PHILOSOPHICAL TOPICS
VOL. 38, NO. 1, SPRING 2010

Does Moral Theory Corrupt Youth?

Kieran Setiya
University of Pittsburgh

ABSTRACT: I argue that the answer is yes. The epistemic assumptions of moral theory deprive us of resources needed to resist the challenge of moral disagreement, which its practice at the same time makes vivid. The paper ends by sketching a kind of epistemology that can respond to disagreement without skepticism: one in which the fundamental standards of justification for moral belief are biased toward the truth.

My title is adapted from Elizabeth Anscombe's infamous 1957 radio talk, "Does Oxford Moral Philosophy Corrupt Youth?" Her answer was no: it merely reiterates their already depraved opinions. Unlike Anscombe, I will not be concerned with such "point[s] of method" as the advice to "concentrate on examples which are either banal […] or fantastic" (Anscombe 1957, 162–63), or with what she called "consequentialism" (in Anscombe 1958). We begin, instead, with some remarks by Annette Baier, in "Theory and Reflective Practices," that appear to answer yes: moral theory does corrupt youth.

> The obvious trouble with our contemporary attempts to use moral theory to guide action is the lack of agreement on which theory to apply. The standard undergraduate course in, say, medical ethics, or business ethics, acquaints the student with a variety of theories, and shows the difference in the guidance they give. We, in effect, give courses in comparative moral

> theory, and like courses in comparative religion, their usual effect on the student is loss of faith in *any* of the alternatives presented. We produce relativists and moral skeptics, persons who have been convinced by our teaching that whatever they do in some difficult situation, some moral theory will condone it, another will condemn it. The usual, and the sensible, reaction to this confrontation with a variety of conflicting theories, all apparently having some plausibility and respectable credentials, is to turn to a guide that speaks more univocally, to turn from morality to self-interest, or mere convenience. […] In attempting to increase moral reflectiveness we may be destroying what conscience there once was in those we teach. (Baier 1985, 207–8)

This passage may seem to contain some not-so-harmless exaggeration. To begin with the coldest comfort, we are entitled to wonder whether students of medical or business ethics always take it quite so seriously. Whatever effect I have on the undergraduates to whom I teach introductory ethics, I doubt that many of them are turned "from morality to self-interest or mere convenience." Nor is my influence compelling enough to destroy whatever conscience they have.

As I said, the comfort here is cold. More reassuring, perhaps, is that the existence of profound disagreement about morality is not the fault of moral theorists. It is evident to anyone who reads the newspaper, or has an elementary knowledge of history. If moral disagreement makes an epistemic problem, we might reply, it does so anyway, and moral theorists cannot be blamed. It is in any case unclear just what the problem is meant to be. Baier's remorse may attach to nothing more than a psychological fact: that confrontation with disagreement withers confidence. This "fact" may not obtain; but even if it does, how would it show that loss of confidence is, in Baier's words, not only "usual" but "sensible"? Why is she is so willing to concede that students are *justified* in becoming moral skeptics on the basis of their education? Finally, despite some suggestive remarks, Baier operates with an insufficiently examined notion of "moral theory." What exactly is it that moral philosophers are doing that they should not? What would a properly anti-theoretical moral philosophy be?

We have, then, three questions to ask. What is meant by "moral theory" in this context? Why suppose that it generates an epistemic problem of disagreement that is distinct from, or worse than, the problem we would otherwise face? And is there any way to avoid moral theory without sheer anti-intellectualism?

In what follows, I argue that there is a tempting and prevalent conception of moral theory and its method that generates exactly the problem that Baier fears. It threatens to corrupt youth in that its epistemology deprives them of resources needed to resist the epistemic challenge of moral disagreement, which its practice at the same time makes vivid. I prosecute this charge in three stages: in section 1, an account of moral theory as it figures in the indictment; in section 2, the problem of disagreement for moral theorists; and in section 3, a solution to this problem that moral theorists cannot give.

I. WHAT IS MORAL THEORY?

Baier is not the only recent philosopher to set herself against moral theory. Despite attempts to find unity, however, the work classified as "anti-theoretical" is exceptionally diverse, and its targets rarely well defined. What common cause joins *The Sovereignty of Good*, *Ethics and the Limits of Philosophy*, and *After Virtue*?[1]

There are dangers in trying to define moral theory against the backdrop of such diversity—for instance, that we will end up with something too ambitious, a characterization that would-be moral theorists can simply disavow.[2] If a moral theory is a procedure for deciding moral questions without the need for judgment or moral sensitivity, there may be no such thing; but who ever thought otherwise? What we need is an account of moral theory that is sufficiently determinate to be the object of critical attention, but one whose aspirations do not seem quixotic. It is inevitable that the critique of moral theory, once defined, will not apply to everything that self-styled "moral theorists" do. The most we can hope is that it will apply to an enterprise that is both recognizable and influential.

The target of the present essay, in brief, is the kind of theory that takes our moral intuitions seriously as starting points and aims to produce a systematic body of principles that vindicates these intuitions by endorsing them, undermines them by failing to do so, and yields justified claims where they are silent. This description needs refinement, but it should be more or less familiar. It is meant to characterize the work of act-consequentialists like Shelly Kagan (1989), rule-consequentialists like Brad Hooker (2000), non-consequentialists like Samuel Scheffler (1982), and virtue theorists like Michael Slote (1992, 2001).

While there are variations within the family of moral theorists, as for instance in their conception of moral intuitions and the nature and degree of deference owed to them—matters to be examined shortly—they share an epistemic commitment to theoretical virtues of simplicity, power, consistency, and explanatory depth.[3] The aspiration of moral thought, for moral theorists, is the construction of a relatively simple consistent body of moral principles on the basis of which we can justify a wide range of verdicts about particular cases. "Justify" may but need not mean "deduce": it is a matter of explaining why certain verdicts or subsidiary principles are correct, and an answer to the question "why?" does not always take the form of a valid proof.[4] Explanatory depth in moral theory is measured by the extent to which it provides such justifications. Even though the principles that justify are no different in kind from the principles and verdicts justified, in that they are further moral claims, and even though justifications must come to an end, the point at which they do so can be more or less superficial.[5] It counts in favor of a moral theory, other things being equal, that its purported justifications are deep.

