



Moral Cognitivism and Motivation

Sigrun Svavarsdottir

The Philosophical Review, Vol. 108, No. 2. (Apr., 1999), pp. 161-219.

Stable URL:

<http://links.jstor.org/sici?sici=0031-8108%28199904%29108%3A2%3C161%3AMCAM%3E2.0.CO%3B2-S>

The Philosophical Review is currently published by Cornell University.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/sageschool.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

Moral Cognitivism and Motivation

Sigrún Svavarsdóttir

The impact moral judgments have on our deliberations and actions seems to vary a great deal. Moral judgments play a large part in the lives of some people, who are apt not only to make them, but also to be guided by them in the sense that they tend to pursue what they judge to be of moral value, and shun what they judge to be of moral disvalue. But it seems unrealistic to claim that moral judgments play a pervasive role in the lives of all or even most people. There are considerable variations in how strong a tendency people have to think in moral terms, and in how such thoughts affect their decisions and actions. For every moral hero who single-mindedly pursues moral values, there are thousands of less committed people who only do so when it does not cost them too much in material comfort, personal relations, or social standing. And of course, what counts as too much varies from person to person. On top of such variations, there are those who consistently display moral indifference—people who concede, for example, that certain investment policies have morally problematic consequences, but who can readily and without compunction ignore that in their business decisions. There even seem to be moral subversives, people who intentionally and knowingly pursue what they acknowledge to be morally wrong or bad, and do so for that very reason.

Such variations in moral motivation—motivation by moral judgments—strike me as constituting good reasons for thinking, first, that moral judgments need to be supplemented by a distinct conative state (desire in the broadest sense of that term) in order to play a motivational role; second, that this conative attitude is *not*

Material from versions of this paper was presented at the CUNY Graduate Center's Colloquium Series (October 11, 1995); to the New York Society for Philosophy and Public Affairs (April 19, 1996); and at the Rutgers University Colloquium Series (April 30, 1998). I thank each of these audiences for helpful comments and questions. Special thanks are due to Paul Boghossian, Ruth Chang, Carsten Hansen, Frances Kamm, Barry Loewer, Thomas Nagel, Derek Parfit, Peter Railton, John Richardson, Stephen Schiffer, and especially to anonymous referees for the *Philosophical Review*.

necessarily present in those who make moral judgments; and, third, that its motivational strength varies from person to person. Of course, the conceptual content of the judgment and the conative state have to be appropriately related: the judgment has to represent the object of evaluation in a way that engages the conative attitude in question. This suggests that making a moral judgment is a matter of conceiving of something as being a certain way and that this way of conceiving of an object is no more dependent on the motivational states of the thinker than is conceiving, say, of the day as sunny; nor need it affect the motivational states of the thinker any more than the latter judgment does. But although an agent need not be guided in the sense of being motivated by his moral judgment, moral judgments—unlike judgments about the weather—are in some sense invariably action-guiding: for any type of moral judgment, there are fairly determinate ways in which an agent, irrespective of his ends, may act *in accord with* or *contrary to* it. Thus, even when the moral subversive sets the house on fire because she believes it is evil, she acts in some sense contrary to the judgment that it is evil to put the house on fire.¹

These observations provide, I submit, a good starting point for an inquiry into the nature of moral judgments. In light of them, it seems reasonable to adopt the working hypothesis that moral thought involves conceptual resources employed in the formation both of cognitive states (belief, doubt, worry, etc.) and conative and emotional states (desire, hope, wish, anger, relief, etc.). The challenge is to understand what is distinctive about these conceptual resources and especially in what way they are more intimately bound up with the regulation of conduct than are those involved in conveying information, say, about climatic conditions, given that judgments about both moral and climatic matters motivate only when they engage a desire with appropriately related content.

In taking this approach, I am rejecting the *internalist* thesis that there is a necessary connection between moral judgment and motivation, a thesis that is widely accepted as a constraint of adequacy on accounts of moral judgments. The main goal of the present paper is to undermine this constraint. This task is undertaken in sections 3–6. A secondary goal—tackled in sections 6–7—is to allay

¹Here I have in mind Dostoyevsky's fictional character Lisa (the little she-devil) in *The Brothers of Karamazov*.

worries about tracing moral motivation to a desire, and more specifically to a desire formed by employing distinct moral concepts. The first two sections are devoted to preliminaries.

1. Motivational Internalism

The internalist thesis in question makes a claim about the connection between the mental act of making a moral judgment and motivation. *Motivational internalism* seems, therefore, an appropriate label; and *motivational externalism* will serve as a name for the opposing view.² In its strongest form, motivational internalism states that moral judgments are “intrinsically” motivating; in other words, they motivate on their own rather than in collaboration with a distinct conative state. Although the label ‘internalism’ suggests this strong version, it is more in line with the existing literature to formulate the thesis as claiming that moral judgments are *necessarily connected* to motivation to pursue or promote what is judged favorably and to shun or prevent what is judged unfavorably. Thus formulated, motivational internalism is compatible with the view that the motivation imported by a moral judgment has its roots in a distinct conative state that is, however, necessarily connected to the moral judgment. The necessity at stake is supposed to be conceptual necessity. Presumably, the idea is that the ascription conditions for moral judgments are such that an agent could not be considered to have made a specific moral judgment unless he were motivated in a specific way—or, some might prefer, the possession conditions for moral concepts are such that an agent could not sincerely and competently apply them without being motivated in a specific way.

Some further modifications of the thesis have been proposed in response to Michael Stocker’s observation that under conditions

²For useful general discussions of these positions as well as surrounding issues, see William K. Frankena’s classic article “Obligation and Motivation in Recent Moral Philosophy,” in *Essays in Moral Philosophy*, ed. A. I. Melden (Seattle: University of Washington Press, 1958), 40–81; Stephen L. Darwall, *Impartial Reason* (Ithaca: Cornell University Press, 1983), chap. 5; David O. Brink, *Moral Realism and the Foundations of Ethics* (Cambridge: Cambridge University Press, 1989), 37–80; Michael Smith, *The Moral Problem* (Oxford: Basil Blackwell, 1994). In Darwall’s terminology, the thesis under discussion is *judgment internalism about moral judgments*; in Brink’s terminology, it is *appraiser internalism about motives*.

of deep depression, severe cases of weakness of will, and other maladies of the spirit, the connection between moral judgment and motivation is often broken, even in individuals who are normally motivated by their moral judgments.³ Michael Smith, for example, has urged that motivational internalists recognize that the connection between moral judgment and motivation is a defeasible one. In the formulation of his “practicality requirement on moral judgment”—which is a close relative of motivational internalism—he adds the important qualification that the connection between moral judgment and motivation holds only *in (practically) rational agents*.⁴

I am not inclined to resist a modification of the internalist thesis meant to block Stocker’s counterexamples. However, the introduction of the practical rationality condition opens up a can of worms. The question becomes, What is built into that constraint? That is to say, what sort of notion of practical rationality is being used? We need a more informative formulation of this condition before evaluating an internalist thesis that incorporates it. Indeed, motivational internalists should be very cautious in accepting this vague qualification of their thesis if they want to use the internalist thesis as a constraint on accounts of moral thought. For once the rationality condition is introduced, it may be argued that the internalist constraint has no bearing on what sort of mental act a moral judgment is, but rather reveals some of the norms or requirements of

³Michael Stocker, “Desiring the Bad: An Essay in Moral Psychology,” *Journal of Philosophy* 76 (1979): 738–53.

⁴Smith, *The Moral Problem*, sect. 3.1. Besides introducing the rationality condition, Smith’s statement of the practicality requirement differs from my statement of motivational internalism in two ways. First, Smith does not use any modal notion in stating the content of the practicality requirement. However, he claims that the requirement has the status of a conceptual platitude. But presumably, if it is a conceptual truth that a certain connection holds between moral judgment and motivation, then that connection holds in every conceptually possible world; that is, the connection holds of conceptual necessity. Second, Smith’s practicality requirement specifies that the connection in question holds only between moral judgments *about the behavioral options of the agent making the judgment* and motivation. I do not restrict motivational internalism in this way because it seems artificially to weaken its putative support for noncognitivism (see next section). But such a restriction would not affect my arguments. In his development of his practicality requirement, Smith is influenced by Christine Korsgaard’s work on practical reason. See her “Skepticism about Practical Reason,” *Journal of Philosophy* 83 (1986): 5–25.

rationality.⁵ Instead, internalists might simply attach an exception clause for agents suffering from motivational disorders that affect them more generally. Incorporating this qualification into my statement of motivational internalism, the thesis under scrutiny is that *moral judgments are of conceptual necessity connected to motivation to pursue or promote what is judged favorably and to shun or prevent what is judged unfavorably, except in individuals suffering from motivational disorders that affect them more generally.*^{6,7} (From now on, I will not

⁵Maybe these two issues cannot be pried apart. Maybe one cannot plausibly argue that the internalist constraint is or reflects a requirement of rationality unless one gives an account of moral judgments that reveals why that would be the case. Smith at least seems to think that these two issues go together; he gives an account of the content of moral judgments that, he argues, makes sense of why the practicality requirement is a conceptual truth (see note 61, below). Still, I counsel caution in formulating the qualification in terms of conditions of practical rationality unless, of course, a more informative formulation of it is forthcoming. Needless to say, if conditions of rationality are just conditions under which the agent's conduct is interpretable as an intentional action, then I have no qualms about incorporating it into the internalist thesis. However, such an addendum is unnecessary, since it adds nothing of substance to the thesis.

⁶Notice that as I understand motivational internalism, it makes a claim about the connection between motivation and *moral* judgments, rather than evaluative or normative judgments in general. Some would advocate an internalist constraint on all evaluative or normative judgments. Obviously, I am contesting both that moral judgments are necessarily connected to motivation and that the distinguishing mark of evaluative or normative judgments in general is their connection to motivation. But I would like to leave it open (though I remain doubtful) whether there is a subcategory of evaluative or normative judgments for which the internalist constraint holds. Now, it is of course notoriously difficult to say what distinguishes moral judgments from other evaluative or normative judgments. If pressed, I would gesture toward such factors as their "aspiring" to some sort of impartiality and being vaguely "concerned with" human well-being. But here I will have to rely on an intuitive understanding of moral judgments' being distinct from prudential, aesthetic, and other categories of evaluation. Thanks are due to John Richardson for pressing me to make this clarification.

⁷James Dreier offers a weaker internalist thesis, *modest internalism*, as a constraint of adequacy on accounts of moral language. See his "Internalism and Speaker Relativism," *Ethics* 101 (1990): 6–26. Modest internalism adds to motivational internalism the condition that the moral judge be "in normal context." This proposal need not be taken seriously if "normal contexts" are specified merely as "those circumstances under which a person is motivated by his moral judgments." Dreier recognizes this, but admits that he does not know how to specify rigorously the independent conception of normality that he needs. In the absence of such specifica-

explicitly state the exception clause, but it should be understood as implicit.)

Notice that motivational internalism, as stated, does not claim that the motivation necessarily accompanying moral judgments overrides all other motivation. Indeed, the position is silent on the strength of the accompanying motivation, and even allows that its strength varies from agent to agent.⁸ Internalists will, therefore, acknowledge most of my observations about variations in moral motivation. However, they have to contest my claim that some people, not suffering from general motivational disorders, are unmoved by their moral judgments or are even moved by them to pursue evil and do wrong. This is supposed to be a conceptual impossibility.

The internalist thesis is often invoked in the debate between noncognitivists and cognitivists about moral thought and language. *Cognitivism*, as understood here, maintains that moral judgments employ representational resources, expressible in the moral vocabulary of our public language. Moral judgments are related to moral beliefs in the standard way that beliefs and judgments with the same content are related. Both moral judgments and beliefs, as well as the sentences expressing them, are truth-evaluable.⁹ In con-

tion, the condition is too obscure for modest internalism to play an important role in deciding between competing accounts of moral thought and language. For further remarks on Dreier's view, see note 30, below.

⁸Richard W. Miller suggests that it is a condition on the ascription of moral judgments that the agent be motivated by her cognizance of moral considerations when it has no cost to herself. See his *Moral Differences: Truth, Justice and Conscience in a World of Conflict* (Princeton: Princeton University Press, 1992), 95. I am inclined to read him as accepting motivational internalism as formulated in the text: moral judgments necessarily import some motivation, however weak and easily overridden it may be in situations that involve some cost to the agent (see, however, note 32, below). But Miller can also be read as accepting a weaker version of motivational internalism: moral judgments need only motivate in circumstances when the moral judge perceives no cost to himself. This weaker version also falls within the target of my argument in subsequent sections.

⁹Notice that I have formulated cognitivism so that it does not involve any ontic commitment to moral facts or properties. It is a view about the nature of moral thought and the semantics of moral language. Prima facie, a cognitivist could be either a realist or an error theorist about moral thought and discourse. The cognitivist camp includes thinkers as diverse as Richard N. Boyd, "How to be a Moral Realist," in *Essays on Moral Realism*, ed. Geoffrey Sayre-McCord (Ithaca: Cornell University Press, 1988), 181–228; Brink, *Moral Realism*; John L. Mackie, *Ethics: Inventing Right and Wrong* (New York: Penguin, 1977); John McDowell, "Projection and Truth in Eth-

trast, *noncognitivism*, as understood here, maintains that moral judgments manifest some sort of conative attitudes taken towards the object of evaluation under a nonmoral mode of presentation.¹⁰ The moral terms of the public language are not representational devices; instead, they function semantically as mood indicators that signal a grammatical mood, employed to express these sorts of conative attitudes. In spite of our frequent use of the truth predicate in moral discourse, neither moral judgments nor moral sentences provide truth-evaluable representations of the object of evaluation.¹¹ It should be fairly obvious that the internalist constraint

ics," *The Lindley Lecture* (University of Kansas: Lawrence, 1988); G. E. Moore, *Principia Ethica* (Cambridge: Cambridge University Press, 1903); Peter Railton, "Moral Realism," *Philosophical Review* 95 (1986): 163–207; Nicholas L. Sturgeon, "Moral Explanations," in *Morality, Reason and Truth* ed. David Copp and David Zimmerman (Totowa: Rowman and Allenheld, 1985), 49–78; David Wiggins, "A Sensible Subjectivism?" in his *Needs, Values, Truth: Essays in the Philosophy of Value* (Oxford: Blackwell, 1987), 185–214; Smith, *The Moral Problem*; David Copp, "Moral Obligation and Moral Motivation," in *New Essays on Metaethics*, ed. J. Couture and K. Nielson (*Canadian Journal of Philosophy*, supp. vol., 1996); and Dreier, "Speaker Relativism." The last three thinkers, unlike the rest, accept reductivism about moral concepts: they believe that an adequate analysis can be given of moral concepts. However, all three use concepts from other parts of normative discourse in the analyses. Reductive (or analytic) cognitivism should not be equated with reductivism about moral properties or facts, which is, for example, accepted by Railton.

