Internalism and Speaker Relativism*

James Dreier

INTRODUCTION

In this article I set out a reason for believing in a form of metaethical relativism. In rough terms, the reason is this: a widely held thesis, internalism, tells us that to accept (sincerely assert, believe, etc.) a moral judgment logically requires having a motivating reason.¹ Since the connection is logical, or conceptual, it must be explained by a theory of what it is to accept a moral claim. I argue that the internalist feature of moral expressions can best be explained by my version of moral relativism, which I call “speaker relativism.”

Like noncognitivist theories, my view builds into the very semantics of moral terms a connection to the speaker’s motivating reasons. Unlike any noncognitivist theory, my account of the semantics is truth conditional; I offer an analysis of moral expressions which specifies the conditions in which sentences containing them are true on a given occasion of use. My theory differs from noncognitivism in that it presents a way in which tokens of sentences containing moral terms can be straightforwardly true or false. I argue that this difference is an advantage for my account.

Speaker relativism is the theory that the content of (what is expressed by) a sentence containing a moral term varies with (is a function of) the context in which it is used.² The content of a moral term itself depends on the most salient moral system in the context of use. Just what a moral system is, and which systems are most salient in a context, I shall explain presently. For now, let me note why I call this view a form of relativism.

* I would like to thank Mark Kalderon, Steve Rieber, Gideon Rosen, Michael Smith, Jamie Tappenden, and Mark van Roojen for helpful discussions of a number of ideas presented in this article. I also thank three anonymous reviewers for their comments on an earlier draft.

1. By a motivating reason I mean one which will pass muster as what Bernard Williams calls an “internal reason” in “Internal and External Reasons,” in Moral Luck (Cambridge: Cambridge University Press, 1981), pp. 101–13. Someone has a motivating reason, r, to do x, just in case r could be cited in an explanation of his or her intentionally doing x.

2. My account owes a great deal to that of David Wong in Moral Relativity (Berkeley: University of California Press, 1984). It can, indeed, be seen as a modified version of his theory.

Ethics 101 (October 1990): 6–26
© 1990 by The University of Chicago. All rights reserved. 0014-1704/91/0101-0003$01.00
For one thing, the content of a moral claim or belief is, on my view, relative to a context. For another (and this is really just a consequence of the first), two people in different contexts may utter "x is good" and "x is not good" and both speak truly. This possibility has sometimes been considered the hallmark of relativist metaethics.\(^4\) Contrary to some popular arguments,\(^4\) I believe this type of relativism can be coherently maintained. But why should it be maintained? The fact that a theory is coherent is not itself a sufficient reason to believe it! So my aim in this article is to explain why such a theory is attractive.

The reason, again, is that the theory is best able to account for internalism. Of course, that is a very poor reason unless internalism itself is a plausible doctrine. Through much of this century, internalism was one of the least controversial data in metaethics. But recently it has been challenged. Part of my job is to reinstate the doctrine, or a suitably modified version sensitive to the relevant objections. I begin with some general remarks designed to capture the flavor of the project. I then present internalism in a crude version, set out the objections to that version, and propose a modified formulation. Next, I show how speaker relativism can explain internalism and argue that it provides a better explanation than do leading competitors. Finally, I work into my account some complexities designed to capture the subtleties introduced by the discussion of the objections to internalism.

A TWO-LEVEL SEMANTICS FOR MORAL TERMS

We want to know what could be the semantic value of moral terms, what someone has to know to be competent in their use, given a widely recognized feature of moral judgments. That feature I will call "internalism": there is an internal or logical connection between sincerely asserting a sentence of the form "x is (morally) good" and being motivated to promote or perform x.

The semantic value of a moral term seems to have two parts. On the one hand, it includes a content—an extension, and an intension or function from circumstances of evaluation to extensions, and perhaps other contentful aspects such as a descriptive sense. The content is roughly the contribution the term makes to what is expressed by sentences in which it occurs on an occasion of utterance.

On the other hand, the semantic value of a moral term includes some part that captures the internalist feature. That part explains why someone who asserts sincerely "x is good" will normally have the relevant motivation. Imagine a group of people who recognize natural facts of human suffering and fulfillment, equality of distribution, and so on, but

do not care about those things. If they have a term that applies to all and only things with those natural properties, it might have the same extension, intension, the same content as our word “good,” but there is a sense in which it does not mean what “good” means. You might say, it lacks the internalist character of “good.”

Hume took the internalist feature of moral terms to show that moral properties are projected by us onto the world. Our representations of vice and virtue, he says, “engage us to avoid the one, and embrace the other.”