In what follows, I use the term "coherence" as shorthand for the theoretical virtues described above, passing over differences in the weight assigned to each of them by different moral theorists. In the limiting case, a theorist might give very

minimal weight to simplicity or power, accepting complex explanations without attempting to extend them into novel ground.[6] Still, even theorists of this kind try to avoid gratuitous complexity. More commonly, appeals to simplicity and power are allowed to put pressure on our intuitive beliefs. We could illustrate this phenomenon with various examples, of which I give three. The first is both simple and familiar: even critics of act-utilitarianism often concede that its elegance and scope are attractive, that they are reasons to accept it even if, on balance, the cost to moral intuition is too high. The second example is more complicated: debates about the allegedly paradoxical character of "agent-centered restrictions"—roughly, moral principles that forbid us from acting in certain ways even to prevent more actions of the very same kind—rest on the apparent difficulty of assimilating these restrictions to a coherent system of principles. What is the rationale for prohibiting action of a certain kind if not to minimize its occurrence, which agent-centered restrictions perversely rule out?[7] Disputes about this question, even those which stress the force of moral intuition, typically give weight to considerations of explanatory power. Finally, a less familiar example: in proposing an account of right action as action motivated by moral virtue, Michael Slote (2001, 28) argues, on grounds of simplicity or theoretical unification, that benevolence is the only virtue there is. His argument is meant to provide defeasible support for the reduction of other putative virtues, like justice or courage, to "good or virtuous motivation involving benevolence or caring" (Slote 2001, 38). If this seems odd to you, as it does to me, you should ask how it differs from claims about the appeal of act-utilitarianism that many are given to accept.

Although moral theorists agree that coherence is epistemically significant in moral thought, and in the method of building our moral intuitions into an adequate theory, they disagree about what moral intuitions are and why we should defer to them. Broadly speaking, there are two views about this. The first treats moral intuitions as nothing more than our initial moral beliefs and advocates an epistemology of pure coherence: one's moral convictions are justified insofar as they belong to a system of beliefs that is simple, powerful, consistent and explanatorily deep.[8] The answer to the question "Why defer to moral intuitions?" is that coherence is the only epistemic pressure we face: only a failure of coherence can require us to reject our initial moral beliefs, and what we should believe is a function of them.

The second view is more ambitious about the analogy between moral thinking and empirical knowledge. It treats moral intuitions as a matter of how things seem, morally speaking—where such "seemings" are cognitive states distinct from beliefs—and gives them the epistemic role of perceptual appearances. Thomas Nagel defends this idea in *The View from Nowhere*:[9]

> In physics, one infers from factual appearances to their most plausible explanation in a theory of how the world is. In ethics, one infers from appearances of value to their most plausible explanation in a theory of what there is reason to do or want. […] If we start by regarding the

> appearances of value as appearances of something, and then step back to form hypotheses about the broader system of motivational possibilities of which we have had a glimpse, the result is a gradual opening out of a complex domain which we apparently discover. The method of discovery is to seek the best normative explanation of the normative appearances. (Nagel 1986, 146)

On what I will call the empirical model, one's moral intuitions constitute evidence for one's moral beliefs in much the way that perceptual appearances constitute evidence for beliefs about the physical world, and the standards of coherence figure in determining what this evidence supports. Again, there will be variations here. How far can the exigencies of coherence require us to reject the appearances? How are the elements of coherence to be balanced together? For the moment, however, we need not say more. Moral theory can be defined roughly, and disjunctively, as endorsing either pure coherence in moral epistemology or the empirical model. Either way, "we need theory in ethics […] and theoretical virtues like simplicity and unifying power have some weight in deciding what kind of ethical view to adopt" (Slote 2001, 10).

II. HOW MORAL THEORY CORRUPTS YOUTH

Moral theory can seem innocent to the point of inevitability. Against this, I will argue that moral theorists cannot give an adequate epistemology of disagreement. The reasons for this are several, but they have a common source: a failure to find sufficient asymmetry in the situations of those who disagree. In this section, I press the problem of disagreement against the two kinds of moral theory in turn.

This undertaking rests on assumptions about the interpretation and metaphysics of moral judgment, but they are relatively modest: that we can speak of moral propositions as being true or false; that moral beliefs can be epistemically justified or not; and that in cases of apparent disagreement about morality, even between members of quite different communities, at least one party must be mistaken—though not, perhaps, irrational—in her beliefs. Though disputable, these assumptions are shared by the moral theorists I am arguing against.

Begin with the empirical model. In the background is an abstract epistemic picture, according to which one's degrees of belief should be proportioned to one's evidence; and one's evidence is identified with a set of propositions or mental states to which one bears some special relation. The indeterminacies of this picture could be resolved in various ways. As it stands, it does not say that one should be confident of *p* only if one has evidence of its truth; there is room for nonevidentially justified beliefs. Nor does it entail "objective Bayesianism," according to which one's degrees of belief should match permissible prior probabilities conditionalized on evidence propositions. It assumes very little about the nature of evidence and evidential support. What the empirical model adds to this abstract picture is that, for

moral beliefs, one's evidence is ultimately supplied by moral intuitions, intellectual seemings that play the epistemic role of perceptual appearances, and that the "proportions" that one's degrees of belief should bear to these appearances are fixed by permissible standards of coherence and of deference to how things seem. There is a complication here, that some moral beliefs, and perhaps some moral intuitions, rest on nonmoral beliefs, as when I think you acted wrongly partly on the ground that you killed my friend. If the latter belief is unjustified, that will tend to undermine the former. Still, on the empirical model, this is the only way in which nonmoral evidence is relevant to moral epistemology. If our nonmoral beliefs are similar, and similarly justified, they can simply be ignored. Holding these points fixed, one's degrees of moral belief should be those best proportioned to one's moral intuitions.