¹⁰'Moral judgment' has become a term of art in the metaethical literature. It is used to refer to the mental and speech acts central to moral evaluation, whatever their nature may be. Since I am concentrating on moral motivation in this paper, I will be mostly using it to refer to the relevant mental acts. It seems more accurate to think of judgments as mental acts rather than mental states, although they are, of course, the onsets, expressions, or activations of mental states. Perhaps they should be equated with occurrent mental states. Nevertheless, I distinguish between moral judgments and the corresponding mental states in this paper, and talk about the former as manifesting the latter.

¹¹I have been careful to formulate the noncognitivist position so that it encompasses not only the emotivist view of A. J. Ayer (see *Language, Truth, and Logic* (London: Gollancz, 1946), chap. 6) and Charles L. Stevenson (see "The Emotive Meaning of Ethical Terms," *Mind* 46 (1937): 14–31), and R. M. Hare's prescriptivism (see *The Language of Morals* (Oxford: Oxford University Press, Clarendon Press, 1952)), but also Simon Blackburn's quasi-realism (see *Spreading the Word* (Oxford: Oxford University Press, 1984), chap. 6, and *Essays in Quasi-Realism* (Oxford: Oxford University Press, 1993)) and Allan Gibbard's norm-expressivism (see *Wise Choices, Apt Feelings: A Theory of Normative Judgment* (Cambridge: Harvard University Press, 1990)).

prima facie favors noncognitivism: if moral judgments manifest conative states of mind, they will necessarily have some motivational force, whereas if they manifest beliefs, it needs to be explained why they would—unlike most beliefs—be necessarily motivating. If the internalist constraint is accepted, the burden is on the cognitivists to show that they can meet it.¹²

Indeed, the internalist constraint is often touted as providing conclusive rather than only prima facie support for noncognitivism. But as several cognitivists have pointed out, it can provide conclusive support for noncognitivism only if it is supplemented by the thesis that the motivating power of beliefs invariably relies on the contingent presence of a desire whose content is appropriately related to the belief's content. Many of these cognitivists actually accept the internalist constraint, but reject the Humean view that motivation is always rooted in a desire and maintain that moral judgments manifest motivating beliefs: "besires."¹³ Some may even

¹²All but Ayer invoke the internalist constraint in favor of their view, but it is most explicitly done by Stevenson ("The Emotive Meaning," 13)—cf. his magnetism requirement—and Blackburn (*Spreading the Word*, 188). However, in a 1991 article ("Just Causes," reprinted in *Essays in Quasi-Realism*, 198–209), Blackburn modifies his position in such a way that one may wonder whether he remains a motivational internalist. When responding to an objection by Nicholas Sturgeon (see note 41, below), Blackburn seems to grant that Socrates and Thrasymachus (in Plato's *Republic* I) are both making judgments about justice, although only Socrates is motivationally affected by them. But Blackburn has not really given up on motivational internalism; for he holds that Satan can want evil to be his good only so long as he attaches some negative emotion to what he deems evil: "Milton's Satan can be represented as wanting to make evil his good because of his actual self-disgust, indeed his suffering, at being forced (by the need to have something over which to reign) to make such a choice. . . . But if his strategy is successful, and he succeeds in driving out not only remorse but all the other elements that enable him to see his own plight as desperate, then indeed the interpretation would start to waver. He would no longer be *seeing* evil as his good, but merely *doing* evil and seeing it as good" (201). And Blackburn is quite ambivalent about whether Thrasymachus should, indeed, be interpreted as making judgments about justice: "Thrasymachus himself faces this peril [of being best reinterpreted]: if he sticks to the view that justice is what is in the ruler's interest, perhaps he is not best seen as discussing justice at all. Add the quirk that he is contemptuous of whatever it is that he is thinking of, and the difficulty increases" (199).

¹³John McDowell is probably the most prominent cognitivist and internalist who reconciles these two views in this manner. See his "Are Moral Requirements Hypothetical Imperatives?" *Proceedings of the Aristotelian Society*, supp. vol. 52 (1978): 13–29, and his "Virtue and Reason," *Monist* 62

attempt to reconcile the internalist constraint, the Humean view of motivation, and moral cognitivism. It might, for example, be argued that although moral judgments are belief-like and therefore—given the Humean thesis—motivationally inert *on their own*, their content is such that they could not be made in the absence of a desire that provided the source of moral motivation.¹⁴

My working hypothesis that moral thought involves conceptual resources employed in the formation of cognitive, conative, and emotional states clearly puts me in camp with cognitivists.¹⁵ However, I am *not* inclined to take on the burden of explaining how moral cognitivism is to be reconciled with internalism, since I am of a firm externalist conviction and—as explained in the introduction—believe that it helps to provide an interesting starting point for an inquiry into the nature of moral judgments. Instead, I will seek to undermine the internalist thesis as a constraint of adequacy on accounts of moral judgments. I am not alone in accepting motivational externalism. In the recent literature, David Brink and Peter Railton have, for example, committed themselves to this externalist position. However, my approach to moral motivation is markedly different from theirs.

Both Brink and Railton trace the source of moral motivation to

(1979): 331–50. See also M. Platts, *Ways of Meaning* (London: Routledge and Kegan Paul, 1979). And for an argument against the Humean view of motivation, see T. Nagel, *The Possibility of Altruism* (Oxford: Oxford University Press, Clarendon Press, 1970). David Lewis, who is a critic of the view, invented the convenient term ‘besire’ to refer to this unique kind of mental attitude that has both an assertoric and a motivational force. See his “Desire as Belief,” *Mind* 97 (1988): 323–32.

¹⁴For a simple illustration, take a subjective naturalist who claims that the belief that *a* is good amounts to the belief that *a* meets one’s approval. The subjectivist may attempt to demonstrate that it is metaphysically or even conceptually impossible to have a belief with such content without actually approving of *a*, which is a type of desire within a Humean framework. Relying on a different analysis of the content of moral beliefs, Michael Smith attempts a reconciliation between cognitivism, the Humean view of motivation, and the practicality requirement (see note 61, below). Miller (*Moral Differences*) attempts a reconciliation along entirely different lines (see note 32, below).

¹⁵I am, moreover, inclined toward nonreductive or nonanalytic cognitivism (see note 9). Notice that the unanalyzability of moral concepts does not preclude the possibility that their semantic values are properties for which we also have (or could develop) other, nonmoral, concepts. I prefer to remain noncommittal on this point.

a conative attitude whose conceptual content does not involve moral concepts. Brink suggests that the motivational source depends on the agent's substantive moral views. For example, sympathy for others is the source of moral motivation for those who accept other-regarding moral principles.¹⁶ Railton, on the other hand, suggests that a desire to be able to justify one's conduct from a general or impartial standpoint, rather than merely a personal standpoint, is what drives people to engage in moral evaluation and conduct themselves accordingly.¹⁷ Although I find both suggestions illuminating, I would not want to adopt either one as a general account of moral motivation. I have a much more simple-minded proposal: the disposition to be motivated by one's moral judgments is grounded in a conative attitude (desire) taken towards objects under a moral mode of presentation.¹⁸ For convenience, call it *the desire to be moral*. I expect that it varies from agent to agent which other mental states sustain that desire. In some cases, it may be sustained by a healthy dose of sympathy with others, as well as an acceptance of a norm of benevolence. In other cases it will, instead or additionally, be sustained by a desire to be able to justify one's conduct from an impartial standpoint. But undoubtedly, there are also cases in which it is sustained, say, by an awe of God and a theological view of the foundations of morality, or by a fear of punishment. It may also just stand there pretty much on its own, not dependent on any other motive at that particular stage in the person's life. I doubt there is any unique psychology that sustains a disposition to be motivated by one's moral judgments. And the project of accounting for motivation involving moral judgments should not be confused with the project of characterizing the mo-

¹⁶Brink, *Moral Realism*, 49.

¹⁷Railton, "Moral Realism," 202.

¹⁸My view on moral motivation is closest to W. D. Ross's and H. A. Prichard's, but this does not mean that I accept their metaphysical and epistemological views. In his 1928 article "Duty and Interest" (reprinted in *Readings in Ethical Theory*, ed. W. Sellars and J. Hospers (New York: Appleton Century Crofts, 1952)), Prichard gives up his earlier view of moral motivation and identifies the desire to do what is right as the source of moral motivation. Ross subscribes to the same view in his *The Right and the Good* (New York: Oxford University Press, 1930). I identify the desire as a conative attitude taken towards an object under *some* moral mode of presentation; it need not involve the concept of rightness. I also depart from Prichard and Ross in that I do not take this desire to be universal or somehow inherent in human nature.

tivational structure of the morally ideal or reflectively sophisticated agent. It is the former subject that I have been broaching.

This view of moral motivation seems to attract criticism that goes beyond the internalist objection. For example, David Copp, who embraces motivational externalism, refers to such accounts of moral motivation as “crude moral psychology.”¹⁹ Of course, my view is a crude psychology in that it gives us the most superficial understanding of why someone is motivated on account of his moral judgments, but I hope to convince the reader that it is a plausible account of motivation by moral judgments. This will be done mostly in section 6 in the context of responding to Michael Smith’s recent argument against motivational externalism, which trades on the claim that externalists are committed to a hopelessly wrong-headed view of moral motivation—namely, the one to which I have just committed myself. But first we need to get some further preliminaries out of the way and discredit motivational internalism.

2. Other Internalisms

In order to bar some misunderstandings about the intended scope of my arguments, this section distinguishes motivational internalism from some other related theses that have gone under the internalist rubric, but that are not under discussion in this paper. First of all, motivational internalism has to be distinguished from the thesis that in order for a moral judgment to be appropriately applied to the conduct of an individual, that individual has to be motivated under specific conditions to undertake (or avoid) the action positively (or negatively) appraised. Let’s call this latter thesis *application internalism*, since it is about the conditions of application of a moral judgment to an individual.²⁰ In contrast, moti-

¹⁹See Copp, “Moral Obligation.” On Copp’s own view, moral motivation has its source in an intention (or a policy) to conform, and support conformity, to certain standards. The relation between this intention and the agent’s moral judgments is that the judgments imply that the standards in question are justified and call for (or prohibit) the action under evaluation. His account thus presupposes a partial analysis of the content of moral judgments; mine doesn’t.

²⁰This distinction is closely related to Stephen Darwall’s distinction between judgment and existence internalism about moral judgments (*Impartial Reason*, 54). (Brink also makes a similar distinction; in his terminology, this is the distinction between appraiser and agent internalism about motives (*Moral Realism*, 40).) However, my distinction is drawn at the level of

vational internalism advances a claim about the motivational role of a moral judgment *in an individual who makes it*. I limit my attention to motivational internalism in this paper because I am primarily interested in undermining a certain argumentative strategy for noncognitivism and cognitivist views that link the exercise of moral concepts to motivational responses *of the thinker*. But I also find it a much more compelling and interesting thesis than application internalism. It seems wildly implausible to claim that it is inappropriate or somehow odd, say, to judge it wrong of an individual to sexually abuse children just on the grounds that he is not (either under the actual or some non-normatively specified counterfactual conditions) motivated to refrain from such conduct. Moral judgments are unconditional exactly in that they apply to agents regardless of what their motivations are.

It is important to appreciate that it is motivational internalism rather than application internalism that *prima facie* lends support to noncognitivism. Given the assumption that we have an infallible access to our own mental states, application internalism implies that an agent will not make a sincere and competent moral judgment about *himself* unless he has the appropriate motivation. This may be thought to link application internalism to noncognitivism, since some proponents of the latter have emphasized the connection between motivation and moral evaluation only insofar as the latter applies to one's own conduct. Richard Hare, for example, cashes out the prescriptivity of moral judgments ultimately in terms of its being impossible to assent sincerely to a (positive) evaluation of an action without actually undertaking the same type of conduct, when feasible.²¹ And Allan Gibbard seems only to acknowledge the

thought and concepts, whereas Darwall's is drawn at the level of reference. Application internalism pertains to the application conditions of moral judgments or concepts to an individual, whereas existence internalism pertains to the conditions for a moral obligation or value existing or being instantiated. *Prima facie*, my distinction leaves it open to noncognitivists and other antirealists to accept application internalism. For an example of application and existence internalism, see Gilbert Harman's "Moral Relativism Defended," *Philosophical Review* 85 (1975): 3–22.

²¹To be more exact, Hare maintains that the prescriptivity of a moral judgment consists in its entailing an imperative or command addressed to the person to whom the judgment applies; but it is distinctive of commands that an assent to a command addressed to oneself is sincere only if one abides by it. Notice that Hare maintains that there is a necessary connection between a moral judgment and action—or, translated into the language of motives, between moral judgment and overriding motivation.

constraint that “[w]hen the normative judgments we interpret a person as making at a time apply to that person himself at that very time, he must be normatively motivated.”²² However, all non-cognitivists maintain that moral judgments manifest some sort of conative states.²³ It is hard to see how that could be unless moral judgments had some motivational force regardless of the target of evaluation.

For example, when I judge that a friend is treating his partner unjustly, I am—according to noncognitivists—expressing either some sort of a preference that my friend not treat his partner as he does or some sort of a motivationally charged sentiment against his behavior. Such a preference or sentiment should be manifested in an inclination to interfere, to avoid appearing to condone his behavior, and to show some support to the partner.²⁴ (Of course, none of these inclinations need translate into action; for there might be overriding motivations not to interfere, due perhaps to the moral opinion that it would not be right of me to interfere in the matter). Now, if motivational internalism is right, this consequence of noncognitivism is one of its strengths. Application internalism does not give similar support to noncognitivism. For it does not imply that there is such a pervasive connection between making a moral judgment and being motivated in a certain way.

This is stronger than motivational internalism as stated above. Hare, as far as I know, is alone in holding this stronger thesis. See his *The Language of Morals*, § 2.2 and § 11.2; and his *Moral Thinking: Its Levels, Method and Point* (Oxford: Oxford University Press, Clarendon Press, 1981), § 1.6.

²²Gibbard, *Wise Choices*, 101.

²³Ayer (*Language, Truth, and Logic*), Stevenson (“The Emotive Meaning”), and Blackburn (*Spreading the Word*) maintain that moral judgments are expressive of emotive states: certain motivationally charged sentiments. Hare (*Moral Thinking*) takes them to express preferences of a unique type. While Gibbard (*Wise Choices*) contends that they express the acceptance of norms for certain sentiments, which is a state of a motivational system he refers to as *a normative control system*.

²⁴Admittedly, it is not clear that Gibbard’s view has this upshot. On his view, I am expressing a positive attitude towards my friend’s feeling guilty about treating his partner as he does and towards others’ feeling resentful about his treating the partner as he does. Gibbard thinks that I can have this second-order attitude even if I do not feel any resentment against my friend’s conduct. So maybe he would contend that this second-order attitude need not be manifested in the sort of inclinations mentioned in the text. Still, Gibbard must rely on motivational internalism rather than application internalism for the second reason given in the text.

More importantly, the connection that application internalism claims to hold between making a moral judgment about one's own conduct and being motivated does not in the least suggest that moral judgments manifest some sort of conative states. For the connection is due not to the attitudinal force of the judgment, but rather to the application conditions for moral judgments, as well as the agent's privileged access to his mental states. On this view, moral judgments about the merits of our own options track some of our motivational propensities instead of being motivating themselves. Application internalism, therefore, does not lend any support to noncognitivism.

