. I may have a reason to  $\phi$  without it being rational for me to desire to  $\phi$  if successful  $\phi$ ing comes only to those who do not aim at this result. But Smith’s view avoids this problem, I think, if it is interpreted as the claim that I have reason to  $\phi$  just in case there is some true description  $\psi$  of  $\phi$  under which I would desire to  $\psi$  if I was fully rational. Smith offers his own response at pp. 212–13, which is discussed by Geoffrey Sayre-McCord, “The Metaethical Problem” (in this issue).



he does think that judgments of practical reason normally motivate, he does not think that they must. On this point, he and I are in agreement.

If I think that from the point of view of what fundamentally matters in practical deliberation I have reason to  $\phi$ , then it seems natural for me, for that reason, to desire or aim to  $\phi$ .<sup>15</sup> As I noted earlier, one can think that this connection between practical judgment and motivation is necessary without supposing there is a necessary connection between motivation and action. It is not plausible to think that judgments of practical reason are sufficient for action. Weakness of the will is possible; it is sometimes true that, though I judge X to be better than Y, when the time for action arrives I choose to do Y. In some cases of weakness of will, the breakdown comes between motivation and action, not between practical judgment and motivation. But breakdown can come between practical judgment and motivation. What seems plausible, and what Smith defends, is the claim that *ceteris paribus* judging an action reasonable or rational produces a motivation to perform that act. But other things are not always equal. As Smith points out, in cases of severe depression and apathy, it seems, one can make practical judgments without motivational effect (pp. 119–20).<sup>16</sup> Moreover, there is now neuropsychological evidence that suggests that patients with damage to a number of distinct brain regions (especially the prefrontal lobe of the cerebral cortex) continue to make the same practical judgments they made before their injuries or lesions but without motivational effect.<sup>17</sup> Where there is such physical and psychological interference, practical judgment does not produce motivation. If so, we must deny that judgments of practical reason entail motivation. If moral judgments are judgments of practical reason, we must deny that moral judgments entail motivation.

This is already to recognize one kind of *amoralist*—someone who recognizes moral requirements yet remains indifferent. The possibility of this sort of amoralism undermines the strong internalist assumption about moral motivation. If, as I believe, the puzzle about moral motivation is best construed as resting on this strong internalist assumption, then recognition of the possibility of this sort of amoralism is sufficient to vindicate the externalist solution. Because Smith formulates his

15. As n. 14 acknowledges, my pursuit of what I believe I have reason to do ought sometimes to employ indirection.

16. Compare Michael Stocker, "Desiring the Bad: An Essay in Moral Psychology," *Journal of Philosophy* 76 (1979): 738–53; and Alfred Mele, "Internalist Moral Cognitivism and Listlessness," *Ethics* 106 (1996): 727–53.

17. The case of Phineas Gage and other clinical data, for which this is a natural interpretation, are described in Antonio Damasio, *Descartes' Error: Emotion, Reason, and the Human Brain* (New York: Putnam, 1994), chaps. 1–4. I would like to thank Allan Gibbard for drawing my attention to this work.



version of the moral problem with a weaker version of the internalist assumption, he does not see this kind of amorality as a threat to internalism.

## V. EXTERNALISM ABOUT MOTIVATION

So far, my disagreement with Smith is largely (though perhaps not entirely) verbal. But though I find his form of rationalism more plausible than those that rely on a purely cognitive account of moral motivation, I find the externalist solution more plausible. My skepticism about this form of rationalism does not merely reflect skepticism about the strong version of normative internalism, about which he and I agree. Unlike Smith, I am also skeptical about appraiser rationalism. I think we can have a belief that an action is morally required without the belief that it is rationally authoritative.<sup>18</sup> To defend this claim is, in effect, to reject internalism about moral motivation, even Smith's weaker version.

One way to explain the appeal of this solution is by appeal to the possibility of another form of amorality. The kind of amorality whose possibility Smith concedes is a kind of *unprincipled* amorality. It is unprincipled, not in the sense that it is random or lacks a psychological explanation, but in the sense that it is due to psychological interference with the normal process by which results of practical deliberation affect an agent's motivational set; indifference does not reflect principles the agent accepts. But another form of amorality is *principled*. In fact, principled amorality actually rests, in part, on (weak) normative internalism. It's in part because our motivational states normally track our beliefs about what we have reason to do that we can imagine a principled amoralist who is indifferent to what she judges morally required. Because moral motivation is predicated on the assumption that moral requirements generate reasons for action or have rational authority, it is possible to make moral judgments and yet remain unmoved as long as there are possible conceptions of morality and practical reason according to which moral requirements need not have rational authority.