The problem for the empirical model can be seen by reflecting on recent work in the epistemology of disagreement. This work is sharply divided about the extent to which one should defer to "epistemic peers" whose beliefs or degrees of belief are different from one's own. In particular, there are serious objections to the so-called Equal Weight View on which one should give the same weight to others' opinions unless one has more independent evidence of one's own reliability.[10] Still, even those who reject such deference as a general policy concede that it is the best response to cases of *perceptual* disagreement. To take an example from Tom Kelly (2010, §5.1), imagine that you and I are watching two horses cross a finish line. It looks to me as though Horse A finished ahead of Horse B, but it looks the opposite way to you. Before we talk, I am justified in believing that Horse A won and you are justified in believing the same about Horse B. But if we compare notes, and we come to know how things seemed to our interlocutor, our confidence should fade. This verdict does not depend on the Equal Weight View and its attendant controversies but on the way in which our evidence has changed:

> I have gained evidence that suggests that Horse B won the race, while you have gained evidence that Horse A won the race. Moreover, given the relevant background assumptions and symmetries, it is natural to think that the total evidence that we now share favors neither the proposition that Horse A finished ahead of Horse B nor the proposition that Horse B finished ahead of Horse A. (Kelly 2010, §5.1)

If I have no independent reason to discount your perceptions, or to think that they are less reliable, I should give how things look to you as much evidential weight as how things look to me.[11] This is consistent with holding my line, if I have evidence that Horse A won the race, or that Horse B lost, apart from how things look. But, as it happens, that is not the case. Abstracting from asymmetries of self-knowledge—it may be easier for me to know how things look to me than how they look to you—my total evidence requires me to suspend belief.[12]

How does this bear on the empirical model in moral epistemology? It shows that the model is subject to a skeptical problem of disagreement. We can see this in two stages. First: not all significant moral conflict can be traced to differences in

nonmoral belief or in our perhaps implicit standards of coherence and deference to moral intuition. It is not that things seem the same way to all of us, morally speaking, and that we merely differ in the beliefs we derive from the appearances, as different scientists might draw different conclusions from the same observational evidence. Instead, our moral intuitions vary widely, so that serious moral disagreement does and would survive agreement on the nonmoral facts, and in how to proportion moral beliefs to the evidence that intuition provides. Even if we were perfectly coherent by the lights of the empirical model, with systems of moral belief that are simple, powerful, consistent, and explanatorily deep, ideally proportioned to our different moral intuitions, we would continue to disagree. Second: if moral intuitions play the epistemic role of perceptual appearances and provide us with evidence in a similar way, the horse race argument applies. Before they discover their disagreement in moral outlook, those whose intuitions differ widely may well have justified moral beliefs. But when they confront one another, and come to know that things seem different to others, their confidence should fade. With no more evidence of their own reliability, apart from the intuitions in dispute, they should incorporate their interlocutors' intuitions into their total evidence. The likely effect is that, abstracting from asymmetries of self-knowledge, their evidence will no longer support their previous views.[13]

The foregoing remarks are inevitably schematic. On the empirical model, what I should believe in the face of moral disagreement turns on the extent to which others have conflicting moral intuitions, and how far our disagreements can be explained by differences in nonmoral belief and in failure to conform to ideal standards of coherence and deference to intuition. These are epistemic and sociological questions of enormous complexity. It is speculative to claim, as I do, that things seem sufficiently different to others—past and present—as to undermine our moral confidence if we give those appearances the same weight as our own. But the objection to the empirical model does not, in the end, require this. Consider instead a hypothetical disagreement. You belong to an utterly homogeneous moral community whose moral intuitions are those you actually have, and whose moral beliefs are proportioned to them by permissible standards of coherence and deference. Your nonmoral beliefs are also well supported by nonmoral evidence. For the first time, you meet someone from another community. He agrees with you about the nonmoral facts, but his moral intuitions are shocking. Fill in the details accordingly. Perhaps it seems to him that children are property and can be sold or given away, that women have no moral standing whatsoever, or that one should always act so as to maximize aggregate happiness even if many are trampled along the way. For more dramatic possibilities, we can turn from morality to practical reason. Perhaps it seems to him that what you call "justice" is a farce and that one should be exclusively selfish, or act so as to satisfy one's final desires, whatever they are. Despite this, his ethical beliefs are as well proportioned to his intuitions as your beliefs are to yours.[14] It turns out that he, too, belongs to a homogenous community, as extensive as your own. What should you now believe? On the empirical

model, you must become agnostic about the questions on which you disagree, since apart from the intuitions that conflict, you have no more evidence of your reliability than his. But this is not the right response. We should not defer to moral monsters but condemn them, however coherent and numerous they are. The empirical model assigns the wrong weight to moral intuitions in giving them the epistemic role of perceptual appearances.

These arguments against the empirical model—that it threatens skepticism in the face of actual disagreement, and that it gives the wrong verdict in cases of hypothetical disagreement—prompt a natural response. They take the analogy between moral intuitions and perceptual appearances seriously and ask what we should believe in perceptual disagreements of the relevant kind. The results are then applied to moral disagreements, conceived on the empirical model. Advocates of this model may protest that they did not intend the analogy so literally. Apart from a limited role for nonmoral evidence, their claim is that one's moral beliefs should be proportioned to the evidence supplied by one's moral intuitions, where the proportions are fixed by permissible standards of coherence and deference. Nothing follows about the epistemic significance of others' intuitions, even if one has no evidence, apart from one's own intuitions, about their relative reliability. If we refuse to give weight, or equal weight, to the intuitions of others, we can block the consequences lamented above.