I would credit Kant with the aforementioned insight that moral judgments are unconditional in that they apply to agents regardless of what their motivations are, although the Kantian formulation of it in terms of the categoricity of moral judgments imports the more controversial thesis that moral judgments give all agents to whom they apply reasons for action regardless of their aims. This rationalist thesis is also sometimes referred to as internalism. Presumably, the thesis is not that any moral judgment has such reason-giving powers, but rather that correct moral judgments do. This thesis speaks of the relation between moral judgments and reasons for action rather than between moral judgments and motivation. However, it is often thought that the existence of a reason for an individual to undertake an action has implications about the motivations of that individual. But even if that were true, this rationalist thesis would have implications about the motivations of the person to whom the moral judgment applies rather than about the motivations of the person who makes the judgment. In other words, it might imply application internalism or a kindred thesis, but it certainly does not imply motivational internalism.²⁵ Thus, for reasons

²⁵Harman ("Moral Relativism") indeed arrives at application internalism in such a way. But a Kantian would presumably resist the implication that the applicability of a moral judgment to an individual is conditioned by his motivational states or capacities. Kantians have attempted to get around this by maintaining that a moral judgment correctly applies to an agent only if he would be motivated to act in a suitable way insofar as he were rational. In that way a triadic connection is established between the correct applicability of a moral judgment to an individual, the individual's having certain reasons for actions, and the individual's having certain motivations under conditions of rationality. See Korsgaard, "Skepticism"; and Smith, *The Moral Problem*.

similar to those given above, I will not discuss this rationalist thesis any further here.

Somewhat less obviously, motivational internalism needs to be distinguished from the rationalist thesis that a moral judgment creates a reason for action for the person who makes it. This distinction would, of course, not be valid, if motivating an individual to act and providing an individual with reasons to act were one and the same thing.²⁶ But I think that, at least initially, these two notions should be kept apart. Similar remarks are in order with respect to a related rationalist thesis that claims that moral judgments amount to, or at least involve, the recognition of a reason for (or against) the action evaluated. Rationalists of either variety may, however, grant the distinction between their thesis and motivational internalism, but still maintain that motivational internalism is implied by their rationalism; for they may accept the auxiliary thesis that either having or recognizing that one has a reason for acting in a certain way implies that one is motivated to undertake that action.²⁷ Indeed, I expect that some will claim that motivational internalism has its appeal only insofar as it is seen as consequent upon one or the other of these rationalist theses. For the purposes of this paper, however, I would like to avoid the thorny issues concerning the nature of reasons for action, and their relation to moral judgments and to motivation. Motivational internalism has certainly been advanced in complete independence of any rationalist thesis and can be evaluated on its own. Should I succeed in casting doubt on the former thesis and should it, indeed, be implied by either one of the above rationalist thesis in conjunction with the auxiliary thesis, then so much the worse for one or both conjuncts.

²⁶For example, Miller (*Moral Differences*) formulates his internalist thesis in terms of a necessary connection between an agent's having moral beliefs and "moral considerations['] . . . provid[ing] her with reasons for choice in costfree situations," but he obviously takes this trivially to imply a necessary connection between moral belief and motivation in costfree situations.

²⁷It seems to me that Nagel, in *The Possibility of Altruism*, is best read as advocating motivational internalism on the grounds of the second rationalist position discussed in this paragraph; of course, this is not to deny that he also accepts the rationalist thesis discussed in the last paragraph.

3. An Argument Against the Internalist Constraint

Motivational externalists typically argue their case by giving a counterexample to the internalist thesis: they sketch an actual or fictional case of a person described as being adept at making moral judgments but entirely unmoved by them, although not suffering from any general motivational disorder.²⁸ The problem with such counterexamples is that many people cannot recognize them as such: descriptions of such cases strike them as incoherent. Here intuitions conflict, leading to yet another apparently intractable philosophical disagreement. Externalists and internalists may, therefore, seem to have equal burden to explain the contrary intuition away or support their position with a positive argument.²⁹ My aim in this section is to show that the burden of argument lies squarely on the internalists' shoulders and that consequently the internalist thesis cannot be invoked as a constraint of adequacy on accounts of moral judgments.

Let's start with an example:

The Example of Patrick: Virginia has put her social position at risk to help a politically persecuted stranger because she thinks it is the right thing to do. Later she meets Patrick, who could, without any apparent risk to himself, similarly help a politically persecuted stranger, but who has made no attempt to do so. Our morally committed heroine confronts Patrick, appealing first to his compassion for the victims. Patrick rather wearily tells her that he has no inclination to concern himself with the plight of strangers. Virginia then appeals to explicit moral considerations: in this case, helping the strangers is his moral obligation and a matter of fighting enormous injustice. Patrick readily declares that he agrees with her moral assessment, but nevertheless cannot be bothered to help. Virginia presses him further, arguing that the effort required is minimal and, given

²⁸Brink's objection against motivational internalism consists essentially in bringing up the case of the amoralist (*Moral Realism*, 46–48): "someone who recognizes certain considerations as moral considerations and yet remains unmoved by them and sees no reason to act on them" (27).

²⁹This is essentially Smith's response to Brink's amorism objection (Smith, *The Moral Problem*, 68–71). Smith goes on to shoulder his part of the burden of argument by giving an argument against externalism. I address that argument in section 6 of this paper.

his position, will cost him close to nothing. Patrick responds that the cost is not really the issue, he just does not care to concern himself with such matters. Later he shows absolutely no sign of regret for either his remarks or his failure to help.

Notice that Patrick has not been described as making a moral judgment without being motivationally affected by it. Nothing has been said about Patrick's mental states. Instead, I have described his overt—verbal and nonverbal—behavior and the features of his situation that would be readily discernible to observers. The description of the case should be readily acceptable to both externalists and internalists.

Patrick's callousness is, of course, baffling, and his exchange with Virginia suggests that he has some major character flaws. But how exactly are we to understand his conduct? Is his assent to Virginia's moral judgment insincere? Is he trying to outrage her? Is he held back by an unreasonable fear of the possible consequences for himself, even if it is obvious that there is no danger to himself involved? Is his moral commitment so weak that the cost in time and energy is enough to override it? Is he hiding his shame and regret? Or is he just callous and cynical about moral matters? It is not clear what to say. However, our epistemic situation in this case is not significantly different from what it is in any other circumstances when we are trying to figure out the psychological states of others. Our understanding of others relies on observations of what they say, their body language, and their overt behavior, preferably over time. It also relies on an epistemological analogue of Davidson's principle of charity. And, furthermore, we have to rely heavily on the assumption that there is, typically, a great stability in people's mental lives over time. Although our mental states are subject to various changes, there is a certain consistency for most of us in how easily and under what conditions these changes are effected, which suggests the existence of longstanding mental dispositions on our part: abiding character traits, deep-rooted concerns or attitudes, and fundamental beliefs.

Let's, therefore, add some further information about how Patrick has behaved in the past.

Additional information about Patrick: Besides being known for courage and conservative estimates of risk to himself, Patrick is independently minded and earnest to a fault—indeed, non-

est to the point of tactlessness and even cruelty. And in any case, he has nothing to gain from misleading Virginia in the given circumstances. Moreover, Patrick makes claims couched in moral terms infrequently and impassionately, and never gives them as reasons for his actions. He has frequently been observed taking actions that seem pretty uncontroversially wrong (or in other ways morally problematic) without displaying any signs of hesitation or regret. In contrast, he has often displayed obvious signs of regret and shame when his plans have misfired, he has overestimated risk to himself, or he has publicly embarrassed himself in matters he finds important. He has also passed up numerous opportunities to perform obvious and uncostly moral deeds. However, when prodded, he will engage in prolonged and intelligent conversations about moral matters and seemingly take an independent stand on the moral status of a controversial public policy or an action. Nonetheless, he usually ends such conversations by volunteering the opinion that he has long ago rid himself of any aspiration to live by moral standards.

I submit that given this information about Patrick, it would be reasonable to exclude not only the hypothesis that he was overcome by unreasonable fear while too ashamed to admit so, but also the hypotheses that his assent was insincere and that his moral commitment wavered or was very limited to start with. Indeed, it seems most plausible to conclude that he has no moral commitment and is completely cynical about moral matters: he knew what was right to do in the circumstances, but could not have cared less.

Motivational internalists will tell us that this cannot be the right conclusion to draw under any circumstances. They will insist that Patrick has—in spite of his disavowal—some inkling of motivation to do what he judges morally right or good, or he is not making a sincere and competent moral judgment. Now, I readily grant internalists that in spite of all outer signs of sincerity and competency in judgment and of lack of motivation, it is *possible* that Patrick is—purely on a whim or for some obscure reason—misleading his interlocutor, or that he is constantly fighting, even repressing, inclinations to pursue what he judges to be of moral value. But I would think that the explanation that strikes me as the most plausible one is also in the running. This internalists will deny.³⁰ Notice

³⁰Dreier acknowledges the epistemic possibility externalists advocate

that at this point, I am not so much interested in which explanation of Patrick's behavior is the most plausible one, but rather in which explanations cannot be readily ruled out as false or in other ways defective. Internalists and externalists have conflicting intuitions in this matter. But when there is a conflict of intuitions (among intelligent and sensible people) about which hypotheses are in the running as an explanation of some observable phenomenon, the burden of argument is on those who insist on a more restrictive class of explanations. This seems to me entirely reasonable as a methodological principle governing empirical investigation.

An example unrelated to the externalist-internalist dispute might help to make this plausible: Imagine that Alice points out to a fellow microbiologist, Gary, an explanation of his data that competes with the explanation he has advanced. Assuming that Alice is no crackpot but, rather, an intelligent and sensible fellow researcher, Gary cannot responsibly dismiss Alice's challenge without having some reason for ruling out her hypothesis. He might be able to rule out the hypothesis as radically misconceived, as not even being a logically possible explanation. Or, he might be able to rule it out as having bizarre metaphysical commitments, being inconsistent with a well-confirmed background theory, being far-fetched, assuming too many coincidences, or overlooking some aspects of the data. But then he might not be able to give any credible reason for ruling out Alice's explanation, in which case he would have to admit that both explanations are in the running,

and, thus, rejects the internalist position that I am investigating in this paper. However, as explained in note 7, he accepts a weaker form of motivational internalism, one that includes the condition that the moral judge must be motivated in the normal context. His putative evidence for the claim that we operate with some sort of normality condition is that the most serious counterexamples to motivational internalism could not be construed as externalists interpret them except against "the background of central internalist cases" conceived of as providing the normal situation ("Speaker Relativism," 13). Dreier would think it significant that my description implies that the cynic in question used to take moral judgments more seriously. He seems to think that our understanding of the cynic as making genuine moral judgments in spite of his lack of motivation depends on seeing his state of mind against the background of a more normal one he used to have and that most other people in his community have (cf. his point against Stocker, at 12). Dreier might be on to something important here, but it is far from clear that it is best captured by a normality condition on motivational internalism.

even if he continues to favor his original one and seek evidence to confirm it. It may be objected that Alice has an equal burden to support her contention that the class of explanatory hypotheses in the running needs to be expanded.³¹ I readily concede that she would have to support such a general claim by proposing at least one new hypothesis and articulating clearly how it is supposed to account for the data (as she has *ex hypothesi* already done). But Alice does not have any further burden of supporting her claim that the hypothesis is in the running unless she has been given some clearly articulated reasons for dismissing it. It strikes me as reasonable that there is this asymmetry in the burden of argument when the issue is simply what explanations cannot readily be ruled out as false or in other ways defective. A methodological principle that condones this asymmetry in the burden of argument serves to counter our lack of imagination, narrow-mindedness, biases, and intellectual laziness: one cannot responsibly dismiss something that strikes other intelligent and sensible people as a feasible explanation without having some story about why it should be ruled out. There is less danger in overlooking the truth if we hang on to a proposed explanation in face of skepticism from intelligent and sensible people so long as no (unanswered) reasons have been given for dismissing it. The contested explanation will minimally serve as a healthy reminder that the support for favored candidates is still inconclusive.

By casting us in the role of observers trying to understand Patrick's conduct, I have shifted our perspective from a philosophical investigation of moral judgments to an empirical investigation of observable behavior. In this context, the conflicting externalist and internalist intuitions are triggered by the question whether a certain hypothesis is in the running as an explanation of the behavior. This has enabled me to appeal to a methodological principle governing empirical investigations to shift the burden of argument onto the internalists. In order to make my point as forcefully as possible, I have concentrated on a case in which the epistemic possibility in dispute seems to me not only one of the hypotheses that need to be considered, but actually the most plausible one,

³¹Thanks are due to Ruth Chang and anonymous referees for the *Philosophical Review* for pressing this objection. And thanks are due to Carsten Hansen for a helpful discussion of it.

given the information provided about the agent's past and present behavior. But that intuition need not be universally shared for my point to go through: *it is motivational internalists who are restricting the range of the hypotheses that are in the running for explaining Partick's conduct, so the burden is on them to justify that restriction.*

Motivational internalists will, of course, claim that the disputed hypothesis is conceptually incoherent. This hardly amounts to meeting the burden of argument, given that intelligent and sensible people in command of the relevant concepts have proposed the explanation and it has the familiar structure of a belief-desire explanation. The conceptual mistake has to be laid bare. At this point, the internalists may elaborate that the hypothesis flies in the face of the conceptual truth that moral judgments necessarily motivate those who make them. It is not question-begging to support the internalist restrictive intuition about the class of feasible explanations by thus invoking the internalist thesis itself. But the move invites the question, Why think this is a conceptual truth? And obviously, it would be question-begging for internalists to appeal to their intuition at that point, since the internalist thesis has already been invoked in defense of that very intuition.