It seems to me that such conceptions of morality and practical reason are not only possible but also familiar. For instance, as long as we associate morality with an impartial point of view that imposes other-regarding duties and accept an agent-centered conception of practical reason that rests on instrumental or prudential conceptions

18. See Philippa Foot, "Morality as a System of Hypothetical Imperatives," *Philosophical Review*, vol. 81 (1972), reprinted, with postscript, in her *Virtues and Vices* (Berkeley: University of California Press, 1978); and my *Moral Realism and the Foundations of Ethics* (New York: Cambridge University Press, 1989), chap. 3.

of practical reason, it seems we must recognize the possibility of moral requirements that it is not irrational to disobey.

1. Moral requirements include impartial other-regarding obligations that do not apply to agents in virtue of their aims or interests.
2. Rational action is action that achieves the agent's aims or promotes her interests.
3. There are circumstances in which fulfilling other-regarding obligations would not advance the agent's aims or interests.
4. Hence, there can be (other-regarding) moral requirements such that failure to act on them is not irrational.

It is important to see that the antirationalist conclusion of this argument does challenge the authority of ethics. The antirationalist conclusion would not be troubling if morality and rationality were two independent but coordinate evaluative perspectives. For then it might seem to be an open question whether an agent should side with morality or rationality when they conflict. But practical rationality is not just one standard or perspective among others, with no obviously privileged position; it should be understood to concern whatever fundamentally matters in practical deliberation or whatever it is ultimately reasonable to do. So, for example, if I have doubts about whether I have reason to act on a particular norm, I should be interpreted as having doubts about whether that is a norm of practical rationality, rather than as having doubts about rationality. If so, antirationalism denies that morality should always have authority in our deliberations.

So understood, the antirationalist argument does challenge the appraiser rationalism on which internalism about moral motivation rests. But that challenge does not require that the antirationalist argument be sound; in particular, the challenge does not depend upon the truth of an impartial conception of morality or an agent-centered conception of practical reason. The challenge requires only that we treat the rational authority of morality as an open question. One reason to do so is to recognize antirationalist challenges of this familiar sort. What matters are what beliefs one can hold, not their truth; if a person's beliefs about morality, rationality, and auxiliary issues lead her to doubt the authority of morality, we can see how she could fail to be motivated by her moral judgments.

The internalist must hold not only that the best conceptions of morality and of practical reason converge in their demands but that we cannot hold conceptions of moral requirements and practical reason that could diverge in this way.<sup>19</sup> We could conceive of practical

19. We cannot allow the best conception of a concept to establish truths about the concept that constrain eligible conceptions of that concept and the ascription of beliefs to inquirers employing those concepts, for that would make impossible the ascription

reason in agent-centered terms, and we could conceive of morality in impartial terms, but not both. If we think of practical reason in agent-centered terms, then we would have to think of morality in the same agent-centered terms, and if we think of morality in impartial terms, then we must conceive of practical reason in impartial terms. Anyone who claims to conceive of morality and practical reason in terms that allow them to diverge must be mistaken; their claims about morality or practical reason should be reinterpreted. But while this maneuver would preserve internalism about moral motivation intact, I see little to recommend it. If practical reason admits of agent-centered conceptions and morality admits of impartial conceptions, why can't someone hold both conceptions? What prevents this?

One thing that would prevent this is if we were forced to think of moral requirements as requirements of practical reason.

1. If I am under a moral requirement to  $\phi$ , there is a moral reason for me to  $\phi$ .
2. If there is a moral reason for me to  $\phi$ , there is a reason for me to  $\phi$ .
3. If there is a reason for me to  $\phi$ , it would be *pro tanto* irrational for me to fail to  $\phi$ .
4. Hence, if I am under a moral requirement to  $\phi$ , it would be *pro tanto* irrational for me to fail to  $\phi$ .

But this defense of appraiser rationalism and internalism about moral motivation is not compelling. Sometimes when we say that I have a reason to  $\phi$ , we mean

- a) There is a behavioral norm that enjoins  $\phi$ ing and applies to me.

Of course, in this sense of reason, moral norms do imply reasons. There are as many kinds of reasons as there are norms, including moral reasons, legal reasons, reasons of etiquette. But we often have something more in mind in ascribing reasons; we mean

- b) There is a behavioral norm that enjoins  $\phi$ ing, it applies to me, and it would be *pro tanto* irrational for me not to  $\phi$ .

If there is reason, in this sense, to act on a norm, then practical reason endorses this norm. But not all reasons for action in the first sense are reasons for action in the second sense.<sup>20</sup> For instance, it is arguable

---

of mistaken conceptions of a concept or mistaken beliefs about the extension of the concept.