If we take this path, however, we not only compromise the analogy between moral intuitions and perceptual appearances on which the empirical model drew, but fall into the trap of *epistemic egoism*. In order to resist the problem of disagreement and the corresponding threat of skepticism, the revised empirical model must say that, when my intuitions conflict dramatically with yours, I am justified in sustaining most of the beliefs that rest on them, even if, apart from that conflict, I have no more evidence of my reliability than yours. (Since the situation is symmetric, you are equally entitled to your beliefs, in the corresponding predicament.) But now we can invoke a qualified "reflection" principle for epistemic justification. To a first approximation, if A knows that he would be justified in believing *p* in circumstance *q*, then he is already justified in being conditionally confident of *p* given *q*.[15] This cannot be exactly right. We need to allow for the case in which A also knows, or has reason to suspect, that his evidence in *q* would be misleading, so that he would be justified in believing *p* even though it might very well be false. Roughly speaking, to say that A's evidence would be misleading is to say that if he were in *q*, there would be some fact about his situation knowledge of which would defeat or undermine his justification for believing *p*. No doubt further refinement is necessary, but we can work with this second approximation:

> If S knows that he would be justified in believing *p* in circumstance *q*, then he is already justified in being conditionally confident of *p* given *q*—so long as he has no reason to think that his evidence in *q* would be misleading.

In order to avoid complications about evidential defeat and undermining, we will ask what I would be justified in believing apart from the evidence given by my actual intuitions. (On the revised empirical model, they threaten to defeat or undermine the justification supplied by other intuitions I might have had.) To make this question vivid, imagine that I am reflecting on morality before I have moral intuitions, like some prodigious two-year-old epistemologist. This need not be possible, of course; the point is that we are interested in antecedent justification. Suppose, then, that I know the truth of the revised empirical model, and that my knowledge does not rest on moral intuitions. In particular, I know that I will be justified in believing the moral outlook my intuitions support, whatever it is, even if things seem very different to you and I have, apart from this conflict, no more evidence of my reliability than yours. It follows by the reflection principle that I am already justified in being conditionally confident of any moral outlook that my intuitions support, even if yours do not. In other words, quite apart from the evidence given by my actual intuitions, I am justified in believing that the moral outlook my intuitions support is more likely to be correct than the moral outlook supported by yours: that they are a more reliable guide to the moral facts. And since the situation is symmetric, you are justified in believing the same about yours.

This consequence is intolerable. Perhaps it is true that we are *a priori* justified in trusting the reliability of our perceptual and intellectual faculties, but not that, apart from the evidence they supply, we are justified in believing that our faculties are more reliable than others', should their outputs diverge. Such comparative confidence in our own capacities, regardless of what they tell us, amounts to epistemic egoism. In the absence of evidence, we are not entitled to believe that how things seem to us is a better guide to how they are than how they seem to anyone else.

The empirical model in moral epistemology thus confronts a dilemma. If it treats moral intuitions like perceptual appearances, it faces a problem of moral disagreement. If it averts this problem by insisting on an asymmetric treatment of our own and others' intuitions along the lines explored above, it leads to epistemic egoism. Conclusion: the empirical model does not allow for a plausible, nonskeptical response to moral disagreement.

The moral theorist's alternative is a pure coherence view, which drops the idea of moral intuitions as intellectual seemings. Instead, it begins with our moral beliefs, and claims that they are justified to the extent that they form a coherent system, where coherence is a matter of simplicity, power, consistency, and explanatory depth. As with the empirical model, there is a complication when moral beliefs rest on nonmoral beliefs that are not supported by nonmoral evidence. But again, we can set this aside. If our nonmoral beliefs are similar, and similarly justified, they can be ignored. One's degrees of moral belief should be those achieved by revising one's initial beliefs to the limit of coherence. (Variations are possible, as for instance those that give special weight to a subset of moral beliefs, one's "considered judgments.") Since the beliefs of others do not have the epistemic standing of one's

own—they count as evidence only when one has evidence of their reliability—one need not be disturbed by the fact of disagreement, as such.

> In the end, whether one is justified in retaining one's original view in light of another depends on whether one's own evidence tells in favor of the other view or not. In the face of (even) coherent alternatives, one justifiably rejects the others, when one does, on the basis of what one justifiably believes. […] This means, of course, that had one's initial beliefs been different, had one believed one thing rather than another, one would have justifiably rejected the views that one actually (and with justification) accepts. (Sayre-McCord 1996, 172)

By the same token, one's interlocutor is justified in his beliefs, however different they are from yours, to the extent that they are coherent; the situation is symmetric. Since one's beliefs are justified if coherent, the bare existence of someone with a coherent and therefore justified alternative theory is epistemically harmless.

No doubt the specifics of the pure coherence view need work.[16] What matters here is that its solution to the problem of disagreement echoes that of the revised empirical model, and like that model it buys its refusal of skepticism at the cost of epistemic egoism. Faced with the problem of disagreement, the pure coherence view maintains that I am justified in sustaining my moral beliefs insofar as they belong to a coherent system of beliefs, even if I know that your beliefs are very different from mine and, apart from those differences, I have no more evidence of my reliability than yours. But now recall the qualified reflection principle for conditional credence:

> If S knows that he would be justified in believing *p* in circumstance *q*, then he is already justified in being conditionally confident of *p* given *q*—so long as he has no reason to think that his evidence in *q* would be misleading.

To avoid complications about evidential defeat and undermining, we ask what I would be justified in believing apart from the evidence given by my actual beliefs. We can make this vivid if we imagine that I am reflecting on morality before I have beliefs about its content. (As before, this need not be possible; the point is to focus on antecedent justification.) If I know the truth of the pure coherence view, I know that my moral beliefs will be justified, so long as they are coherent, even if your equally coherent beliefs are different and I have, apart from those differences, no more evidence of my reliability than yours. It follows by the reflection principle that I am already justified in being conditionally confident of any moral outlook, given that it is coherently believed by me but not you. In other words, even apart from the evidence supplied by my actual beliefs, I am justified in believing that my moral outlook is more likely to be correct than yours, if both are equally coherent: that my beliefs are a more reliable guide to the moral facts. And since the situation is symmetric, you are justified in believing the same about yours. In resisting the challenge of moral disagreement, the moral theorist is forced, once again, into epistemic egoism.[17]

We might ask how this argument differs from the standard objection that epistemic coherence theories fail to connect justification with truth. After all, that objection is often pressed by noting the apparent possibility of distinct and conflicting systems of coherent belief, most of which would have to be false. How could it be anything more than luck that my beliefs are not just coherent but correct? The answer lies in a contrast already drawn, between the proposition that our basic faculties are reliable and the proposition that ours are more reliable than others' when their outputs diverge. A principle of interpretive charity on which our beliefs necessarily tend toward the truth, or on which they tend to count as knowledge, would begin to account for our reliability as something more than accidental, and perhaps give *a priori* grounds for trusting our perceptual and intellectual faculties.[18] In doing so, it would help to dissolve the standard objection to epistemologies of pure coherence. It would, however, do nothing at all to explain how each of us could be justified, without evidence, in taking ourselves to be more reliable than others when our moral intuitions or beliefs diverge.