If the internalist thesis is supposed to express a conceptual truth, there must be some concept or concepts that exclude as incoherent the explanatory hypothesis externalists favor.³² It seems the only way to go for the internalist is to identify these concepts and defend an analysis of them that yields the internalist thesis as a

³²Miller (*Moral Differences*) would contest this claim. He provides a neo-Davidsonian argument for motivational internalism: he derives it from a general condition on the ascription of belief rooted in norms of rationality to which an individual must to a great extent conform for his behavior to be interpretable as intentional action. Miller does not offer the thesis as a conceptual truth about moral judgments in the sense that it would fall out of an analysis of either the concept of a moral judgment or the concepts employed in moral judgments. But then it is doubtful that it can be invoked as a condition of adequacy on accounts of moral judgments that *prima facie* favors noncognitivism. Indeed, Miller can be seen as engaged in the project of explaining how a cognitivist can accept motivational internalism, while denying that moral judgments are ways of conceiving of the world that depend on or affect the motivational states of the thinker. I am not inclined to follow his lead, since I am convinced that we can readily interpret an individual as making sincere and competent moral judgments even if he fails to be motivated by them when he perceives it to be cost-free. At the end of section 5, I respond to Miller's argument for the claim that this is impossible.

corollary. There are two candidates. One is the class of concepts employed in moral judgments, for example, *good*, *right*, *just*. The other is the concept of a *moral* judgment: we simply could not conceive of a mental act as a *moral* judgment unless it had the appropriate motivational impact.³³ If the internalist thesis can be shown to fall out of *the best* account of one or the other of these candidates, externalists will have to acknowledge that they were touting an incoherent explanatory hypothesis. I am not going to argue here that such a defense of motivational internalism is bound to fail. Rather I simply want to emphasize that such a defense would be non-question-begging *only if* the internalist thesis had not been invoked as a criterion for deciding between competing accounts of these concepts. It would be blatantly circular to argue for the internalist constraint on the basis that it falls out of the best account of moral concepts, but then defend that account partly on the ground that it, in contrast with some or all of the contenders, meets the internalist constraint. But then internalists are not in a position to advocate their thesis as a constraint of adequacy on accounts of moral thought and language. Even if we could discover on the basis of an a priori investigation that the internalist thesis is a conceptual truth, it is not an obvious conceptual platitude that can be invoked as a constraint of adequacy on accounts of moral thought and language at this stage of our limited understanding of the relevant conceptual resources.³⁴

³³These are genuinely distinct candidates. The best way to appreciate this is to consider the theoretical possibility that moral judgments employ concepts like *good*, *right*, and *just*, but share these concepts with nonmoral judgments; what distinguishes them from judgments that use the same conceptual resources is their attitudinal force, which is in part motivational. On this view, motivational internalism falls out of the concept of a moral judgment rather than out of the concepts employed in moral judgments. Noncognitivists would also prefer the former candidate, since they do not think moral judgments employ any distinct concepts (assuming concepts are representational devices). It is worth noting in this context that Brink interprets motivational internalists (or “appraiser internalism about motives,” as he refers to the position) as claiming that the internalist thesis holds “in virtue of the concept of morality” (Brink, *Moral Realism*, 40).

³⁴Undoubtedly, some philosophers have the hunch that motivational internalism provides a genuine insight into the nature of moral judgments. And proceeding on that hunch might prove fruitful. I have my doubts. Be that as it may, my main point is that it is only a hunch and that the internalist thesis cannot play a crucial role in the context of justification, namely, in evaluating competing accounts of moral judgments.

Internalists may resist this conclusion by questioning whether the only way for them to discharge the burden of proof is by defending on independent grounds an account of moral thought and language that implies the internalist thesis. Instead, they could go on the offensive and attempt to discredit externalism either by revealing more directly a mistake externalists are making when extending the class of feasible explanatory hypotheses or by giving a *reductio* of the externalist position. The familiar “inverted commas” strategy for explaining away the externalist intuition may be construed as a response along the former lines, and Michael Smith’s recent argument against motivational externalism can be construed along the latter lines. In sections 5 and 6, I will argue that both attempts fail. But first I would like to supplement my attack on motivational internalism by offering a debunking diagnosis of the internalist intuition.

4. Explaining Away the Internalist Intuition

Indeed, it does not seem such a hard task to explain away the internalist intuition. First of all, I suspect that in many instances the internalist intuition reflects not a firm grip on moral concepts, but rather a deep moral commitment that makes it hard for the individual in question to imagine how anyone could be motivationally unaffected by his moral judgments. Our attitudes and commitments all too often cloud our imagination in this way. The internalist’s mistake is to think that the possibility he cannot envisage is inconceivable. Secondly, the internalist intuition may also reflect the optimism of the overzealous moralist that moral motivation is somehow guaranteed, if only we get people to see moral matters aright. And thirdly, the internalist intuition may be bred by the wish to close an embarrassing skeptical question. The internalist thesis has the consequence that a certain sort of moral skeptic can simply be dismissed as suffering from conceptual confusion. This is not the skeptic who questions the existence of moral facts or moral properties (call this one a *metaphysical skeptic*), but rather the skeptic who questions our commitment to morality (call this one a *commitment skeptic*). This skeptic does not dispute that there are morally better and worse alternatives, but wonders why that should affect our decision making and action.³⁵ He need not be cynical

³⁵ Obviously, the commitment skeptic is not wondering whether he mor-

about moral judgments; he may be thoroughly committed to morality, but still wonder whether it makes sense—whether there is any justification for being thus committed. But his question presupposes that it is at least conceptually possible to be a moral cynic. So, if internalists are right, this skeptical question can simply be dismissed as an incoherent worry. Now, philosophy is littered with attempts to lay to rest uncomfortable and persistent skeptical questions by unmasking them as based on some conceptual confusion, but most of these have been in vain. A controversial intuition that serves exactly the purpose of closing such a question is extremely suspect.³⁶

Although I have much confidence in the above debunking explanations of the internalist intuition, I doubt they give the full story behind the pervasive appeal of internalism. Admittedly, I would probably experience puzzlement upon encountering an individual who made a moral assessment of his circumstances, yet appeared indifferent to moral values. A part of the puzzlement is over a psychological makeup: how could anyone be so emotionally impoverished as not to connect motivationally with the kind of considerations driving moral evaluation? It is somewhat on a par with the puzzlement some of us have experienced over people who seem to be able to recognize the beauty of a rugged landscape

ally should take into consideration the moral value of the alternatives. Traditionally, the commitment skeptic has been understood as asking whether he rationally should commit to morality; in other words, he is implicitly asking about the relation between morality and rationality.

³⁶Both Nicholas Sturgeon and David Brink have drawn the above distinction between two types of moral skepticism. Sturgeon uses the traditional term ‘amoralism’ to refer to skepticism about moral commitment, while Brink uses the term ‘amoralist skepticism’. I do not use the traditional term ‘amoralist’ because it has been used both to refer to commitment skeptics and (as Brink does) to people who are indifferent to their moral judgments—cynics, in my terminology. But I want to emphasize that commitment skepticism and moral cynicism are two distinct phenomena, as I explain in the text. Brink draws attention to commitment skepticism in order to convince internalists to take his amoralism objection seriously (*Moral Realism*, 46–48). Sturgeon, on the other hand, uses commitment skepticism to make a point similar to the one I have just made. Sturgeon observes that noncognitivists are committed to dismissing amoralism as impossible, and notes that their tactic can be seen as an instance of a discredited anti-skeptical strategy. (See his “What Difference Does it Make Whether Moral Realism is True?” *Southern Journal of Philosophy* 24, supp. (1986): 121.)

without being deeply moved by it: the person who nods with a smile, snaps a picture, and then hurriedly moves on to a more welcoming spot. But the puzzlement seems to run even deeper, especially when the person volunteers the moral judgments: Why would anyone who is completely indifferent to moral considerations bother to take note of the moral conditions of his surroundings? This perplexity may, however, be traced to the assumption that an individual who bothers to make a moral judgment possesses some degree of moral commitment, an assumption that conflicts with the description of the individual as making a moral judgment and not being in the least motivated by it. For a moral commitment is a commitment to the realization of moral values; and it is *a priori* true that one could not be committed to something unless one were, *other things equal*, motivated to do what one judges will *best* protect or enhance that something, and abstain from whatever one judges to be detrimental to that something. This, surely, falls out of the concept of a commitment.

It seems reasonable to operate with the assumption that an individual who bothers to make moral judgments is morally committed to some degree, but I see nothing against giving up this assumption when we get significant evidence to the contrary.³⁷ And once we have done so, there is nothing puzzling about the idea that the person in question fails to be motivated by his sincere and competent moral judgments. It is when we ignore the role of this crucial assumption that we feel the pull toward the internalist position. Put differently: The internalist thesis would be uncontroversial *if* its domain were restricted to morally committed individuals. Those who advocate motivational internalism as a conceptual platitude are ignoring this crucial restriction. Of course, the restriction should be lifted, if it were established that a moral commitment is a precondition for the capacity for making sincere and competent moral judgments. But, given my earlier argument, that would have to be done on grounds that are neutral with respect to the debate between internalists and externalists.³⁸

³⁷Peter Railton has correctly pointed out to me that an explanation is needed of why it is reasonable to operate with this assumption. This should, I believe, be done in the context of explaining how the conceptual resources employed in moral thinking are more intimately bound up with the regulation of conduct than are those involved in conveying information, say, about climatic conditions.

³⁸Interestingly, David McNaughton rests his case for motivational inter-

More puzzling than the cynics we have encountered so far is an individual who judges that one course of action is of lesser moral value than another, but is motivated *on account of that very judgment* to take the former alternative.³⁹ And he is surely rivaled by an agent who becomes all enthused upon judging that an action is of some minor moral value, but loses considerable interest when realizing that it is of much greater moral value than he initially supposed. But rather than pose problems for externalists, these cases nicely illustrate the strength of the above explanatory framework. The fact that the former character is motivated by his moral judgment enforces our initial assumption that he has a moral commitment of some degree, but that is *other things equal* inconsistent with his being more motivated to pursue what he judges to be the lesser of two moral values. (Of course, our diverse commitments may come into conflict and various other things may interfere with our commitments, with the result that overall we are—notwithstanding our moral commitment—motivated to pursue a thing we judge to be of lesser moral value than something else available to us. But it is particularly puzzling when, as in the above case, the very judgment about the relative value of two alternatives does not motivate an agent, committed to that value, to pursue the one he judges to be the greater of the two.) At the same time, it seems more difficult than in the previous cynic cases to withdraw the assumption that the individual in question is morally committed, given that he is motivationally affected by his moral judgment. Similar observations can be made about the second case above. Both imagined scenarios put us in the paradoxical position of *prima facie* being able to treat the individual encountered neither as a moral cynic nor as a morally committed person. Thus, this explanatory framework allows us to account for why these two cases are more puzzling than the previous ones. Nevertheless, there is a way of dispelling the puzzlement in both cases—namely, by telling a story that shows that we should give up our (reasonable) assumption that the agent is morally committed. For example: The person in question is a cynic who is normally indifferent to moral values; however, he has

nalism on the claim that an adequate moral theory must account for the connection between moral commitment and action. See his *Moral Vision* (Oxford: Basil Blackwell, 1988), 20–22. Thanks are due to an anonymous referee for the *Philosophical Review* for pointing this out to me.

³⁹Thanks are due to Paul Boghossian for raising this case.

made a bet with a friend that he will do something of minor moral value in the next few days, while having avowed not to do anything of great moral value in his life. But although this scenario is possible, it is certainly not the most obvious thing to occur to us, nor is it very likely to be true.

The beauty of the above explanation of our likely puzzlement upon encountering moral cynics is that besides relying on a relatively uncontroversial claim about what it is to be committed to something and making sense of the difference in the degree of puzzlement over the above cases, it enables us to understand the pervasive appeal of internalism. Maybe the internalist intuition is not merely the offspring of wishful thinking, moralistic optimism, and substantive commitments. Its roots probably lie in a perfectly legitimate intuition pertaining to the concept of commitment.

5. Explaining Away the Externalist Intuition: A Response

For the remainder of this paper, I assume a more defensive stance and ward off attempts to discredit externalism. I start by considering a common strategy for explaining away the externalist intuition. When carefully developed, it is admittedly quite compelling. Nonetheless, I will argue in this section that it ultimately fails. Those who play this strategy concede to externalists that there are cases in which we have every reason to believe that an agent, first of all, is making competent use of moral terms in voicing *some sort of* an opinion; secondly, is sincerely voicing this opinion; and, thirdly, is not motivated on account of it. But they claim that externalists are, however, making a mistake when they take these to be cases of someone failing to be motivated by his *moral* judgment.

Initially at least, this sounds paradoxical. Aren't moral judgments whatever mental acts are sincerely expressed by competent use of moral language? This may seem the only credible way of identifying in a theory-neutral manner our subject of study and of dispute, namely, moral judgments. If this is right, the above internalist strategy is blocked: one cannot make the above concession to externalists and then charge that they are mistaken in concluding that these cases involve moral judgments. However, this attempt to undermine the internalist strategy does not work. The problem is that it is in general possible to use the same sentence to express sincerely different types of mental states. For example, the sentence

'The streets are gray and deserted' may be competently used to express sincerely either a perceptual state or a detest for one's environment (or both). But then we cannot in general pick out a mental type by describing it as the type of mental state sincerely expressed by competent use of a certain vocabulary. Thus, in order to identify moral judgments in a theory-neutral way along the lines suggested above, we would at least have to add 'literal' or 'standard' to 'competent use of moral language'. But once that has been done, the internalist can charge that externalists fail to appreciate that the cynic is not and cannot be speaking literally or making standard use of moral terms: he is using them to express something other than a *moral* opinion.

This is the spirit of Hare's suggestion that moral cynics (or amor-alists, as he calls them) use moral terms in inverted commas. Hare actually maintains that moral terms have a different meaning when thus used. He gives a noncognitivist account of moral terms in their normal use, but in inverted commas use he gives a cognitivist analysis of them as meaning something like 'it is generally accepted in this society that x is right/good/just' when the moral terms are used in their normal senses in the analysans.⁴⁰ On this account, the cynic is expressing not a moral judgment, but rather an opinion about what kind of moral judgments his fellows would make in his circumstances. (Hare could, indeed, have maintained that the cynic is using moral terms nonstandardly to express such an opinion, without claiming that moral terms are ambiguous.) Though this is an ingenious attempt to block possible counterexamples to Hare's account of moral thought and language, it just does not provide a plausible account of all possible cases of moral cynicism.

My example of Patrick is a case in point. Recall that when prodded, Patrick will engage in prolonged and intelligent conversations about moral matters and seemingly take an independent stand on moral issues. This part of the story can easily be embellished: Patrick has proven to be a very acute moral thinker. He is very good

⁴⁰ *The Language of Morals* (New York: Oxford University Press, 1952), especially, 124–25 and 164–65. Hare's exact words are: "Thus it is possible to say 'You ought to go and call on the So-and-sos' meaning by it no value-judgement at all, but simply the descriptive judgement that such an action is required in order to conform to a standard which people in general, or a certain kind of people not specified but well understood, accept" (164).

at pointing out inconsistencies in moral views, extending moral conclusions to new cases, pointing out morally relevant differences between cases, challenging particular moral judgments on grounds anyone recognizes as potentially relevant, and disputing the relevance of certain factors. In these disputes, he often takes the position of the moral maverick, although his opinions are neither crazy nor flippant. Furthermore, he seems deadly serious when discussing these issues, although he does not aspire to live by the conclusions of his inquiry. Now, certainly Hare's claim that Patrick is making judgments about the opinions of others and using moral terms in a different sense is not plausible. It is ad hoc to maintain that Patrick, when defending a maverick moral position, is talking about what moral judgments people would generally hold in his society, rather than about moral matters as his morally committed interlocutors are doing. It is a rather desperate move to treat this scenario as a case of massive miscommunication.⁴¹ Furthermore, this move would, rather implausibly, have to ascribe to the maverick cynic a wildly mistaken view about the dominating moral view in his society.