20. Hart distinguishes between the internal and external aspects of rules; see H. L. A. Hart, *The Concept of Law* (Oxford: Clarendon Press, 1961), pp. 55–57. Whereas belief that there is an (*a*)-type reason to act seems to require only an external view of norms, belief that one has a (*b*)-type reason to act seems to require or perhaps produce an internal view of norms.

that failure to conform to requirements or reasons of etiquette or law need not be *pro tanto* irrational. It is clear that moral requirements are moral reasons and that moral reasons are reasons in the sense of behavioral norms. It is not clear that they generate reasons in the sense that failure to conform to the behavioral norms is *pro tanto* irrational. So the sense of reason for action in which (3) is true need not be the same sense of reason for action in which (2) is true. If so, it's arguable that it is only by failing to distinguish these two senses of reason for action that the rational authority of morality could fail to seem an open question.

I see no reason to deny the possibility of the principled amoralist. If the rational authority of morality is an open question, then it's possible to make moral judgments without being motivated to act. If so, the (principled) amoralist is conceivable, even if her indifference can be shown to rest on mistaken conceptions of morality, rationality, or other auxiliary issues. If so, the externalist solution to the puzzle about moral motivation is the most plausible. But notice that if the externalist solution is correct, the noncognitivist solution is not only unnecessary but unavailable. For if it is possible to be unmoved by moral judgments, they cannot essentially express pro-attitudes, as the noncognitivist claims.

## VI. SMITH'S ANTIEXTERNALISM

Smith rejects the possibility of principled amoralism; he claims, in effect, that it would never be reasonable to interpret someone's judgments, to which she was genuinely indifferent, as moral judgments. Though a putative amoralist might make roughly the same discriminations among actions, policies, and people that moralists do, might call her judgments moral judgments, and might insist that they reflect her own moral standards, not those of someone else (e.g., prevailing community standards), Smith insists that it could not be right for us to interpret her judgments as moral judgments. Instead, he thinks we should always interpret her use of moral language as being in inverted commas; she is indifferent to the moral standards that others around her accept, not to standards she herself treats as moral.<sup>21</sup>

21. Sometimes the debate between internalists and externalists depends on the two employing different criteria for identifying judgments as moral judgments. Often, internalists employ functional criteria—roughly, those judgments are moral judgments that the appraiser treats as fundamentally important—whereas externalists employ contentful criteria—roughly, those judgments are moral judgments that concern certain sorts of matters, for instance, having to do with the welfare of affected parties. Consider our putative amoralist—someone who professes indifference to requirements that track matters plausibly thought to be moral in content, on the ground that she thinks they sometimes lack rational authority, yet she persists in regarding them as moral requirements. Whereas functional criteria require interpreting the judgments she calls moral

Smith defends this interpretation of the putative amoralist by appeal to an allegedly analogous debate about whether certain kinds of visual experience are essential to having color concepts and expressing color judgments (pp. 68–70). On one view about color concepts, it is constitutive of possession of color concepts that one have certain visual experiences under suitable conditions. On this view, a blind person who uses color terms, perhaps even in reliable ways, to pick out colored objects does not possess color concepts and so does not succeed in making genuine color judgments when she uses color terms. Similarly, Smith wants to claim, being motivated by what one judges moral is partly constitutive of possessing moral concepts and expressing moral judgments. As a result, the putative amoralist must lack moral concepts and be unable to express moral judgments to which she could then express indifference. Smith's main point does not require accepting the claim about visual experience being constitutive of color concepts. His point is that it would be question begging for someone to appeal to the possibility of a blind person making color judgments as evidence for thinking that color concepts did not presuppose visual experiences. To interpret the blind person's judgments as color judgments is, without further argument, just to deny that visual experience is necessary for color concepts; it is not an argument for that claim. Similarly, he wants to claim, it is question begging for me to appeal to the amoralist in support of externalism; to interpret the putative amoralist's judgments as moral judgments is, without further argument, just to embrace externalism, not to argue for it. This conclusion, by itself, would not help establish internalism, but it might take some wind out of the sails of externalism.