These arguments put us, finally, in a position to say why and how moral theory corrupts youth. It does so by tacitly or explicitly invoking epistemic theories—the empirical model, perhaps revised; the pure coherence view—on which it is impossible to block the skeptical problem of moral disagreement without appeal to unacceptable forms of epistemic egoism. If our pedagogy instructs our students in one or other of these epistemologies, their loss of faith when confronted with a diversity of moral intuitions and moral outlooks in *Ethics 101* is, as Baier claimed, not only predictable but sensible. Resisting or failing to see the temptations of egoism in epistemology, they follow our epistemic standards into paralyzing doubt.

III. WAY OUT

Before the prosecution rests, we need to consider a possible gap in the case against moral theory. The verdict of corruption is premature until we have at least a rough conception of how to respond to moral disagreement without skepticism. If the skeptical conclusion is true, it is no defect in the moral theorists' epistemology that it respects this fact.

Here we must ask, in abstract terms, what justifies moral beliefs. In doing so, we adopt the mild evidentialism that figured in the background of the empirical model, above. According to this framework, there is a relation one can bear to certain propositions or psychological states, which thereby constitute one's evidence, and to which one's degrees of belief should be proportioned. Assuming that evidence can be classified as moral or nonmoral, the possibilities are these:

(i) Certain attitudes with moral content provide us with evidence, to which our moral beliefs should be proportioned; though, in order to avoid the

problems of section 2, the attitudes cannot be identified with mere beliefs or with moral intuitions as intellectual seemings.

(ii) Our nonmoral evidence supports some moral beliefs in contrast to others: that is, the proportions by which one's beliefs should track one's evidence are not a matter of simple coherence, but of accepting the moral conclusions made probable by nonmoral claims about one's circumstance.

(iii) Some moral beliefs are epistemically, but nonevidentially justified: our justification for holding them does not rest on evidence of their truth.[19]

The questions raised by these ideas are too large to examine in detail here. But we can make some preliminary notes.

In many ways, the simplest position to develop is (ii). According to a plausible doctrine of supervenience, each moral proposition is entailed by a series of nonmoral descriptions to whose disjunction it is strictly equivalent. The proponent of (ii) may hold that such descriptions give conclusive evidence of the corresponding moral claim. For the most part, however, we make moral judgements with only partial knowledge of the nonmoral circumstance and our ignorance or conjecture extends to matters that would make a moral difference. In that case, our evidence is defeasible: it supports *m* just to the extent that it supports the disjunction of nonmoral propositions that is equivalent to *m*. Knowing right from wrong on the basis of such defeasible evidence requires a capacity to tell when other facts about the circumstance are likely to be relevant. And that depends on knowledge of the nonmoral workings of the world.[20]

There is scope for an argument that takes us from (ii) to (iii). Suppose that *N* is nonmoral evidence for moral proposition, *m*, and that I am justified in believing *m* on the basis of *N*. According to a natural though controversial principle, I must have antecedent justification to believe that if *N*, *m*.[21] If this belief is justified by further evidence, *N'*, we can apply the principle again. At some point, we are bound to find a moral claim that is epistemically justified without evidence, as in (iii). This would be a form of intuitionism very different from the ones that were criticized in section 2.[22]

Finally, it might be argued that nonevidentially justified beliefs provide us with evidence. It is not clear to me that this argument is sound; but it also is not clear to me how to articulate a more plausible version of (i). For instance, it does not help to suggest, with Williamson (2000, ch. 8), that the attitude of *knowledge* provides us with evidence, even when its content is a moral proposition. For we can still ask how, or on what grounds, a moral proposition is known. Barring circularity or regress, we must eventually appeal to moral knowledge that rests on some other source of moral evidence, on nonmoral evidence, or on no evidence at all.

These remarks suggest that the options distinguished above, as (i) to (iii), may not be as various as they seem. In light of their convergence, we can say what kind of epistemology would solve the problem of moral disagreement. Focusing on (ii)

and (iii), it would be an epistemology in which the fundamental standards of justification—either of proportioning beliefs to nonmoral evidence, or of believing without evidence—favor some moral outlooks over others. In particular, they favor moral outlooks that apply moral concepts correctly on the basis of nonmoral facts. Moral epistemology is biased toward the truth.[23]

The situation of those who confront disagreement is thus potentially asymmetric. In the hypothetical case considered in section 2, we should not defer to the advocates of slavery, misogyny, or injustice, however coherent they are, if the nonmoral evidence in fact supports our beliefs, or if those beliefs are nonevidentially justified. By contrast, they should never have held their awful opinions, nor should they do so now. It matters who is wrong and who is right.[24]

This conclusion about what is involved in resisting the challenge of moral disagreement helps to motivate and clarify some obscure images in moral philosophy: that the standards of justification for moral belief are "internal" to morality; that there is no scope for "external" critique.[25] Such claims are hard to interpret and easy in read in ways that make them seem implausible or banal—for instance, as endorsements of a pure coherence view, as conflating moral and epistemic reasons, or as reminding us that reflection must begin with our own beliefs.[26] The substantive truth in the vicinity is that moral education is education not only into the space of moral reasons but into the space of reasons for and against moral beliefs. As McDowell writes,