But the spirit of Hare's proposal may be saved, though his account of inverted commas use be rejected. It may be argued that inverted commas use is a matter of engaging in moral discourse to simulate moral judgments: Those who make inverted commas use of the moral vocabulary are not expressing any sort of judgment or opinion. Rather they are play-acting or parodying moral evalu-

⁴¹I have come to realize that this point is a close relative to Sturgeon's objection to noncognitivism in "What Difference Does it Make Whether Moral Realism is True?" There he points out that it is implausible to deny that Socrates and Thrasymachus (in Plato's *Republic* I) agree on various things concerning justice; for example, "that the actions of the aspiring despot, in 'appropriating the possessions of the citizens' while enslaving the owners of these possessions as well (344b-c), are unjust" (120). But noncognitivists seem committed to doing so; for while Socrates disapproves of that action as well as any other action he calls unjust, Thrasymachus admires it and other actions he calls unjust. Noncognitivists will have to claim that Socrates is expressing his disapproval, while Thrasymachus is "using the terms 'just' and 'unjust' (that is: the corresponding Greek terms) in some secondary 'descriptive' sense, and hence that his argument with Socrates is only apparent" (121). In other words, Sturgeon's objection is that noncognitivists will implausibly have to treat the agreement, as well as the disagreement, between Socrates and Thrasymachus as a case of massive miscommunication.

ation.⁴² Such a use of moral language is perhaps best illustrated, not by a moral cynic, but rather by a moral reformer who parodies the moral judgments of the conformists. Unless she is engaged in deceptive behavior, the reformer will signal that her judgments, couched in moral terms, are not to be taken as genuine moral judgments. She will do this by ironic emphasis on the moral terms, raising of eyebrows, or other gestures. Certainly, we must concede that such a use can be made of moral language, but it does not follow that all cynics are exploiting moral language in this way.

Undoubtedly, there are cynics who use—indeed, cannot but use—moral language in inverted commas. I have in mind cynics who like the reformer reject conventional moral views, but go further and reject any possible moral view, maintaining that moral discourse and thought is not responsive to any moral facts and should be discarded as nonsensical or fictional.⁴³ Like the reformer who apes the judgments of the conformists, these cynics do not really buy what they are saying.⁴⁴ But that is because of their nihilist

⁴²In *The Moral Problem*, Smith takes a somewhat similar track. When responding to Brink's objection to internalism, Smith claims that amorali-
 sts (cynics, in my terminology) "do not really make moral judgements. . . .
 The point is not that amorali-
 sts really make judgements of some other
 kind: about what other people judge to be right and wrong, for example.
 The point is rather that the *very best* we can say about amorali-
 sts is that they try to make moral judgements but fail" (68). Of course, there is a
 difference between Smith's proposals and the one in the text: Smith's sug-
 gestion is that cynics make honest attempts, but fail at making moral judg-
 ments, whereas my suggestion, on behalf of internalists, is that cynics are
 play-acting or parodying moral judgments. But the proposals are similar in
 that both reject Hare's original suggestion that the cynics are making judg-
 ments, albeit not moral ones. Now, Smith's proposal would not fare any
 better than the one I have suggested. As Smith himself acknowledges, the
 issue of whether his or Brink's description of the cynic's (amoralist's) em-
 ployment of moral language is the right one boils down to the issue of
 whether "being suitably motivated under the appropriate conditions is . . .
 a condition of mastery of moral terms" (70). I agree. My main point in
 this paper is that the latter issue can only be settled by defending an ac-
 count of moral thought and language on grounds that are neutral as far
 as the debate between motivational internalists and externalists goes.

⁴³Miller's moral nihilist is this type of cynic (*Moral Differences*, 84). I am
 in agreement with Miller that his moral nihilist cannot but use moral lan-
 guage in inverted commas. However, given his acceptance of motivational
 internalism, Miller is committed to claiming that no cynic can make a
 genuine moral judgment. I disagree with him on this point. Momentarily,
 I will address Miller's argument to the effect that it is incorrect to ascribe
 moral belief to any cynic.

⁴⁴These characters are much like a scientist who for illustrative purposes

or eliminativist attitudes towards morality, rather than because of their moral cynicism or even skepticism about moral facts. Let me first explain why it is not because of their (metaphysical) skepticism.

Those who question the existence of moral facts or properties need not be moral eliminativists. Noncognitivists, for example, do not take their skepticism about moral facts to undermine moral practice, since they think that moral judgments do not “purport” to be about moral facts anyway. It is also possible for skeptics to maintain that moral judgments are about fictional entities, but argue that engaging in moral discourse still serves important purposes and that we should, therefore, retain it. These two types of metaphysical skeptics take moral practice seriously enough to stay within it. It would be implausible to deny that their contributions to moral discourse are genuine moral judgments. Our capacity to make moral judgments is hardly that sensitive to our philosophical views of them; nor is the nature of our moral judgments directly affected by our metaethical views.⁴⁵ Skepticism about moral facts, therefore, does not doom a person to use moral language in inverted commas. Nor does moral cynicism; for it need not be bound up with an eliminativist attitude toward morality.

Certainly, there are people who have never thought much about the nature or grounding of moral evaluation, but have simply grown disaffected and cynical about morality, people who take it for granted that stealing is wrong and telling the truth is right, but have not found much payoff in staying away from what is wrong or in doing the right thing. Nor would I, from my armchair, dismiss

gives an explanation of some phenomena in terms of a scientific theory she does not accept. The scientist is not strictly speaking describing how the phenomena in question would be explained by the rejected theory, but rather framing an explanation within that theory. However, she is not *offering* it as an explanation of the phenomena; for giving an explanation in terms of a theory that one disowns hardly amounts to offering an explanation of the phenomena. It can at most be an act of illustration or deception.

⁴⁵ Obviously, I am rejecting global and other strong versions of semantic holism. If the reader finds it implausible that noncognitivists, error-theorists, and realists about moral discourse are making the same sort of judgments within that discourse so long as they continue to participate earnestly in it, I urge him or her to consider whether the same is to be said about belief-ascriptions made by philosophers who have different views about belief-ascriptions.

the possibility of meeting a sophisticated but cynical moral thinker, who in spite of disaffection with morality does not question moral discourse in the way that the moral nihilist does—a person who, say, was raised among extremely morally reflective and morally conscientious people, but ultimately grew disaffected and weary of living his life according to moral strictures, although never really questioning the existence of moral values and moral requirements. Such a person may have gotten an early overdose of morality, so to speak. Notice that the suggestion is not that moral judgments fail to motivate this agent because he perceives moral action to be costly to himself. Although a sense of costliness might have helped to wear down his commitment to morality, his disaffection with morality is total at this point: he would not be moved by his moral judgments even in situations in which he perceives moral action to be cost-free. Such a person need not have turned mean or cruel or self-indulgent. Indeed, he may live a morally valuable or acceptable life, but it is not a life *guided by moral judgments*. Still, I would think that such a cynic is fully capable of making genuine moral judgments. Internalists will insist that this is a mistake on my part: I am failing to appreciate that this cynic is using moral terms in inverted commas. But why think he is? Unlike the reformer and the eliminativist, this cynic buys what he is saying; for he does not question the moral requirements or values of which he speaks. Undoubtedly, internalists will dig in their heels at this point and claim that this cynic does not *really* buy what he is saying; it is not possible to acknowledge moral requirements or values without being somewhat motivated to honor them. But how are they going to defend that claim? They had better not rely on the internalist intuition at this point in our dialectic.

Richard Miller has suggested that it follows from a condition on ascription of beliefs that a person cannot have moral beliefs if she fails to be motivated by moral considerations at least when she perceives this as involving no cost to herself.⁴⁶ The proposed condition is that the ascription of belief must play an appropriate role in giving rationalizing explanations of the person's conduct. Miller points out that the ascription of moral belief to the cynic could serve to provide rationalizing explanation only of her verbal conduct. But "the desires explaining why she makes the usual moral

⁴⁶ *Moral Differences*, 95.

utterances are desires to conform, to manipulate, to be left alone, or the like.”⁴⁷ And—Miller goes on—the “beliefs combining with such desires to make her utterances rational conduct for her are not the moral beliefs the utterances standardly express but non-moral beliefs concerning what people like to hear, and so forth.”⁴⁸ Miller concludes on these grounds that it would be incorrect to ascribe to the cynic moral beliefs.

It is not clear that we should accept the proposed condition on the ascription of belief, but let’s grant it for the sake of argument. I also grant Miller that the motivation for the cynic’s relevant verbal behavior has its source in conative states whose content does not immediately relate to the content of moral beliefs in the appropriate way for rationalizing her verbal conduct. But this does not imply that a genuine moral belief does not enter into the rationalizing explanations of the utterances in question. This would be true even if we accepted Miller’s suggestion about what sort of desires would have to be ascribed to the cynic. For with little ingenuity we can tell a story that suggests that articulating her sincere (even unorthodox) moral belief serves her sinister purpose so that an ascription of moral belief to her has a role in a rationalizing explanation of her conduct. For example, the cynic may, in expressing her moral beliefs, have the sinister purpose of enchanting an impressionable youth who is intrigued by people’s moral opinions. Of course, there would be competing rationalizing explanations of the same conduct that would not involve an ascription of moral belief, and it is open to Miller to argue that they would always be the more plausible ones. But that would have to be argued by appealing to something more than the proposed condition on belief ascription. Some might appeal to considerations of simplicity, preferring to avoid ascribing a whole new class of beliefs, namely, moral ones, as part of the overall explanation of the cynic’s conduct and dispositions. But such a decisive role for parsimony stands, at the very least, in need of a justification that Miller does not provide.⁴⁹

I, moreover, reject Miller’s suggestion that we always have to as-

⁴⁷Ibid.

⁴⁸Ibid.

⁴⁹Thanks are due to a referee for the *Philosophical Review* for helping me to formulate this point.

cribe sinister motives to cynics in order to give a rationalizing explanation of their verbal behavior involving moral terms. Say our non-eliminativist cynic is Patrick. Why did Patrick say what he did in response to Virginia's queries? Given what we know about Patrick's past, one possible explanation involves ascribing to him the desire to speak his mind. Another would ascribe to him the desire to respond sincerely to Virginia's questions. Similarly, his engagement in moral debates, when prodded, may be explained by ascribing to him a love of debate, a desire to set others straight, or a taste for finding problems in others' reasoning. And I would think that such desires rationalize the engagement in moral discourse rather than the simulation of engagement in moral discourse, if the former is an available option to the agent. The only reason for preferring rationalizing explanations that ascribe some sort of simulation of moral belief rather than moral belief to Patrick would be that we have independent reasons for thinking that Patrick does not have the capacity to form moral beliefs, or at least does not have moral beliefs.

We have seen that the inverted commas strategy for explaining away the externalist intuition ultimately requires a defense of the claim that it is impossible to acknowledge moral requirements or values without being motivated to honor them—that is to say, make a moral judgment without being motivated by it. And I see no shorter route for defending that claim than to argue for an account of moral thought that shows that a moral commitment is a precondition for having the capacity to make genuine moral judgments. Thus, it has turned out that the inverted commas strategy does not help internalists to avert the conclusions of section 3.

6. Smith's Fetishism Charge

The conclusion would, however, be averted if Michael Smith's recent argument against motivational externalism were successful. The argument exists in two versions.⁵⁰ Although the second version

⁵⁰The first version is given in §3.5 of his *The Moral Problem*, while the second version appears in his "The Argument for Internalism: Reply to Miller," *Analysis* 56 (1996): 175–83. Parenthetical references in the text to "MP" and "RM" are to these two works, respectively. Smith accepts a thesis closely related to motivational internalism (as I have characterized that position), which he calls "the practicality requirement on moral judgment" (see note 4 for a discussion of the subtle difference between these

is offered as a clarification of the first, it is clearly distinct from and not obviously an improvement over the first. Therefore, both versions need to be addressed. Since Smith's argument—in both versions—targets specifically the view of moral motivation to which I have committed myself, my discussion will serve the double purpose of responding to Smith's counter to externalism and making a case for my view of moral motivation.

6.1 *The First Version*

In *The Moral Problem*, Smith's basic idea is that externalists cannot satisfactorily account for the motivational dispositions of morally good people.⁵¹ His argument starts with the observation that

[b]y all accounts, it is a striking fact about moral motivation that a *change in motivation* follows reliably in the wake of a *change in moral judgment*, at least in the good and strong-willed person. (MP, 71)

Any plausible theory of moral judgment, Smith insists, "must . . . explain this striking fact." Internalist theories, of course, take this as an instance of a more general phenomenon of moral judgments being necessarily connected to motivation.⁵² And they are already

two positions). Smith's argument does not hang on whether externalism be understood as contrasting with motivational internalism or with the practicality requirement. A third statement of the argument—a hybrid of the two versions that does not avoid the problems with either one—appears in "In Defense of *The Moral Problem*: A Reply to Brink, Copp, and Sayre-McCord," *Ethics* 108 (1997–98): 84–119.

⁵¹Smith has declared in print that he regrets his choice of terminology when giving the anti-externalist argument in his book. He claims he did not really mean "the term 'good' to pick out those who are good in the more substantive sense of having the motivations that one true morality tells them that they should have" ("Reply to Miller," 177). In this context, Smith refers to objections he has received from David Copp, objections that depend on the above reading of "good people" (see Copp, "Moral Obligation"). I suspect that Smith will also dismiss my response to his argument on similar grounds. However, I am confident that, as the quotes from his book demonstrate, his mishap is not merely an unfortunate choice of terminology. I, therefore, stand by my response. For different criticisms of Smith's argument see Alexander Miller's "An Objection to Smith's Argument for Internalism," *Analysis* 56 (1996): 169–74; David Brink, "Moral Motivation," *Ethics* 108 (1997–98): 4–32; and David Copp, "Belief, Reason, and Motivation: Michael Smith's *The Moral Problem*," *Ethics* 108 (1997–98): 33–54.

⁵²Smith would add "in rational people." Since by this qualification he intends to exclude, among others, the weak-willed, he specifies that the person would have to be strong-willed as well as good in order for the connection to hold reliably.

committed to giving an account of moral judgments that explains it. In contrast, externalists have to treat this as a special case and explain why mental acts, which are generally neither necessarily nor reliably connected to motivation, are reliably connected to motivation in the good and strong-willed person. Smith argues that externalists have to attribute the reliability of this connection in the good and strong-willed person to a unique motivational disposition in virtue of which a person counts as good (MP, 73). He furthermore insists that this motivational disposition would have to be specified as the disposition to be motivated *to do the right thing*. But this, Smith maintains, is fatal to the externalist position; for it gives an absurd account of the motivation of the good person.

However, if this is the best explanation the strong externalist can give of the reliable connection between moral judgement and motivation in the good and strong-willed person then it seems to me that we have a straightforward *reductio*. For the explanation is only as plausible as the claim that the good person is, at bottom, motivated to do what is right, where this is read *de dicto* and not *de re*, and that is surely a quite implausible claim. For commonsense tells us that if good people judge it right to be honest, or right to care for their children and friends and fellows, or right for people to get what they deserve, then they care non-derivatively about these things. Good people care non-derivatively about honesty, the weal and woe of their children and friends, the well-being of their fellows, people getting what they deserve, justice, equality, and the like, not just one thing: doing what they believe to be right, where this is read *de dicto* and not *de re*. Indeed, commonsense tells us that being so motivated is a fetish or moral vice, not the one and only moral virtue. (MP, 75).