Am I simply begging the question? I don't think so. The internalist makes a strong generalization about the connection between moral judgment and motivation. By appealing to the principled amoralist, I am offering what I take to be a counterexample. My claim is not that it is incoherent to deny the possibility of the amoralist or that the

---

judgments as inverted commas judgments and interpreting her judgments of practical reason, whatever their content, as her true moral judgments, contentful criteria require or at least allow us to take her self-description at face value, as an expression of principled indifference to what she judges morally required. An interesting and potentially puzzling feature of Smith's view is that he seems to combine functional and contentful criteria. Whereas his rejection of principled amoralism seems to reflect at least implicit appeal to functional criteria, elsewhere, as we have seen, he suggests that moral requirements are reasons with a particular substance or content (p. 184). I'm not sure whether this is a tension in his view or not. Whereas I think I am committed to rejecting this functional criterion, I don't know that I have to suppose that there is some content that moral judgments essentially have. I suppose I am suspicious of any significant conceptual constraints on when we must interpret a person's judgments as moral judgments.

internalist is forced to be inconsistent. Rather, it is my view that internalism is implausible, because it denies the possibility of what seems to me possible. Whereas I'm not sure that I have any very firm belief about whether blind people could make color judgments, it does seem quite plausible to me that someone could make moral judgments without always being motivated to act accordingly. Moreover, I don't think I've ever been quite content simply to appeal to the possibility of the amoralist, without further argument or at least explanation. I have always tried to explain the amoralist's indifference as reflecting doubts about the authority of morality. Here and elsewhere, I have tried to elaborate upon this explanation. Because motivation can and does normally track beliefs about what one has reason to do, it is possible to make moral judgments and yet remain unmoved as long as one's beliefs about morality, practical reason, and auxiliary issues imply that some moral requirements lack rational authority. In arguing this way against weak internalism, I think am reasoning much as Smith himself does in arguing against strong internalism. He rejects strong internalism because he thinks that we can, through apathy or depression, make moral judgments (or, more generally, judgments of practical reason) without our wills being engaged. Of course, the strong internalist could insist that motivation is constitutive of moral judgment and that, as a result, despite the fact that the apathetic person calls her judgments moral judgments, they are not really moral judgments but only inverted commas judgments. The strong internalist might then claim that appeal to the apathetic moralist is question begging. I assume that Smith might reasonably reply that, whereas the strong internalist has a consistent position, it commits him to claims we find implausible or at least insufficiently motivated and that, all else being equal, this is reason to reject the strong internalist claim. That's what I want to say about the (weak) internalist claim that principled amorality is impossible.

As I say, Smith's complaint about my appeal to the principled amoralist does not require him to endorse the claim that visual experience is essential to color concepts. Nonetheless, many have endorsed the analogy between colors and morals, and someone might appeal to the claim that visual experience is essential to color concepts and judgments to argue for the internalist thesis that motivation is essential to moral judgment. However, I'm skeptical of this use of the color analogy against externalism on several grounds.

First, it's not clear to me that certain visual experiences are constitutive of color concepts if possession of those concepts is necessary for making color judgments. If I'm blind, and have been so from birth, then I do lack visual experiences associated with colored objects. But I might have learned a good bit else, presumably from sighted people, about colors—for instance, about relations between primary and sec-



ondary colors, about warm and cold colors, about how colors of objects supervene on properties of light refraction, and about the colors that objects in my environment have. Though I clearly lack some sensory information about colors that the sighted have, this doesn't prevent me from using color terms meaningfully, as when I ask my (sighted) son to sort the objects on the table into those that are blue and those that are not or as when I assert that my Gala apple is more red than your Granny Smith. Perhaps the possibility of my use of color terms to refer is in some way parasitic on my being part of an intellectual and linguistic community with sighted people. Even so, this doesn't show that my use of color terms involves inverted commas. For I do not ask my son to sort objects that give him a certain visual experience; rather, I ask him to sort objects according to whether they are in fact blue. But then even if the amoralist lacks motivation that the moralist possesses, why should we deny that she makes moral judgments?

Second, it seems most plausible to deny the blind person color judgments if the person is completely blind from birth. It seems hard to deny that the person has color concepts or to insist that she is unable to express color judgments using color terms on the ground that she now lacks requisite visual experience if blindness is a recent development, as it might be for an aged person. When a person who has only recently lost her sight says that her faithful Irish setter is reddish brown, it's hard to doubt that she makes a color judgment. But the principled amoralist need not be congenitally amoral. She might long have been a moralist, either principled or unprincipled. This is the way it is with many unprincipled amoralists. They made moral judgments with motivational effect prior to the onset of depression or neurological damage. They continue to make the same discriminations and judgments using moral terms but now without motivational effect. Indeed, it is this continuity of cognitive ability, despite a change in motivational attitude, that makes it natural to interpret them now as (still) making moral judgments that (now) have no motivational effect. The same can happen with principled amoralists. Having always been motivated by one's moral judgments in the past, perhaps because one tacitly assumed that moral requirements were authoritative, one might come to question the authority of those same requirements. If one then continues to make the same discriminations and judgments using moral terms as before, it seems natural to interpret them still as moral judgments, though now they are made without motivational effect.