> Ethical thinking is local in two ways: first, its characteristic concepts are not intelligible independently of particular cultural perspectives; and, second, it aims (explicitly or implicitly) to be directed by standards of good and bad argument, and the standards available to it are not independent of its own substantive and disputable conclusions. (McDowell 1986, 380)

The deepest mistake of the moral theorist is the assumption that moral thought has the same epistemology as empirical science, that the fundamental standards of epistemic justification are topic-neutral.[27] This is what we have to deny in order to explain how we can be justified in maintaining our beliefs in the face of radical disagreement. We must hold that, at the most basic level, nonmoral evidence supports particular moral beliefs—ones that tend to be correct—or that such beliefs are justified without evidence. Of course, there is no guarantee that we are in the right. Perhaps our interlocutors' beliefs are justified, while ours are not. We have no way to address that question that is independent of whether their beliefs are true. But so it goes. There are no guarantees in the epistemology of any beliefs. We do the best we can.[28]

Despite their schematic character, these reflections afford some purchase on the question that inspired my interest in moral theory, and to which I return in closing: what is the role of *coherence* in moral epistemology? Granting the significance of outright inconsistency as an argument against a moral outlook, do structural features like simplicity, power, and explanatory depth count as arguments in

favor? This is the fallback position for the moral theorist: although it needs to be supplemented, and not by the moral intuitions of the empirical model, coherence is part of what matters in the justification of moral beliefs. If that is so, we can continue to argue for act-utilitarianism by appeal to its simplicity, even if the argument is not decisive; and we can argue against "common sense morality" on the ground that it is an incoherent, superficial, complicated mess. Once we reject the empirical model and the pure coherence view, however, what is the motivation for giving coherence even this subsidiary role? If the analogy between moral thinking and scientific theory has failed, we can ask, with Bernard Williams (1985, 106), "Why should theoretical simplicity and its criteria be appropriate?" In particular, having acknowledged that the epistemic standards that govern moral belief favor certain moral outlooks, and in that sense have moral content, it is not clear that we need coherence to pick up the epistemic slack. Although I have not argued directly against coherence (in the technical sense: simplicity, power, explanatory depth) as an epistemic virtue of moral thought, I hope that the course of this paper presents it in a new light. The idea that we should defer to coherence in moral thought is not the inevitable application of rational standards that are everywhere at home, but a questionable vision of the shape of morality and moral virtue.

It is, I think, this vision that animates depictions of the moral philosopher as moral expert. In part, such depictions may rest on nothing more than the philosopher's luxury of time to think and to investigate the nonmoral facts, or her relative immunity to the unsound arguments that explain at least some of our moral opinions.[29] But they may also rest on the idea that thinking well about morality is thinking coherently (again, in the technical sense) and that philosophical training nurtures one's aptitude for this. If we doubt that coherence has great significance in moral epistemology, this reasoning will seem misguided. Philosophical education is not a source of moral wisdom, except in indirect ways. We can then accept what is at least the evidence of my experience, that philosophers are not much better at knowing how to live than anyone else.[30]

ACKNOWLEDGMENTS

For reactions to this material in earlier forms I am grateful to audiences at the New York Institute of Philosophy, the University of Pittsburgh, the University of Missouri at St. Louis, the University of Miami, the University of Wisconsin at Milwaukee, Rutgers University, the University of Texas, Johns Hopkins University, and the University of Illinois, Urbana-Champaign. Special thanks to Cian Dorr, Mike Otsuka, Karl Schafer, David Sosa, and Brian Weatherson.

NOTES

1. Respectively: Murdoch 1970, Williams 1985, and MacIntyre 1981. For the claim of unity, see Clarke and Simpson 1989.
2. The strategy of disavowal is pursued at length in Louden 1992.
3. For this list, see Kagan 1989, 11–13.
4. Think, for instance, of the rationale for an "agent-centered prerogative" in Scheffler 1992, ch. 3, which is not supposed to meet the deductive standard.
5. Again, see Kagan 1989, 14; also Scheffler 1982, 112.
6. That might be true of Kamm 2007.
7. Classic sources for the "paradox" are Nozick 1974, ch. 3, and Scheffler 1982, 82–83, 87–88 and ch. 4, passim.
8. See Brink 1989, 130–31: "Any moral belief that is part of reflective equilibrium is justified according to a coherence theory of justification in ethics." A similar view appears in Sayre-McCord 1996, 176: "a person's moral beliefs are epistemically justified if, and then to the extent that, they cohere well with the other things she believes." Brink's reference to "reflective equilibrium" is an invocation of Rawls, who can also be read as a pure coherence theorist. His classification is, however, complicated. First, Rawls gives special weight in reflective equilibrium to our "considered judgments," filtered for personal bias, unfamiliarity, and doubt. Second, his views have changed over time. Most significantly, in Rawls 1951, 182–83, considered judgments are further confined to those about which there is general agreement, whereas in *A Theory of Justice* he writes, memorably:

    > I shall not even ask whether the principles that characterize one person's considered judgments are the same as those that characterize another's. […] We may suppose that everyone has in himself the whole form of a moral conception. So for the purposes of this book, the views of the reader and the author are the only ones that count. The opinions of others are used only to clear our own heads. (Rawls 1970, 50)