The problem with externalism is thus supposed to be that it is committed to a completely wrong-headed account of the motivation of the good person. Indeed, the motivational state it identifies as characteristic of the good person is instead characteristic of people with a peculiar moral vice.

I readily grant Smith that “a *change in motivation* follows reliably in the wake of a *change in moral judgment* . . . in the good and strong-willed person.” Whatever else there is to it, being motivated to pursue what is, by one’s own lights, of moral value is a part of being a good person.⁵³ I furthermore accept that externalists

⁵³Copp (“Moral Obligation”) does not grant Smith this. He argues that the connection between moral judgment and motivation would not be reliable in a good, but weak-willed, person, and concludes that it is strength

should explain the reliability of this connection by ascribing to the good person a conative state whose content involves moral concepts. This is not, *pace* Smith, the only route that an externalist could go;⁵⁴ but it is, I believe, the right route. However, I take exception to Smith's claim that the conceptual content of this conative state must be "to do the right thing." For other moral concepts could also figure in its content.⁵⁵ The externalist account I

of will rather than goodness of character that is responsible for the reliable connection. He claims that an externalist must and can explain the reliable connection in terms of the nature of the strong-willed person. Now, Smith would not object to the claim that the reliable connection is absent in a good, but weak-willed, person; his practicality requirement (as well as motivational internalism, in my formulation) has a built-in exception for the weak-willed person, be he good or evil. One might, therefore, take Copp's conclusion that it is strength of will rather than goodness of character that is responsible for the reliable connection as a concession to Smith's version of motivational internalism; for Smith, strength of will is a part of practical rationality. But if I understand Copp correctly, he rejects the idea that strength of will is a necessary part of practical rationality; it is a distinct characteristic that only some practically rational moral judges have. Thus, Copp's conclusion should not be interpreted as a concession to Smith's internalism. Copp's point is really that the credibility of Smith's initial premise (that "*a change in motivation follows reliably in the wake of a change in moral judgement . . . in the good and strong-willed person*") depends on a different reading of 'strong-willed' than Smith is ready to give. I am not taking this line, since I am willing to grant Smith's initial premise, given his reading of 'strong-willed'. Notice that this issue between Copp and Smith illustrates what sort of can of worms is opened by the introduction of the practical rationality condition on the internalist thesis. It could, for example, be conceded to Smith that strength of will is a necessary ingredient in practical rationality, but then argued that Copp is right that strength of will is a distinctive characteristic that is responsible for the reliable connection between moral judgment and motivation and that, moreover, is not necessary for having the capacity to make moral judgments. Whoever takes that position accepts Smith's practicality requirement, but rejects his idea that it reveals an important fact about the preconditions for the mastery of moral concepts.

⁵⁴See Brink's and Railton's suggestions reviewed in section 1.

⁵⁵The conceptual content of this conative state in each individual must, I submit, reflect the moral concepts he employs. For someone may be a good person, even if he holds a somewhat mistaken moral view. Say the correct moral view is not duty-based. Still, someone who holds a duty-based morality may be a good person, so long as he displays the virtues that are crucial to being a good person (say, kindness, compassion, and honesty) and cares about doing what, by his lights, is of moral value. However, there might be limits to how mistaken the moral view could be. Is it, for example, possible to be a good person, yet hold a moral view that has strong racist or sexist implications? This is a tricky question, as we can come to appreciate if we reflect on indi-

propose does not ascribe to the good person a particular concern with doing the right thing. Rather it ascribes to him a more general concern with doing what is morally valuable or required, when that might include what is just, fair, honest, etc. However, I do not expect that this amendment will appease Smith. Undoubtedly, he would be ready to replace 'right' with any other moral term in the supposedly implausible claim he ascribes to externalists, namely, the claim that "the good person is, at bottom, motivated to do what is right, where this is read *de dicto* and not *de re*."

One way of understanding this claim is as amounting to the thesis that the good person is deep-down only interested in one thing, and that is to do the right or morally valuable thing; any other concerns she might have depend on this one. This reading is indeed suggested by Smith's elaboration of his objection:

[i]f good people judge it right to be honest, or right to care for their children and friends and fellows, or right for people to get what they deserve, then they care non-derivatively about these things. Good people care non-derivatively about honesty, the weal and woe of their children and friends, the well-being of their fellows, people getting what they deserve, justice, equality, and the like, not just one thing: doing what they believe to be right [or: morally valuable or required], where this is read *de dicto* and not *de re*. (MP, 75)

Here, Smith speaks as if the externalist maintains that good people care non-derivatively about "just one thing: doing what they believe to be right [or: morally valuable or required], where this is read *de dicto*." I entirely agree with Smith that this is an implausible conception of the good person. There is much more to being a good person than having such a concern. A good person is also considerate, compassionate, kind, loyal, and honest. And these characteristics do not involve motivational dispositions that are engaged by distinctively moral representations of one's circumstances or behavioral alternatives. Typically, a kind person does not undertake a kind act because he has conceived of it in positive moral terms. It is simply the conception of the action as yielding comfort,

viduals from an entirely different era. I tend to think that in evaluating persons who hold reprehensible moral views, we take into account whether they are in a good epistemic position for discovering the mistakes in their moral views. A person who is epistemically in a position to recognize the serious problems with his moral view, but doesn't, does not count as good, even if he is morally conscientious, kind, compassionate, and honest.

relief, or encouragement to someone that makes him undertake it. Thus, the thesis that the good person is deep-down only interested in doing the right or valuable thing is implausible. But certainly, externalists are not committed to any such thing just on the basis of explaining the reliable connection between moral judgments and motivation in the good and strong-willed person by ascribing to her the desire to be moral. That only commits them to maintaining that the desire to be moral is *a part of* the motivational structure of the good person.⁵⁶ Thus, insofar as the force of Smith's objection trades on ascribing to externalists a commitment to such a monolithic conception of the good person, it is not to be taken seriously.

Is there another reading of the controversial claim that is fair to the externalist position, yet makes it as implausible as Smith wants to maintain? We should bear in mind that Smith is particularly interested in how externalists would explain the fact that change in motivation reliably follows upon a change in moral judgment in the good and strong-willed person. The proposal under consideration is that the presence of the desire to be moral serves to explain why once a good person has been convinced that ϕ is right rather than wrong, as she used to think, she reliably comes to care about doing ϕ , even if it was unattractive to her before. In these circumstances, the good person is motivated to ϕ just because she wants to do the right thing and has come to believe that doing ϕ is the right thing to do. But then it seems reasonable to say that *in such circumstances* the good person "is, at bottom, motivated to do what is right, where this is read *de dicto* and not *de re*." It is a way of saying that in such circumstances it is her desire to be moral that fuels her motivation to ϕ . But why is it implausible to conceive of the psychology of the good person—who has just undergone a change of mind and heart concerning moral matters—in just this way?

Smith, recall, charges that externalists are committed to ascribing to the good person the vice of making a fetish of morality.⁵⁷ I

⁵⁶Copp ("Moral Obligation") makes essentially the same point with his example of Dana.

⁵⁷Indeed, Smith makes the stronger claim that externalists are committed to giving an account of the motivational system of the good person that amounts to elevating this moral vice into the one and only moral virtue. This stronger claim relies in part on the mistaken assumption that externalists are committed to the view that the desire to be moral is the

find this a curious charge. *Webster's New Dictionary of Synonyms* comments that 'fetish' "[i]n extended use . . . may be applied to whatever is unreasonably or irrationally regarded as sacred or sacrosanct." This seems to corroborate my understanding of what moral fetishism would be like. It would be the characteristic of holding oneself and others to very rigorous moral standards, while being completely unwilling to entertain any reflective question about their nature or grounds. It would be accompanied by a fear of any skeptical questions about morality, and a refusal to take them seriously enough to even attempt a thoughtful answer. The question 'Why be moral?' would be branded as irreverent and illegitimate. But certainly, the desire to be moral will not alone yield such a character trait.⁵⁸ A concern for being moral should not be confused with a rigorous obsession with morality or a resistance to examine hard reflective questions about morality.

At the risk of being repetitive, I'll rephrase my bafflement at Smith's charge with reference to the point of agreement between internalists and externalists. They both agree that one of the traits of the good person is that she is motivationally engaged by her moral conception of her circumstances to the extent that she will form desires to do things that she was before indifferent to, or even abhorred. However, internalists maintain that she shares this disposition with all competent (and rational) moral judges, whatever is to be said about their character, while externalists hold that competent (and rational) moral judges may fail to have this disposition. Externalists propose to explain this psychological difference by ascribing to those who are motivationally engaged by their moral judgments an appropriate conative state that the others lack. Now, I cannot see how it would make one a worse person if the disposition to care about what one deems morally valuable were due to a desire to be moral. I would think that the crucial thing

only defining characteristic of the good person. Again, I do not see why externalists need to be committed to this.

⁵⁸In "Duty and Interest," Prichard actually argues that given the presence of the desire to be moral, it makes no sense to raise the question 'Why be moral?' His argument depends on interpreting the question as a request for an incentive to act morally. But that is a mistaken premise. The question may also serve to initiate critical reflections on our commitment to morality. As such it can be intelligibly raised by people who already have the desire to be moral.

is that one has such a motivational disposition (as well as certain others) and that it is not due to some desire such as to impress other people or to stay out of trouble or to obtain some other personal gain from being perceived as morally conscientious. To say that it is due to the desire to be moral is really to say that one need only conceive of something in moral terms in order to be motivationally affected, that no further motivation is needed to care to do what one judges morally favorable or avoid what one judges morally unfavorable. Internalists, of course, want to say the same. But they do not see any need to postulate a desire to be moral to explain this phenomenon, since they hold that making a moral judgment suffices in general for moral motivation.

Here we might have come to the crux of the matter. It might be thought that an individual cannot be a good person if her appreciation of ϕ 's being of moral value does not suffice to make her want to pursue or promote ϕ . This is suggested by Smith's William-esque rephrasing of his objection to externalism:

For the objection . . . is simply that, in taking it that a good person is motivated to do what she believes right, where this is read *de dicto* and not *de re*, externalists . . . provide the morally good person with "one thought too many." They alienate her from the ends at which morality properly aims. (MP, 76)

It is, however, far from obvious what the alienating thought is supposed to be. Externalists agree with internalists that a good and strong-willed person passes directly from conceiving of ϕ as having moral value to wanting to pursue or promote ϕ . It is just that this transition, according to externalism, would not occur in every moral judge, so being in the state corresponding to the moral judgment cannot suffice for forming the concern for ϕ . Something else must then establish this pathway between moral judgment and motivation in the good person: this is the desire to be moral. But instead of interjecting itself as a thought between the judgment and the desire produced, it establishes a direct psychological transition from the judgment to the desire.

Of course, we should not forget that intentional mental states are peculiar in that they give rise to a certain intentional perspective on the world. The point of ascribing to agents mental states is not merely to explain patterns of similarity and difference in their conduct, but also to reconstruct their intentional perspective on

their circumstances and their actions. Perhaps Smith is calling attention to some undesirable element in the intentional perspective of a person in whom the reliable connection between moral judgment and motivation is forged by the desire to be moral. This idea needs to be examined carefully.

It should not be thought that the intentional perspective of that person is such that she “sees” herself as ϕ -ing because she desires to promote moral values and believes that ϕ is a moral value. From the first-person perspective, we do not normally think about our conduct in terms of the type of psychological state that underlies our thinking about and choosing between our behavioral alternatives.⁵⁹ We simply think about our behavioral alternatives and circumstances, and how they relate to us, various other people, and various projects. However, to what and whom we pay attention reflects our concerns and commitments. So does our first-person conception of why we act as we do. When we “see” ourselves as undertaking ϕ because it has such and such characteristics, we are not offering a third-person causal explanation of our conduct: we are saying neither that its having such and such characteristics caused us to ϕ , nor that our believing that it has such and such characteristics caused us to ϕ . We are, though, in a way explaining ourselves to ourselves and to others; for we are revealing our motivating thought—the conception of the action or its end that engaged us motivationally. And in doing so, we reveal what our concerns are; we reveal, barring self-deception, why we cared to ϕ .⁶⁰ Now, consider again my externalist account: the good person is motivated to ϕ and, let’s assume, actually ϕ s as a consequence of believing that ϕ is μ (when μ is a moral concept) and desiring to μ . Notice that there is an internal link—a relation accessible from

⁵⁹Smith certainly agrees with me on this point. We both reject what he and Philip Pettit have coined “the foreground view of desires.” See P. Pettit and M. Smith, “Backgrounding Desire,” *Philosophical Review* 99 (1990): 565–92.

⁶⁰It is very tempting to use the notions of *reason* and *justification* in this context. I am resisting the temptation, though, since that move raises a lot of subtle questions that I want to avoid at this point. I believe an agent can see himself as doing something because it is β without thinking that its being β carries much justificatory force or gives him anything like a reason for action. Philosophers have been far too ready to equate justifying one’s actions to oneself or others with explaining one’s actions to oneself or others.

inside the first-person perspective—between these mental states in virtue of their content: the content of the belief relates the objects of the two desires. This internal link suggests that the intentional perspective of the good person is, on this account, such that she “sees” herself as undertaking ϕ because it is μ —say, because it is morally valuable. Does this introduce an undesirable element into the intentional perspective of the good person?