Third, it seems most plausible to deny that the blind person makes color judgments if her visual disability is quite general, namely, if she is blind or at least generally color-blind. If her disability with respect to visual and color experience is quite selective—restricted only to certain shades or perhaps to certain objects or perceptual contexts—then it is harder to see the disability defeating the ascription



of color judgments. If there's something that prevents me from seeing the color of this object, but others assure me that it is blue and I have had visual experience of other blue things, then it seems natural to interpret my sentence "This object is blue" as an expression of a color judgment on my part. The amoralist's indifference can be likewise selective. The possibility of a principled amoralist reflects the possibility of conceptions of morality and practical reason according to which failure to conform to moral requirements is not always *pro tanto* irrational. Convergence or overlap in the demands made by one's conceptions of morality and practical reason can be significant without being complete. For instance, it is a common view about the demands of an impartial morality and those of enlightened self-interest not only that they coincide to a considerable extent but also that their coincidence is imperfect and counterfactually unstable. If one's conception of morality is impartial and one's conception of practical reason is prudential and one accepts this common view about the substantial but imperfect coincidence between morality and prudence, the coincidence of moral judgment and motivation will be substantial but imperfect. If the motivationally committed judgments of the amoralist that employ moral terms count as moral judgments, then so too should those to which she expresses principled indifference.<sup>22</sup>

Finally, I must confess some confusion about how the analogy between colors and morals could support an antiexternalist conclusion. Let us concede, at least for the sake of argument, that certain kinds of visual experiences are partly constitutive of the possession of color concepts and the capacity to make color judgments. How does this help show that motivation is partly constitutive of the possession of moral concepts and the capacity to make moral judgments? Visual experience is not motivation. If, as one might think, visual experiences are themselves cognitive states, then the analogy between colors and morals would apparently suggest that certain cognitive states are partly constitutive of representational judgments with moral content. I don't see how this helps establish the internalist thesis that motivation is

22. Selective amoralism could be reconciled with an extremely weak internalist thesis to the effect that it is a precondition of interpreting a set of judgments by an appraiser as her moral judgments that she not be completely indifferent to all of them. This kind of internalism is clearly weaker than even Smith has in mind, and it would clearly be too weak to serve as a premise in familiar arguments for noncognitivism. So, for many of my purposes, I need not deny this weak motivational link between moral judgment and motivation. However, even this thesis seems too stipulative. Though I suspect that the most easily comprehensible forms of amoralism will be selective in this way, I know of no way to rule out the global amoralist. For I don't see on what grounds we can insist that it could never be plausible to interpret someone as holding conceptions of morality and practical reason according to which moral requirements always lacked rational authority.

partly constitutive of moral judgment. Perhaps it might be claimed that visual experience is itself a precognitive or noncognitive state on which representational judgments are based. If so, the analogy between colors and morals might then seem to imply that moral concepts and judgments essentially involve noncognitive states. But this line of reasoning would be problematic in several ways. First, I don't know that it is reasonable to think of visual experiences as precognitive, rather than as representational. Second, even if they are precognitive states, it doesn't follow that they are practical states, such as desires and other pro-attitudes; but then it is again difficult to see how the analogy between colors and morals supports an internalist claim about moral judgment. Finally, if we did think that color concepts and judgments essentially involved noncognitive, practical states, then we might wonder whether the analogy between colors and morals doesn't beg the question against the externalist. For if color judgments, unlike many other sorts of judgments, essentially involve some kind of noncognitive component, then it is not clear why we should assume that moral judgments are especially like color judgments unless we have already accepted internalism. Judgments of etiquette or judgments of law, for example, do not seem to have any noncognitive component essentially; noncoincidentally, there doesn't seem to be any problem interpreting someone as making a judgment of etiquette or law without any commitment of his will. Why think that moral judgments are more like judgments of color than like judgments of etiquette or law?<sup>23</sup>

Remember Smith claims only (though wrongly, I think) that I cannot appeal to the amoralist to support externalism. By itself, this doesn't help internalism or hurt externalism. Smith thinks an independent argument is required (pp. 70–71), and he goes on to argue that the externalist is unable to explain how it is that an agent's motives reliably track her moral beliefs (pp. 71–76). If, having believed that X was morally superior to Y, I change my mind and come to believe that Y is morally superior to X, this normally occasions a change in my motivation: I cease being disposed to bring about X and become disposed to bring about Y. Similarly, if there's been no change in my

23. Smith rejects any conception of moral requirements that represents them as institutional norms, on a par with requirements of law or etiquette (pp. 80–84). One thing that worries Smith is that an institutional conception of morality would represent moral requirements as “only externally related” to our moral motivations (p. 83). However, this is an objection only if we have assumed that internalism is true; but then this aspect of an institutional conception of morality cannot be part of an argument for internalism. A less question begging worry about an institutional conception of morality is that moral judgments seem motivationally more important than judgments of law or etiquette. But the externalist can explain this if he can explain how the institution of morality is more important or authoritative than the institutions of law or etiquette (cf. Sec. VII below).