    Third, it is not clear whether the procedure of reflective equilibrium is intended as an epistemic standard, as a tool for the study of "substantive moral conceptions," with questions of truth being set aside (Rawls 1975, 7), or as a methodological proposal (Sayre-McCord 1996, 140–45). Finally, even if reflective equilibrium is understood in epistemic terms, Rawls might accept a qualified analogy between moral intuitions and perceptual appearances; see Daniels 1979, 269–72.
9. See also Kagan 2001, 45–47.
10. There are clear statements of this approach in Feldman 2006, Elga 2007, and Christensen 2007; for what I take to be decisive objections, see Kelly 2010, §3.
11. The antecedent of this conditional excludes two possibilities: in the first, my evidence suggests that I am more reliable; in the second, whatever our comparative reliability, the evidence for mine is more extensive. In the discussion of morality to come, both possibilities count as ones in which I have "more evidence of my reliability than yours."
12. The qualification about self-knowledge explains why some resist the skeptical verdict when the disagreement is more extreme: the race does not look close to either of us; I believe that Horse A won by a length; this is what you believe about Horse B. Since conflicting appearances are common when a race is close, but rare when it is not, the extent of our disagreement gives me reason to doubt that things seem to you the way you say they do. Once we correct for this uncertainty, the skeptical pressure returns.
13. For a similar argument, see Feldman 2006, 222–24, responding to Rosen 2001, 86–87.
14. Do these hypotheses make sense? It might be argued, against that assumption, that "moral disagreement […] requires a background of shared moral opinion to fix a common […] set of meanings for our moral terms" (Jackson 1998, 132). If we appear to disagree too sharply, the right interpretation has us talking past one another, expressing different moral concepts even if we use the same words (Jackson 1998, 137). Some versions of this line would emphasize agreement in intuition. Others would allow more variation there, claiming only that our intuitions must be

similar enough, and the demands of coherence stringent enough, to yield a close consensus. Neither claim is plausible. Like Williamson (2007, ch. 4), I doubt that concept-possession is generally so constrained. And whatever may hold for specifically moral concepts like *right* and *wrong*, even radical disputes about what there is reason to do, and how one should live, are perfectly intelligible.

15. This claim echoes White (2006, 538–59) on future justification; he also notes the need to qualify. For related though more technical discussion, see van Fraassen 1984 and Briggs 2009.
16. As in Brink 1989, ch. 5, and Sayre-McCord 1996.
17. The charge of egoism implies an unacceptable bias toward oneself, and that may seem to go beyond what the argument shows. The coherence theorist could defeat the accusation of bias by claiming that, while one is justified in being conditionally confident of one's own opinions in case of disagreement, one is also justified in being conditionally confident of the opinions of others. Both attitudes are epistemically permitted. But this reply won't work. In effect, the reflection principle is a principle of transmission, on which knowledge of the epistemic status of *p* in circumstance *q* transmits that status to one's conditional credence in *p* given *q*, with the proviso about misleading evidence. It follows that the coherence theorist can avoid a biased treatment of my antecedent conditional credences only by saying that, in the circumstance of coherent disagreement, I have as much justification for thinking that you are right as I have for thinking that I am. That is a consequence we should hope to avoid. Properly understood, the pure coherence view is permissive in holding that, if I coherently believe *p* and you coherently believe *not-p*, each of us is epistemically justified, not that, in that circumstance, I have no more justification for believing *p* than *not-p*. It is this bias in favor of what I happen to believe that the reflection principle transmits, and that generates the epistemic egoism to which I object.
18. For versions of this idea, see, especially, Davidson 1983, 146–51 and Williamson 2007, ch. 8.
19. Versions of this idea for nonmoral epistemology can be found in Cohen 1999, White 2006, §9, and, with qualifications, in Schiffer 2004, §6–7, and Wright 2004.
20. A similar claim about the connection between moral or practical wisdom and more mundane capacities to interpret and predict one's environment is defended by Hursthouse (2006), in the course of reading Aristotle. There are further questions to ask about so-called doxastic justification: not just having justification to believe, but believing with justification. Where there is adequate nonmoral evidence for a moral proposition, *m*, to believe it with justification is, at least, to believe it on the basis of that evidence. But this is not sufficient, since the path from evidence to moral belief may be confused, or involve defective reasoning: a rule of inference that goes badly wrong elsewhere. If I believe *m* with justification, or know that it is true, it cannot be an accident that my belief is based on evidence that in fact supports it. The issues raised by this requirement are not specific to moral epistemology: they apply to all beliefs. Resolving them would take us far from our present concerns.
21. That is, to reject what has come to be known as a "dogmatic" attitude to the evidence for one's beliefs. This terminology derives from Pryor 2000. For a powerful critique of dogmatism in epistemology, see White 2006; and for complications, Weatherson 2007.
22. Again, there is a puzzle about doxastic justification, and it cannot be even part of the answer that believing *m* with justification is believing it on the basis of sufficient evidence. We need some other account of the difference between justified belief and belief that is accidentally true.
23. Is there room for a hybrid view on which moral intuitions play an evidential role that supplements this nonmoral and perhaps nonevidential bias? In principle, yes, but in practice it is hard to see why the bias should apply only to some moral truths, and then which ones. Nor would a view of this kind entirely solve the problems of section 2: there will be an argument for loss of confidence or epistemic egoism in the region of moral belief that rests on intuition. More attractive is the view that intuitions are relevant to "non-ideal theory," which directs and evaluates subjects for whom propositional justification and apparent truth diverge.
24. This paragraph generalizes the Total Evidence View (Kelly 2010), according to which the correct response to disagreement turns on whose beliefs the evidence in fact supports. It does not follow that, if we are right, we should be utterly dogmatic, since it may be rational for us to doubt our own reliability in the face of less radical and so more troubling disagreement. Like the role for

intuitions mentioned in the previous endnote, this prospect falls within the scope of non-ideal theory.

25. Dworkin 1996 gives trenchant expression to these ideas, which are also associated with John McDowell.
26. Dworkin (1996, 117–20) comes close to the first pitfall in his stubborn response to skepticism, and to the second at various points; see Dworkin 1996, 98, 122, 125. McDowell's invocations of Neurath on repairing the ship at sea (e.g., at McDowell 1994, 80–82 or McDowell 1995, 188–89) can easily but wrongly suggest an epistemology of pure coherence, or the boring methodological point that we have to start with what we think.
27. Sayre-McCord (1996, 138) is unusual in making this explicit: "So far as I can see, the epistemic evaluation of our moral beliefs is of a piece with that of all our other beliefs; there is no distinctive epistemology of moral belief." My resistance to this view is qualified, in that the arguments of section 2 are somewhat general: they count against an epistemology of intuition for any subject-matter that allows for radical disagreement but admits neither epistemic egoism nor the implications of the unrevised empirical model. Wherever these conditions hold, we will be pushed toward epistemic localism of the kind that I defend in the moral case. To that extent, the locality of epistemic standards is a relatively global phenomenon.
28. This points to a grain of truth in the pure coherence view. Since we must rely on our moral beliefs not only as a guide to morality but to the standards of moral epistemology, an egoistic trust in those beliefs may be impossible to avoid. That would make it epistemically blameless, and in a weak sense justified, even when it leads us astray.
29. These suggestions appear in Singer 1972.