Maybe it is best to approach this question by considering what the intentional perspective of the good person would be on Smith’s account. Smith acknowledges that there is no unified internalist position on the psychology underlying the reliable connection between moral judgment and motivation in all strong-willed (or rational) moral judges. He himself maintains that the moral judgment causes a nonderivative desire for the object of evaluation under the nonmoral description that provides the grounds of a positive moral evaluation of it.⁶¹ In other words, the good and strong-willed person’s belief that ϕ is μ causes her to desire to ϕ , which in turn makes her (other things being equal) ϕ . Notice that there is no internal link between these intentional states to suggest that the good person sees herself as ϕ -ing because it is μ . The account suggests that the good person simply “sees” herself as doing ϕ because it is ϕ . Although the belief is the causal antecedent of the desire (and hence of the action), their contents do not relate conceptually such that the content of the belief is seen as

⁶¹According to Smith, all internalists (or defenders of the practicality requirement) depict the connection between moral judgment and motivation as follows: “Thus, if an agent judges it right to ϕ in C, and if she has not derived this judgement from some more fundamental judgement about what it is right to do in C, then, absent weakness of will and the like, defenders of the practicality requirement can insist that she will be motivated non-derivatively to ϕ in C. This is because on the rationalist alternative, a non-derivative desire to ϕ in C is what her judgement that it is right to ϕ in C causes in her, or because, on the expressivist alternative, the judgement that it is right to ϕ in C is itself just the expression of such a non-derivative desire” (*The Moral Problem*, 73). Smith accepts the rationalist alternative. He then argues that this connection holds at least in rational agents because the content of the moral judgment that it is right for S to ϕ in C is (skipping one qualification that need not detain us here) equivalent to the proposition that S would desire that he ϕ in C, if S were fully rational. It would be a sign of irrationality to fail to want *de re* to do what by one’s own lights one would want to do if one were rational; that is, part of rationality is to be disposed to form the desires one (correctly or incorrectly) believes that one would have if one were rational.

providing the motive for the desire to ϕ , and hence for ϕ -ing. Thus, Smith's account does not portray the good person as "seeing" herself as ϕ -ing because it is, say, morally valuable.⁶²

Is it somehow a virtue of Smith's account that it does not suggest that the good person "sees" herself as ϕ -ing because it is morally valuable? Is this one thought too many? I would think the contrary, but to appreciate fully Smith's complaint we need to explore his allusion to Williams's objection to moral theories that emphasize impartiality. Williams charges that a man who rescues his wife rather than a perfect stranger (when he cannot save them both) would have one thought too many if his motivating thought were "that it was his wife and that in situations of this kind it is permissible to save one's wife."⁶³ Williams's point is that this motivating thought would not be compatible with having a deep personal attachment and emotional commitment to his wife. He then employs this point to challenge any moral theory that would recommend or require that we think thus about circumstances in which we have to choose between saving our spouse and a complete stranger, since it would implausibly portray the moral life as incompatible with such strong emotional commitments. Morality, Williams contends, should not alienate us from the commitments that make our lives worth living. Smith plays off this theme when he says that his complaint with the externalist account is that it ascribes to the

⁶²Interestingly, it is far from clear that other internalist accounts will portray the intentional perspective of the good person as Smith does; they might side with me. Expressivists, according to Smith, maintain that the moral judgment is but an expression of a nonderivative desire to ϕ —hence, the reliable connection between making a positive moral evaluation of ϕ and being nonderivatively motivated to ϕ (see the quote in the previous note). Since, on this account, the good person has a nonderivative concern for doing ϕ , one might think that she "sees" herself as doing ϕ simply because it is ϕ . However, in the public language this person expresses the desire to ϕ with " ϕ has moral value," and maybe expressivists will maintain that her intentional perspective is also best expressed by these words. A third internalist alternative (which Smith does not mention in this section of his book) is to maintain that moral judgments manifest desires. On this account as on my externalist account, the intentional perspective of the good person is such that she "sees" herself as ϕ -ing because it is morally valuable. For that is the conception of her action that engages her motivationally. It is just that a desire rather than a pure conative state is postulated to explain this motivational engagement.

⁶³"Persons, Character and Morality," in his *Moral Luck* (Cambridge: Cambridge University Press, 1981), 18.

good person “one thought too many,” which “alienate[s] her from the ends at which morality properly aims” (MP, 76). The idea seems to be this: fundamental moral judgments like “ ϕ is of moral value” direct us towards ends at which morality properly aims. However, if we pursue these aims motivated by the thought that they have moral value, we remain alienated from them. For having this motivating thought is incompatible with having a deep commitment to these ends, since it prevents us from pursuing them for their own sake. But certainly a morally good person must have a deep commitment to morally valuable ends, when this is read *de re*. Consequently, she cannot see herself as ϕ -ing because it is morally valuable, as the externalist account implies she does.

Smith seems to have lost sight of the reason why externalists ascribe the desire to be moral to the good person. It is to explain why change in motivation reliably follows upon a change in moral judgment. It does not seem implausible or undesirable that when such a change occurs the perspective of the good person is as externalists (and indeed most internalists) would predict, namely, “seeing” herself as having become occupied with the end in question because it is morally valuable. It hardly shows a moral failing that her commitment to the end is initially predicated on her realization that it is morally valuable. Take the example Smith discusses: an imaginary case of his giving up the view that libertarians hold the (morally) right values for the view that the social democrats hold the (morally) right values (MP, 71). As a consequence of this change of mind, Smith’s motivations change: he no longer wants to vote for the libertarian party and instead wants to vote for the social democrats. Now, would we conclude that Smith is not a good person if his first-person perspective *during or immediately after his change of mind* were best captured by the thought that he now wants to vote social democratic because social democratic values are more morally acceptable than libertarian values? I doubt so. However, if Smith just could not get beyond this stage in his moral and social development, and could never exert himself for social democratic values without reminding himself of his moral conclusion regarding them, then I would think there was something funny going on in his psychology and I’d hesitate to hold his personality up as a moral ideal. Admittedly, we expect a good person to develop a deep commitment to an end she has come to see as morally valuable and to pursue it for its own sake. But nothing in

the externalist account of moral motivation precludes this. The presence in the good person of the desire to be moral certainly does not prevent her from forming such a commitment.⁶⁴ Although her desire to ϕ may initially be derived from her desire to be moral, it may subsequently come to operate psychologically independently of the latter.⁶⁵ I emphasize once again that externalists are *not* committed to the view that the desire to be moral is and remains the only self-standing conative state of the good person.

I have been construing Smith as trying to bring normative considerations to bear on a metaethical issue. He relies on the claim that a good person would have a certain motivational disposition. This sounds like an evaluative claim specifying one of the necessary conditions for being a good person. The disposition in question has to do with the role of moral judgments in motivation, and Smith seemingly tries to exploit this (partial) conception of the good person to defend the internalist thesis, which he in turn advocates as a condition of adequacy on accounts of moral judgments. However, Smith has recently denied that he has ever tried to bring normative considerations to bear on a metaethical issue in this way.⁶⁶ Given his comments about the concerns of good peo-

⁶⁴Copp ("Moral Obligation") makes a similar point during his discussion of Dana.

⁶⁵Throughout this section I have taken Smith's distinction between derivative and nonderivative desires as intuitively clear and not in need of discussion. Nevertheless, I think that it stands in need of clarification, as this paragraph reveals. One can understand 'a nonderivative desire' as designating a desire that is not acquired as a result of another desire and the belief that this latter desire would be or could only be fulfilled if the object of the former desire were procured. Let's call such desires "acquisition-nonderivative." In this sense, the desire to ϕ is derivative on my account. But one can also understand 'a nonderivative desire' as designating a desire that at a given point in time is *self-standing*, in that its role in the psychology of its possessor at that time is not limited to motivating him to do something he believes is needed in order to satisfy another desire. The desire to ϕ can become nonderivative in this sense in the good person, even if it is acquisition-derivative. For Smith's argument to go through the desire to ϕ must be acquisition- nonderivative in the good person. But I see no merit in that claim, although I agree that the psychology of the good person has to be such that it is possible for the desire to become nonderivative in the sense of self-standing. See also my discussion below of what Smith means by an "instrumental" desire in his recent reformulation of the anti-externalist argument.

⁶⁶"The Argument for Internalism: Reply to Miller" (see note 51). In "In Defense of *The Moral Problem*," Smith appears to take back this disclaimer. There he claims that the problem with the externalist account of

ple and about moral fetishism being a vice rather than the one and only moral virtue, it is hard to take this disclaimer seriously. Be that as it may, I need to consider his restatement of the anti-externalist argument.

6.2 *The Second Version*

Smith's argument now starts with a claim he calls "Weak Moralism Internalism" (WMI): "If an agent judges it right to ϕ in C, and that agent is a moralist, then she is motivated to ϕ in C, at least absent weakness of will and the like" (RM, 176). Smith introduces the term 'moralist' by way of a stipulative definition (which mirrors Brink's definition of the contrasting term 'amoralist'): "moralists are those people who are such that, when they make judgements about what it is right to do they *are* motivated to act accordingly, at least absent weakness of will and the like" (RM, 176). But then of course, WMI is true by definition and both internalists and externalists have to accept it as such. The issue between them concerns whether the contrasting class of amorlists is empty. Now, Smith says something very strange: "My argument for internalism is that it alone is able to give a plausible answer" to the question "*why* . . . think that [WMI] expresses a conceptual truth" (RM, 178). This cannot be right, for both internalists and externalists should give the same answer, namely, that WMI is a conceptual truth because it follows from the stipulated definition of 'moralist'. WMI is an extremely uninteresting thesis for exactly this reason.⁶⁷

moral motivation is that it does not "square . . . with our commonsense idea of moral virtue" (112). As in *The Moral Problem*, his argument depends on the mistaken assumption that those who explain moral motivation—motivation by moral judgments—as having its source in the desire to be moral must think that this desire is the source of all the motivations characteristic of the virtuous, or at least that this desire takes center stage in the motivational system of the virtuous.

⁶⁷However, notice that it is not a far cry from a more interesting internalist thesis, namely, motivational internalism explicitly restricted to morally committed individuals—a thesis that we may dub MCI and that I accepted in section 5. Morally committed individuals are moralists, although the class of moralists, as defined, is broader than that of the morally committed. Amongst moralists there may be morally *uncommitted* individuals who for some perverse reason are, as consistently as any of us, motivated to act as they judge morally valuable or required. The domain of the internalist thesis I am willing to accept is, therefore, even more restricted

When Smith lays out the answers that internalists and externalists will supposedly give to his query about the modal status of WMI, it becomes apparent what he has in mind: internalists and externalists will give a very different explanation of why it is true of a particular individual (who is a moralist) *that, come circumstances C, either he is motivated to ϕ or he is suffering from weakness of will*. Externalists will explain why the italicized proposition is true of the given individual by citing two facts: “the fact that the agent in question judges it right to ϕ in C and the fact that the agent in question is a moralist,” whereas “internalists tell us that the first of these facts is sufficient . . . all by itself” (RM, 178). Of course, internalists will not deny that the agent in question is a moralist, but they will not think that mentioning this adds anything to the explanation. Externalists, on the other hand, need to cite the fact that the individual in question is a moralist in order to complete their explanation of why he has the motivational disposition he has. They can then be pressed to expand their explanation by revealing what psychological mechanism is responsible for the fact that a particular individual is a moralist rather than an amoralist, that is to say, why he is motivated by moral judgments. And this is where the externalist story becomes implausible according to Smith. Certainly, this is an unnecessarily complex way of saying that the externalist account of moral motivation is to be disputed.

Smith claims that the only account of moral motivation available to the externalist traces moral motivation to a desire to do what is right. This is not quite right, for the same reasons that this is not the only available externalist account of the reliable connection between moral judgment and motivation in good people. But minor adjustments in Smith’s argument will make my account of moral motivation its prime target, so it is incumbent upon me to address it. His objection is twofold: first, the externalist account “re-describe[s] familiar psychological processes in ways that depart radically from the descriptions that we would ordinarily give of them”; and second, those who accept it are committed “to an implausible conception of moral perfection” (RM, 180). To substantiate the first objection, Smith tells a story of a friend (let’s call him Mike)

than the domain of WMI. But unlike WMI, the conceptual truth of MCI is not secured by way of a stipulative definition. Its source, I submit, is in our ordinary concept of commitment.

who has radically changed his moral view over the years from act-utilitarianism to a view that sanctions, in some instances, favoring family and friends, even when this cannot be given utilitarian justification. Since Mike is a moralist, his motivational dispositions have changed correspondingly. Smith charges that externalists are committed to saying that before his conversion, Mike “must really only have had an *instrumental* desire to maximize happiness and minimize suffering,” whereas after it, “he must really only have acquired *instrumental* desires to confer special benefits on his family and friends, at least in the first instance.” And he objects that this is “manifestly false, an utterly theory-driven redescription of my friend’s psychological change. . . . [M]y friend clearly has been transformed from being someone with one set of non-instrumental concerns into someone with another set” (RM, 181).

Before directly addressing Smith’s claims, I would like to offer an illustration of what sort of description externalists might give of Mike’s mental states before, during, and after his two moral conversions. I venture the following speculation: Mike has always had some inclination to favor family and friends, but at one point he developed strong inhibitions against acting on these inclinations. These inhibitions were largely the result of being convinced that act-utilitarianism specifies the correct criterion for moral rightness. Having a strong desire to do the right thing and a rigid temperament, Mike quickly developed an avid interest in maximizing total happiness in the world, taking the interest of each person equally into account. In due time, his desire to maximize happiness actually started to dominate all other desires to the point that his friends thought of him as a utilitarian monster. But slowly doubts started to emerge as a result of exposure to arguments against utilitarianism. By and by Mike’s conviction eroded and in the end he accepted a moral view according to which it is often right to be partial to family and friends, even when doing so cannot be given a utilitarian justification. At the same time, he came to see himself as a utilitarian monster, ever ready to sacrifice the interests of friends and family for the utilitarian project. Motivational dispositions he formerly took pride in having developed now became distasteful to him. However, since his desire to do the right thing has continued to be operative in his psyche, these dispositions are slowly eroding and the inhibitions on his inclinations to favor family and friends are undergoing radical change. They are gradually fall-

ing in line with his view of when it is right to give extra benefits to family and friends.

Although I might not have hit on the true description of the psychological development of Smith's friend, I submit that it is not revisionist, nor wildly implausible, nor an utterly theory-driven description of such a change. (True, the story is a bit implausible because it is difficult to believe that anyone is capable of becoming a utilitarian monster, but that is an artifact of Smith's claim about his friend that I have honored.) Indeed, the time-lag that my account allows between conviction and change in motivational dispositions makes it a much more familiar and plausible account of such a psychological change than the internalists'. Internalists will have to insist that, at the time of the first conversion, Mike is not fully convinced of utilitarianism until his motivational dispositions are in line with his utilitarian view. That is, I submit, an implausible and theory-driven description of moral conversion.

Now, let's consider the feature of the externalist description that Smith finds particularly problematic: it is that Mike's earlier desire to maximize happiness and later desire to confer special benefits on family and friends will be "instrumental" on this view "because, possessed as they are by a moralist, they must all have been derived from the one and only non-instrumental desire he has in virtue of being a moralist: a non-instrumental desire to do what is right that leads him to have various instrumental desires . . . when combined with his earlier and later means-end beliefs about how to do what is right" (RM, 181). Notice first that I have speculated that Mike did not need to develop the desire to confer special benefits on family and friends during his second moral conversion, but rather lost inhibitions on his longstanding inclinations to pay special attention to family and friends. Most of us need not be moralists with a moral view that sanctions partiality to friends and family in order to develop such motivations. If Mike is different, he is simply weird irrespective of whether an externalist or an internalist description is given of his psychological profile. But it is true that Mike's desire to maximize happiness is—in my speculative description of his moral development—derived from his desire to be moral in the sense that the latter desire, in conjunction with a belief in the utilitarian principle (notice this is not a means-end belief!), gave rise to the former. But this genesis of the desire to maximize happiness hardly justifies labeling it as instrumental.

To call a desire instrumental suggests that its only role in its possessor's psychology is to motivate him to do something that satisfies another desire of his. A paradigmatic example of such a desire is my desire to go to a certain store, which I believe sells Mare sandals, which I happen to desire. My desire to go to the store is derived from, in the sense explained above, the aforementioned belief-desire pair. But it also has a very tenuous standing in my psychology and a limited role to play there. Its only role is to get me to do what I take to be needed in order to satisfy my desire for Mare sandals. It is a caricature of my externalist account of Mike's motivations to take it to confer the same standing and role to his desire to maximize happiness. Even if his desire to be moral and his acceptance of act utilitarianism originally gave rise to this desire, it need not continue to function as completely subservient to his desire to be moral: Mike may so completely get into the project of maximizing happiness that he even has difficulties abandoning it when he sheds his utilitarian conviction.