motivations with respect to X and Y, this normally means that there has been no (relevant) change in my moral evaluations of X and Y. Of course, an externalist denies that an agent's motives *must* track her moral beliefs in this way. However, Smith thinks that even the externalist must claim that the motives of a "good and strong-willed person" will reliably track her moral beliefs (p. 72). He thinks that the only way the externalist can explain this is by appeal to a general desire, on the part of the good and strong-willed person, to be moral, whatever that turns out to be. He contrasts this sort of concern for something that is derived from a belief that it is morally required with underived intrinsic concern. He then claims that genuinely good people care about the things they care about—themselves, family, friends, and dependents—intrinsically, not because morality requires this. Adapting one of Williams's criticisms of the perspective that impartial moral theories impose on special relationships, Smith claims that the externalist account of the good and strong-willed person ascribes to her "one thought too many."<sup>24</sup> Smith's idea seems to be that in order to explain how some changes of moral belief produce a change in moral motivation the externalist must filter concern for oneself and one's intimates through the lens of moral requirement. This, he alleges, is incompatible with the intrinsic concern a good person should have for himself and his intimates.

I find this puzzling.<sup>25</sup> It is quite plausible that morality itself enjoins intrinsic concern for oneself and one's intimates. If so, it is unclear to me how a moralized concern and an intrinsic concern for oneself and one's intimates are incompatible. Perhaps the worry is not so much about the compatibility between the demands of morality and those of special concern as about the compatibility of two different kinds of motives. In Williams's example, a man chooses to save the life of his wife when he is faced with a situation that forces him to choose between saving his wife and saving a stranger. Williams thinks that the wife might reasonably expect that her husband's motivating thought would simply be "She's my wife." If the husband rescues his wife only after engaging in scrupulous moral deliberation and concluding that it would be permissible to save his wife, then, Williams claims, he has one thought too many. Anyone (not just his wife) might think it rather pathetic if the husband must go through complex and uncertain moral deliberations prior to acting on his special concern for his wife in the case in question, where it is in fact clearly permissible

24. Compare Bernard Williams, "Persons, Character, and Morality," reprinted in his *Moral Luck* (Cambridge: Cambridge University Press, 1981), pp. 1–19.

25. Smith also formulates this issue in terms of the distinction between *de dicto* and *de re* attitudes. I do not discuss this formulation separately, as I find it puzzling and think it is in any case inessential.

to save his wife. But that doesn't show that it is absurd for one's special attachments to be regulated in an appropriate way by moral concerns. For insofar as moral requirements have rational authority, it would be inappropriate to act on special concern for his wife where this was morally impermissible. For instance, presumably it would not be appropriate to save the life of one's wife if this required killing five innocent people. It's not just that morality could limit legitimate forms of special concern in highly unusual circumstances; presumably, it already does constrain the forms of special concern that we think appropriate. Moral beliefs can play this regulative role vis-à-vis special concern without always crowding out the agent's deliberative and motivational field. It need only be true of the agent that he would not have acted on his special concern had he believed doing so was morally impermissible and that he would have engaged in deliberation about whether to act on his special concern if some feature of the situation had seemed to call into question the permissibility of acting on his special concern.<sup>26</sup>

I take something like this to be the right response to the similar incompatibility some see between Kant's insistence that virtuous agents act from the motive of duty and the expression of natural sentiments of benevolence and special concern. Acting from a sense of duty does not require contramoral sentiments and does not preclude actions that express natural sentiments of benevolence and special concern. Rather, a good will requires that the motive of duty be sufficient, and this condition can be spelled out in the counterfactual condition of having one's actions regulated by one's moral beliefs. Action expressing special concern is regulated by the agent's sense of duty when it is true that she would not have had the special concern or at least would not have acted on it had she believed the action was impermissible.<sup>27</sup>

Perhaps Smith's protest is that special concern for one's intimates that is regulated by a desire to do what is morally required is valued for the sake of something else and, hence, could have only instrumental value.<sup>28</sup> He might reasonably object that the sort of special concern that we think appropriate for intimates is noninstrumental. But this argument seems suspect in two ways. First, it assumes that, when concern for one's intimates is regulated by one's moral beliefs and a

26. Compare Scheffler's useful discussion of Williams's worries about how impartial moral theories lead to an overmoralized self (chap. 3).