REFERENCES

Anscombe, G. E. M. 1957. "Does Oxford Moral Philosophy Corrupt Youth?" Reprinted in *Human Life, Ethics and Action*, ed. Mary Geach and Luke Gormally. Exeter: Imprint Academic, 2005. 161–67.

Anscombe, G. E. M. 1958. "Modern Moral Philosophy." Reprinted in *Ethics, Religion and Politics*. Oxford: Blackwell, 1981. 26–42.

Baier, A. 1985. "Theory and Reflective Practices." In *Postures of the Mind*. Minneapolis: University of Minnesota Press. 207–27.

Briggs, R. 2009. "Distorted Reflection." *Philosophical Review* 118:59–85.

Brink, D. 1989. *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.

Christensen, D. 2007. "Epistemology of Disagreement: The Good News." *Philosophical Review* 116:187–217.

Clarke, S., and E. Simpson. 1989. *Anti-Theory in Ethics and Moral Conservatism*. Albany: State University of New York Press.

Cohen, S. 1999. "Contextualism, Skepticism, and the Structure of Reasons." *Philosophical Perspectives* 13:57–89.

Daniels, N. 1979. "Wide Reflective Equilibrium and Theory Acceptance in Ethics." *Journal of Philosophy* 76:256–82.

Davidson, D. 1983. "A Coherence Theory of Truth and Knowledge." Reprinted in *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press, 2001. 137–53.

Dworkin, R. 1996. "Objectivity and Truth." *Philosophy and Public Affairs* 25:87–139.

Elga, A. 2007. "Reflection and Disagreement." *Noûs* 41:478–502.

Feldman, F. 2006. "Epistemological Puzzles about Disagreement." In *Epistemology Futures*, ed. Stephen Hetherington. Oxford: Oxford University Press. 216–36.

Hooker, B. 2000. *Ideal Code, Real World*. Oxford: Oxford University Press.

Hursthouse, R. 2006. "Practical Wisdom: A Mundane Account." *Proceedings of the Aristotelian Society* 106:285–309.

Jackson, F. 1998. *From Metaphysics to Ethics*. Oxford: Oxford University Press.

Kagan, S. 1989. *The Limits of Morality*. Oxford: Oxford University Press.
Kagan, S. 2001. "Thinking about Cases." *Social Philosophy and Policy* 18:44–63.
Kamm, F. 2007. *Intricate Ethics*. Oxford: Oxford University Press.
Kelly, T. 2010. "Peer Disagreement and Higher-Order Evidence." In *Disagreement*, ed. Richard Feldman and Ted Warfield. Oxford: Oxford University Press. 111–74.
Louden, R. 1992. *Morality and Moral Theory*. Oxford: Oxford University Press.
MacIntyre, A. 1981. *After Virtue*. Notre Dame, IN: University of Notre Dame Press.
McDowell, J. 1986. "Critical Notice of Bernard Williams." *Ethics and the Limits of Philosophy*. *Mind* 95: 377–86.
McDowell, J. 1994. *Mind and World*. Cambridge, MA: Harvard University Press.
McDowell, J. 1995. "Two Sorts of Naturalism." Reprinted in *Mind, Value, and Reality*. Cambridge, MA: Harvard University Press, 1998. 167–97.
Murdoch, I. 1970. *The Sovereignty of Good*. London: Routledge.
Nagel, T. 1986. *The View from Nowhere*. Oxford: Oxford University Press.
Nozick, R. 1974. *Anarchy, State, and Utopia*. New York: Basic Books.
Pryor, J. 2000. "The Skeptic and the Dogmatist." *Noûs* 34: 517–49.
Rawls, J. 1951. "Outline of a Decision Procedure for Ethics." *Philosophical Review* 60:177–97.
Rawls, J. 1970. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
Rawls, J. 1975. "The Independence of Moral Theory." *Proceedings and Addresses of the American Philosophical Association* 48:5–22.
Rosen, G. 2001. "Nominalism, Naturalism, Epistemic Relativism." *Philosophical Perspectives* 15:69–91.
Ross, W. D. 1930. *The Right and the Good*. Oxford: Oxford University Press.
Sayre-McCord, G. 1996. "Coherentist Epistemology and Moral Theory." In *Moral Knowledge?*, ed. Walter Sinnott-Armstrong and Mark Timmons. Oxford: Oxford University Press. 137–89.
Scheffler, S. 1982. *The Rejection of Consequentialism*. Oxford: Oxford University Press.
Schiffer, S. 2004. "Skepticism and the Vagaries of Justified Belief." *Philosophical Studies* 119:161–84.
Singer 1972. "Moral Experts." *Analysis* 32:115–17.
Slote, M. 1992. *From Morality to Virtue*. Oxford: Oxford University Press.
Slote, M. 2001. *Morals from Motives*. Oxford: Oxford University Press.
van Fraassen, B. 1984. "Belief and the Will." *Journal of Philosophy* 81:235–56.
Weatherson, B. 2007. "The Bayesian and the Dogmatist." *Proceedings of the Aristotelian Society* 107:169–85.
White, R. 2006. "Problems for Dogmatism." *Philosophical Studies* 131:525–57.
Williams, B. 1985. *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press.
Williamson, T. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.
Williamson, T. 2007. *The Philosophy of Philosophy*. Oxford: Blackwell.
Wright, C. 2004. "Warrant for Nothing (and Foundations for Free)." *Proceedings of the Aristotelian Society, Supplementary Volume* 78:167–212.