If you doubt that there is a legitimate distinction to be drawn here consider the story of Mary, whose desire to be strictly vegetarian originally developed because she believed it was necessary for her health and wanted to be healthy, but who then got into a strict vegetarian diet for its own sake. A few years later, Mary loses this belief and, moreover, becomes convinced that it would be better for her health to ease up on her dietary restrictions. However, having developed a taste for a vegetarian diet and lost her appetite for animal products, Mary's desire to be strictly vegetarian survives her change of mind. It is only because of her desire to stay healthy that her desire to be vegetarian slowly erodes and she finds herself by and by relaxing her dietary restrictions. Mary's story parallels closely the story I told about Mike's conversion, and I submit that it would be highly misleading to describe her desire to be strictly vegetarian as instrumental. The same applies to Mike's desire to maximize happiness on my account.

Smith may object that at least right after Mike's conversion to utilitarianism, this desire had merely an instrumental status in his psyche. Now, if by this Smith simply means that Mike's having his desire to maximize happiness (come what may) was *in the first instance* highly dependent on the desire to be moral and his belief in the utilitarian principle, I do not see why that is such an implausible portrayal of the psychology of someone's conversion to

utilitarianism. But notice that the status of this desire within Mike's psychology would still be very different from the status of the paradigmatic instrumental desire that I discussed above. In particular, its internal relation—the relation accessible from inside the first-person perspective—to the desire on which it depends is very different from the relation between my desire to go to the store and my desire for Mare sandals. In my case, the relation is forged via a belief whose content can fairly be construed as: *that my going to the store is a means to my acquiring Mare sandals*. But the content of Mike's belief in the act-utilitarian principle should *not* be construed as: *that maximizing happiness is the means to doing the right thing*. Rather, it is: *that maximizing happiness is doing the right thing*. When Mike acquires the desire to maximize happiness, he acquires the desire to do on a routine basis that which he in fact believes to *constitute* doing the right thing. It is true that this desire is instrumental in getting Mike to do what he believes to be the right thing and, hence, to do what by his lights satisfies his desire to be moral; but to refer to it as an instrumental desire misleadingly suggests that it gets Mike to do something because he believes it to be the means to satisfy his desire to be moral. Smith should have made clear what exactly he means when describing a desire as instrumental. It makes for a powerful rhetorical move to flag this description, but not for a solid argument against externalism.

Smith's second objection to my account of motivation is that it commits me to an implausible conception of moral perfection. Externalists, Smith claims, must maintain that *morally perfect agents* are to be found among the class of moralists with true moral beliefs (RM, 181–82). But then he adds, "We normally assume that [morally perfect] people are moved by the very features of their acts which make them right. . . . This is part of what makes them morally perfect: morally perfect people are moved by right-making features" (RM, 182). And he claims that externalists have to take the opposite view, that morally perfect people

must instead be moved by a feature that these right-making features possess: the feature of being a right-making feature. But externalists thereby make morally perfect people, in a quite straightforward sense, alienated from the features of their acts that make them right. In desiring to do what is right for the sake of its being the right thing to do, rather than for the sake of the feature that makes it the right thing to do, they desire something that is not of primary moral im-

portance. They seem precious, overly concerned with the moral standing of their acts when they should instead be concerned with the features in virtue of which their acts have the moral standing that they have. In a word, they do not seem to be morally perfect at all. As I say in *The Moral Problem*, externalists thus seem to give a moral fetish the status of the one and only moral virtue. (RM, 182–83)

Smith appears to be drawing our attention to judgments of the form “*a* is morally right because it is ϕ .” And his claim is that the externalist explanation of moral motivation makes moralists desire to do *a* because it is morally right rather than because it is ϕ , even if its being ϕ makes it have the moral status it has. He takes this to show that on the externalist account the moralist is preoccupied with something that is not of primary moral importance (the moral standing of their acts) and alienated from the thing of primary moral importance (the right-making feature). And since morally perfect people must be found among moralists rather than amor-alists, they have also to suffer from such an “alienation,” which Smith thinks is a *reductio ad absurdum* of externalism.

I stand by my earlier claim that moral fetishism is most appropriately thought of as the phenomenon of holding oneself and others to rigorous moral standards, while being completely unwilling to entertain any reflective question about their nature or grounds. This is not the imperfection that Smith takes externalists to ascribe to morally perfect agents. Indeed, it is difficult to see what exactly the imperfection in question is. It is hardly a moral shortcoming to be concerned about the moral status of one’s acts, for example, that one do the right rather than the wrong thing in a given situation. And it seems most peculiar to claim that someone thus concerned is preoccupied with something that is not of primary moral importance. Of course, it would be odd if a person thus concerned were not also concerned about doing ϕ rather than β , if she believed that ϕ is the relevant right-making feature and β the relevant wrong-making feature. But externalists would not deny that, and Smith is not really accusing them of doing so. Rather he is claiming that, on the externalist account, moralists would not be concerned about doing ϕ for its own sake, but rather for the reason that ϕ has the feature of being a right-making feature. And he takes that motivational disposition to be a moral shortcoming. Leaving aside this evaluative claim for the moment,

it should be clear by now that externalists are not committed to the above claim about the concerns of moralists.

As in the case of Mike's concern for his family and friends, a moralist may have had the concern for ϕ long before he came to think of ϕ as a right-making feature. In other cases, even if the origin of the concern for ϕ lies in the desire to be moral and the belief that ϕ is a right-making feature, the concern has taken on a life of its own. In moralists, the desire to be moral is operative *when they resort to moral thinking to guide their conduct*. But many moralists will be influenced by their moral view much more indirectly. Anyone with a moral commitment of a reasonable strength will develop a self-standing interest in ϕ , when ϕ is a feature they believe to be right-making. In many instances, they will be motivated to undertake an action they see as instantiating ϕ without conceiving of it as instantiating a right-making feature. True, externalism allows that there are some moralists who are not like this. But that is certainly a virtue of their account. There probably are warped (rational) people who have the limited concern with right-making features to which Smith is drawing our attention; our account of moral judgments should not rule this out as an impossibility. That is not the way to get rid of human imperfections.

The final step of Smith's argument appeals to a conception of moral perfection. It does not work, since it relies on Smith's mistaken assumption that, on the externalist account, the class of moralists contains only people incapable of being moved by the features they believe to be right-making features without thinking about them as such. However, this step is worth some comment, since it curiously invokes—despite Smith's disclaimer—an evaluative assumption to clinch his argument against externalism. I find Smith's claim about the motivations of morally perfect agents harder to assess than his earlier claim about the motivations of good people. This is because the notion of moral perfection is somewhat obscure. For example, I am unsure what the relation between the class of good persons and the class of morally perfect persons is supposed to be. I do not think that a good person has to have true moral beliefs (see note 55, above), so by my lights these classes are at least not coextensive. But does the latter class form a subcategory of the former? Is our conception of moral perfection supposed to be an extension of our conception of moral goodness, although so demanding that we cannot expect normal humans

ever to instantiate it? Or are we talking about an entirely different phenomenon? Are we perhaps talking about a person who does not need to be guided by moral judgments, someone whose motivational system latches on to the real right-making features and the aims at which morality properly aims without having to resort to moral reflection? Is that moral perfection? If the answer to the last two questions is yes, it cannot be used against the externalist account of moral motivation; for that is an account of how moral judgments motivate. And if the answer is no, I would think that the difference between the morally perfect agent (as well as many morally imperfect, but morally committed, people) and the person who is not motivationally affected by her moral conclusions has to be explained by ascribing the desire to be moral to the former and not the latter. If anything, the above discussion of the failure of Smith's *reductio* has strengthened my conviction that this is the most plausible account of moral motivation.

7. Further Doubts

My experience has been that the view that the desire to be moral underlies the motivational impact of moral judgments is surprisingly often met with a mixture of incredulity and hostility. The reaction sometimes is that it would somehow be undesirable, or would make those motivated on account of their moral judgments somehow less worthy, if this view of moral motivation were correct. This is, of course, Smith's point in a nutshell. His objection, however, is to the content of the desire externalists postulate to explain moral motivation. Others seem bothered just by the idea that underlying moral motivation is a *desire*, and prefer the idea of a motivationally loaded belief—a *besire*—as underlying moral motivation. I find it somewhat of a mystery why the merit of a person or her motives would in any way depend on whether the motivational impact of her moral judgment is best explained as being due to her having a motivationally loaded moral belief, or to her having a motivationally inert moral belief plus a desire to pursue moral value. And I find it equally mysterious to think that it is somehow undesirable, or bad, if the best explanation turns out to be along the latter lines.

There is, of course, a long tradition of drawing a sharp distinction between our rational side and our sensuous side. And it has

been the tenet of rationalism in ethics to establish moral evaluation and motivation as a manifestation of the former and not the latter. Moreover, at least Kantians seem to think that the dignity and authority of morality hangs on the truth of the rationalist thesis. I frankly doubt that a neat distinction can be drawn between our rational and sensuous natures. The attempt to draw this distinction seems to rest on a caricature of our desires and emotions as all being impulses on a par with physical attraction, hunger, thirst, and various cravings, impulses that we often experience as assailing and distracting us, and as reminders of our physical constitution and limitations. If I were employing such a notion of desire in my explanation of moral motivation, I would understand the above reaction. But it should be clear from the discussion that the notion of desire I am employing is that of an intentional state that grounds a disposition to be motivated to undertake actions under a certain description.⁶⁸ It would be distinguished from a belief with an identical content on the basis of the role each respective state plays in our mental lives. On this conception of beliefs and desires, it would be baffling to subtract moral merit from a person on the grounds that the motivational impact of her moral judgments is best explained by ascribing to her both a moral belief and a moral desire rather than merely a motivating moral belief.

But why think that the best explanation of moral motivation lies along these lines? Someone is bound to object that my inclination to go for the belief-desire explanation of moral motivation and conduct simply shows that I have been held hostage to the Humean dogma that all motivation has its root in a desire. I beg to differ. I have not relied on the Humean theory of motivation to support my externalist position. My contention that moral judgments only motivate when accompanied by a desire, with appropriately related content, has been supported by observations of patterns of variations in moral motivation and *not* on the assumption that desires must always be operative in motivation. I have left it open that there might be cases in which we would have no jus-

⁶⁸I would, indeed, advocate a conception of desire according to which a desire to ϕ not only grounds the motivational disposition to do ϕ under certain conditions, but also grounds the cognitive disposition of noticing things that bear on whether ϕ obtains and how ϕ can be brought about, as well as emotional dispositions such as being pleased upon doing ϕ and frustrated when one cannot do ϕ .

tification for tracing the motivating power of a type of judgment to an accompanying desire. But in such cases, we had better not encounter the kind of variation in motivation that suggests that the motivational impact of this type of judgment depends on something more than the making of the judgment itself.⁶⁹

Indeed, the kinds of observations we can make about how moral judgments affect people's practical decisions and conduct seem to provide a textbook case for illustrating the usefulness of distinguishing between beliefs and desires. Both are mental attitudes taken towards conceptualizations or modes of presentation of (actual or possible) objects or events; that is to say, they are both states that have conceptual content. And typically both are involved in the motivation of action. When the contents of a belief and a desire are appropriately related, they motivate behavior; in other words, an agent is motivated to undertake an action he believes will accomplish something, which he happens to desire. If one or the other is missing, there is a failure of motivation, but the repercussions for the overall conduct of the agent typically differ depending on which of them is missing. It is the usefulness in explaining patterns of difference and similarity in the reactions of various individuals—and of an individual over time—to the same conceptualizations of their environment and behavioral options that calls for maintaining a sharp distinction between beliefs and desires rather than some *aprioristic* or metaphysical intuition to the effect that these states are “distinct existences.”⁷⁰ For example, we can exploit such a distinction to explain why people in a similar position and with similar personal commitments may openly agree that it would be right to help the politically persecuted, while only some of them go ahead and try to do so. We may furthermore exploit this distinction to explain variations over time in the behavioral patterns of those who failed to help. It is not only that we can explain the difference between the moral cynic and the morally committed by ascribing the desire to be moral just to the latter.

⁶⁹And even so, considerations of systematicity and uniformity in our theoretical apparatus, used to make sense of each other, might dictate that we maintain the distinction between belief and desire even in such cases.

⁷⁰See McDowell's “Are Moral Requirements Hypothetical Imperatives?” for criticism of the Humean theory of motivation, which charges that it amounts to a mere dogma built on some suspect *aprioristic* or metaphysical intuition.

We can also explain why the motivational impact of moral judgments seems to vary among morally committed people such that in some people it is more easily overridden by contrary motivations than in others. It is because the strength of their desire to do what is morally valuable or required differs; this is what underlies differences in their degree of moral commitment.⁷¹

8. In Closing

I hope to have convinced the reader that, regarded as a thesis in motivational theory, internalism is not plausible or, at least, that it should be rejected as a constraint of adequacy on accounts of moral language and thought. Motivational internalism probably appeals to many because they think it is a way of rendering more precise the plausible—possibly platitudinous—claim that the point of moral evaluation is distinctively to guide conduct. The distinctive directional role of moral judgments is thus equated with a motivational role. This is too hasty. It is entirely unclear what follows from the rather vague claim concerning the point of moral evaluation. It certainly suggests that the primary use of the conceptual resources involved in moral evaluation is to enable us to think through our circumstances and behavioral options in a distinctive way with an eye to deciding how to act. But it is far from clear whether this means that a competent exercise of the concepts involved is tied to our conative states or sentiments in such a way that sincere moral judgments are intrinsically or necessarily motivating. Indeed, even if our primary interests in employing these conceptual resources were echoed in some of the constraints to which they are subject (as I am inclined to think), it need not follow that it is impossible to exploit these resources in sincere, literal, and competent judgments without taking on these interests or having the underlying concerns. And even if it were necessary to have a rather subtle understanding of these concerns, it is not thereby said that one has to have them. Maybe the presence of these conceptual resources in humans is due to our having certain abiding concerns; and maybe the nature of these concepts reflects

⁷¹I would also think that variations in some of their cognitive and emotional dispositions can be explained in a similar way. See note 68 for my conception of desire.

to some extent the nature of these concerns; and maybe the mastery of these concepts requires, or at least is aided by, understanding of these concerns. But nothing of this implies that, in order to use these conceptual resources competently, an *individual* has to be committed to taking sincere judgments employing these concepts as his guide to conduct. Nor does it imply that the application conditions of these concepts are somehow tied to expressions of the concerns in virtue of which we typically have an interest in employing them.

These are deep issues, ones I hope to pursue in the future. I suspect that there is a long and complicated story to be told about the relation between the distinctive character of our moral concepts (and hence moral beliefs) and our being emotional creatures of a particular sort. But I doubt that the relation will be as direct as internalists have it. However, I do not exclude in advance the possibility that the only way of making overall sense of moral thought and language requires that I and other externalists give up our conviction. What I have mainly been resisting in this paper is the claim that the internalist thesis amounts to a platitude that should function as a constraint of adequacy on accounts of moral judgments, moral concepts, and the semantics of moral language.

New York University