27. For a somewhat different defense of the Kantian conception of the motive of duty, see Barbara Herman, "On the Value of Acting from the Motive of Duty," in her *The Practice of Moral Judgment* (Cambridge, Mass.: Harvard University Press, 1993), pp. 1–22.

28. This interpretation was suggested by Smith's reply to me at the APA symposium.

desire to be moral, the first is desired for the sake of the second. But regulation need not involve this. The externalist needn't suppose that one cares about intimates in order to be moral; rather, when special concern is morally regulated, one cares about one's intimates only in ways that are morally permissible. Here, morality acts as a constraint or filter on emotions, not as a goal at which emotions are directed. Second, even if the externalist did have to think of morality as the ultimate goal by which special concern was regulated it wouldn't follow that special concern could only be instrumental. The protest assumes that if one values X for the sake of Y, then one values X only instrumentally. But this assumption is not compelling. Where X and Y are distinct, and X is valued for the sake of Y, X's value is derivative. Where X is valued as a causal means to Y, then X's value is derivative and instrumental. But X can be a proper part of Y, even though X and Y are distinct. If so, one can value X for the sake of Y without valuing X only instrumentally; X will have intrinsic but *contributory* value.<sup>29</sup> But then the externalist could hold that special concern has derivative but intrinsic value. For he could maintain that morally required forms of special concern are to be valued because they are morally required. It is not that these forms of special concern causally bring about dutiful action; it is, rather, that they are part of one's duty. So that if one were to value one's intimates for the sake of what's morally required, one would be according special concern contributory value that is intrinsic, not instrumental, value.

If this is right, then I see no good reason for supposing that the sort of moral motivation that an externalist might ascribe to the good and strong-willed person is incompatible with her having and acting on the basis of special concern for herself and others. But then none of the objections Smith raises to the externalist account of moral motivation seems compelling. Fortunately, as far as I can see, few, if any, of the other interesting and important things he has to say about the semantics of moral discourse (his defense of cognitivism), the nature

29. Compare C. I. Lewis, *An Analysis of Knowledge and Valuation* (La Salle, Ill.: Open Court, 1946), chap. 16. A similar idea is expressed in Aristotle's discussion of complete goods in book 1 of the *Nicomachean Ethics* (1094a18–19, 1097a27–b6). A good is complete if it is chosen for its own sake; it is unconditionally complete if other things are chosen for its sake, and it is not chosen for the sake of something else. Aristotle believes that *eudaimonia* is the only unconditionally complete good; all other goods are chosen for its sake. Some goods chosen for the sake of *eudaimonia*, though not choiceworthy in themselves, are choiceworthy as a causal means to some ingredient of *eudaimonia*; these goods are incomplete, instrumental goods. But other goods—such as the virtues—that are chosen for the sake of *eudaimonia* are choiceworthy in themselves. They are chosen for the sake of *eudaimonia* because they are constituent parts of *eudaimonia*. Such goods are complete or intrinsic goods, though they are not unconditionally complete goods.

of intentional action (his argument for the role of pro-attitudes in motivation), or the connection between normative belief and motivation (his acceptance of weak normative internalism) depend on his case against externalism about moral motivation.

## VII. EXTERNALISM AND THE AUTHORITY OF MORALITY

Does the externalist solution to the puzzle about moral motivation succeed in reconciling intellectual and practical dimensions of morality, or does it force us to abandon the practical dimension of morality? That depends on how the practical dimension of morality is conceived. If it is seen as committed to internalism about moral motivation, then externalism does force us to abandon the practical dimension of morality. But I see no reason to insist that we interpret the practical dimension of morality that way. One way in which morality is practical is that moral judgments are normally motivationally efficacious. On my view, as well as Smith's, moral motivation is predicated on assumptions about the rational authority of moral requirements. Whereas Smith treats these assumptions as supported by conceptual truths, I treat them as defeasible assumptions that require substantive defense. This suggests a more fundamental way in which morality might be practical—it might have rational authority. Morality will be practical in this sense just insofar as its requirements have rational authority. Whereas the motivational efficacy of moral judgment reflects the fact that many of us regard moral requirements as authoritative, the principled amoralist shows us that this assumption is open to challenge.

If I am right, issues about moral motivation, important as they are, have a kind of subsidiary importance. For moral motivation depends upon our beliefs about the rational authority of morality. If so, we need to tackle head-on issues about the rational authority of morality and, in particular, substantive issues about the nature and demands of morality and of practical reason and how they are related to each other.

The externalist about moral motivation can vindicate the practical dimension of morality insofar as she can answer the (principled) amoralist challenge and defend the authority of morality. In order to see ways in which an externalist might defend the authority of morality, it might help to see the antirationalist argument we've already considered (Sec. V) as one solution to another puzzle of moral psychology, this one about the authority of morality. This puzzle also involves a quartet of claims, each of which has struck some people as plausible.

1. Moral requirements include impartial other-regarding obligations that do not apply to agents in virtue of their aims or interests.
2. Moral requirements necessarily provide agents with reasons for action.



3. Rational action is action that advances the agent's aims or interests.
4. Fulfilling other-regarding obligations need not advance the agent's aims or interests.

Claim (1) articulates one conception of ethical objectivity, familiar from Kant, according to which moral requirements appear as impartial constraints on conduct that do not apply to agents in virtue of their aims and interests. For instance, I do not defeat an ascription of obligation to me to help another by pointing out that doing so will serve no aim or interest that I have. Claim (2) expresses a rationalist thesis about the authority of morality. Claim (3) expresses a common view of practical rationality, according to which it is instrumental or prudential. Though prudential and instrumental conceptions of rationality are different, both represent the rationality of other-regarding conduct as derivative. By contrast, practical reason is impartial if it implies that there is a nonderivative reason to engage in other-regarding conduct. Claim (4) reflects a common assumption about the independence of different people's interests and attitudes. Though agents often do care about the welfare of others and there are often connections between an agent's own interests and those of others, neither connection holds either universally or necessarily.

Though each element of the puzzle might seem appealing and has appealed to some, not all four claims can be true. In fact, a number of influential historical and contemporary views can be seen as responses, perhaps tacit, to this puzzle that reject at least one element of the puzzle on the strength of others. Some moral relativists and minimalists reject the existence of impartial moral norms asserted in (1); they claim that genuine moral requirements must be relativized to and further the agent's interests or aims in some way. Antirationalists deny (2) and claim that failure to act on moral requirements is not necessarily irrational. Others, such as Kant, reject both instrumental and prudential conceptions of practical reason in (3) and defend the existence of impartial practical reason. Finally, metaphysical egoists reject (4) and resolve the puzzle by arguing that, properly understood, people's interests are interdependent such that acting on other-regarding moral requirements is a counterfactually reliable way of promoting the agent's own interests.

My own view is to reject relativist and minimalist accounts of the authority of morality as too revisionary but to pool resources to be found in the other three solutions. Unfortunately, there isn't time to sketch, much less defend, these claims. However, it might be worth suggesting how even some versions of the antirationalist solution can attribute significant authority to morality. The antirationalist understands morality and rationality as two independent points of view. One such view represents morality as impartial but rationality as instrumen-



tal or prudential. If so, we might deny that immoral action is necessarily irrational and reject (2). On this view, the authority, but not the scope or content, of morality depends on the aims or interests of agents. If we rely on purely instrumental assumptions about rationality, we can establish the authority of other-regarding moral requirements to those who already have suitable other-regarding attitudes. But the instrumental justification of morality is hostage to these attitudes; it seems unable to explain why those who lack these attitudes should cultivate them or why those who do have them should maintain them. The more traditional defense of morality is to argue that the demands of morality and enlightened self-interest coincide. It seems plausible that the general compliance with familiar other-regarding moral norms of restraint, cooperation, and aid is mutually advantageous. Though each would be better-off if others comply while she does not, the compliance of others is generally conditional on her own. If so, the way to enjoy the benefits of others' compliance is to be compliant oneself. Moreover, because each has an interest in others' cooperation and restraint, communities will tend to reinforce compliant behavior and discourage noncompliant behavior through education and social pressure. If so, the antirationalist can offer a generally reliable but contingent justification of the demands of morality. The coincidence between morality and self-interest, on this view, is probably imperfect and certainly counterfactually unstable, with the result that immorality need not always be irrational. Whereas this antirationalist position allows that immoral action need not always be irrational and so explains why it is possible for someone to make a moral judgment and remain unmoved, it would explain why everyone has some stake in morality, and why people generally have reason to behave morally. As long as we have not tied the content of morality to its rationality, we can reproach the immoralist with immorality. What is lost if we cannot also always reproach him with irrationality?

I would like to, and believe we can, defend a more robust rationalist interpretation of the authority of morality than this, by defending the existence of impartial practical reason and developing the resources of metaphysical egoism.<sup>30</sup> But notice that even this sort of antirationalist can offer an interpretation of how morality has authority and an important practical dimension. If this is the worst we can do, there is little reason to think that the externalist about moral motivation must simply abandon the practical dimension of morality.

30. See my "Kantian Rationalism: Inescapability, Authority, and Supremacy," in *Ethics and Practical Reason*, ed. G. Cullity and B. Gaut (Oxford: Oxford University Press, 1997), and "Self-Love and Altruism," *Social Philosophy and Policy* 14 (1997): 122–57.