Hare wrote, rightly, I think, that evaluative terms generally have a descriptive part and a prescriptive part. How shall we unify the two? We might say, on a given occasion of use a moral term has a content that corresponds to Hare’s descriptive part. It has the same sort of semantic value that any old predicate has, and that is what allows us to speak of the truth or falsity of the sentence in which it is contained. But moral terms also have the internalist character. If a term in some alien language has the character of “good” we can sensibly translate it as “good.” The alien community may not have the same motivations as we have, so the term we translate as “good” may not have the content of “good” as we use it. The character—what stays the same between one context and another—determines the content in a context, and then the content determines the truth conditions of the sentences in which the term occurs. If something like this is right, then moral terms have a two-level semantics of the kind David Kaplan proposes for indexicals. Indexicals have as their primary meaning, according to Kaplan, not a content, but a character, a rule for determining the content given a context. The character is their primary meaning because it is what someone has to know in order to be competent in their use. I will argue that a relativist theory of ethics can account for the internalist feature of moral terms, by maintaining that the connection with a motivation to act is part of their character, which

5. For now, I leave this characterization of the two parts of the semantics of “good” rough; I will fill it in later (pp. 19–22, 23–26). But readers familiar with David Kaplan’s “Demonstratives” will be able to anticipate much of the structure of the account I will give; that paper is now published in Themes from Kaplan, ed. Joseph Almog, John Perry, and Howard Wettstein (New York: Oxford University Press, 1989), pp. 481–563.


7. R. M. Hare, The Language of Morals (Oxford: Oxford University Press, 1952), pp. 111–21; see also p. 146: “‘Good’, as used in morals, has a descriptive and an evaluative meaning, and the latter is primary.”

8. I have not said much about what that character is; I will.

9. By an “indexical” I mean an expression for which the content “is dependent on the context of use and [for which] the meaning [or character] provides a rule which determines the [content] in terms of the context.” The definition is from Kaplan, “Demonstratives,” p. 490, but I have replaced his use of “referent” with the word “content” because I wish to include expressions which are not singular terms, but predicates. I shall say a good deal more about the indexical account of moral predicates later.

10. Ibid., p. 505.
remains constant across use by groups with differing moral interests and concerns. When those concerns do vary, so must the content of the moral terms. That variation is the mark of relativism.

The kind of relativism I advocate is roughly this: the content of a moral term in a context is a function of the affective attitudes of the speaker in the context. Thus, "x is good" means "x is highly evaluated by standards of system M," where M is filled in by looking at the affective or motivational states of the speaker and constructing from them a practical system.11 The main object of this article is to establish a connection between internalism and this kind of relativism.

INTERNALISM

That there is an internal or intrinsic connection between accepting a moral judgment and having a motivating reason is widely accepted, in one form or another, by philosophers of diverse metaethical persuasions. Thomas Nagel writes, "Internalism is the view that the presence of a motivation for acting morally is guaranteed by the truth of ethical propositions themselves."12 John Mackie says that moral properties would have to have a "peculiar evaluative, prescriptive, intrinsically action-guiding" quality.13 Similar remarks are made by W. D. Falk, Edward Westermarck, C. L. Stevenson, and D. C. Williams, to name a few.14 It would be good to have a succinct formulation of a principle all could agree upon, one which would capture the intuition behind all of these remarks. Unfortunately, it is not so easy to say exactly what the internal connection amounts to.

How to Overstate Internalism

It is easy to overstate the case for internalism. Some moral philosophers have asserted that "the good attracts" or that we are necessarily motivated by what we believe to be good. The very strongest version says that we will in fact do what we think is right, that someone's failure to perform a certain action shows that he does not really believe that action to be

11. The directions for "filling in" specify the character of "good," in Kaplan's sense, since they characterize a rule for determining the content given a context of use. Since moral expressions meet Kaplan's definition for indexicals, I shall sometimes say that, on my view, moral expressions are indexicals. By this I simply mean that they share with run-of-the-mill indexicals like "I" and "now" the fact that their character, and not their content, is what remains fixed on different occasions of use.


best. Of course, this strongest doctrine runs afoul of the real phenomenon of weakness of the will.

Instead we should consider the view that a person has some motivation to act as he or she believes will be morally best. There may often be other considerations that affect my deliberations: concern for my other interests does present itself as a rival for practical decisions, and no one runs her life exclusively according to the dictates of morality. So the strong version of internalism says that, necessarily, someone who believes that a certain action is good will be motivated to perform (or try to get others to perform) that action.

This strong version is overstated, too. The problem is that there are situations in which we want to say that an agent can see what is morally good or right yet lacks the relevant motivation. They are not the central cases the internalist was thinking of—indeed, I will argue that they cannot be the normal cases in some sense of "normal"—but they demand some sort of an account. I will describe two classes of counterexamples to strong internalism; there may be others.\(^\text{15}\)

In "Desiring the Bad: An Essay in Moral Psychology," Michael Stocker describes a politician who cared in his youth "about the suffering of people in all parts of the world and devoted himself to making their lives better. But now he concerns himself only with the lives and fortunes of his close family and friends. He remembers his past, and he knows that there is still a lot he could do to help others. But he no longer has any desire to do so."\(^\text{16}\) Of course, there are different ways of characterizing the politician's situation. Maybe he has ceased to believe that it is a good thing to help others. That would be a way of describing the politician's state that does not jeopardize our simple principle. But Stocker's insight is that it makes perfect sense to describe the uncaring man another way. Couldn't he say, "I have not changed my moral beliefs; I still think the morally good thing to do is to help the unfortunate. But I am no longer interested in doing what is morally good." Stocker wants us to notice that it is a perfectly intelligible, all too familiar thing to say. "Lack of [the desire to do what one believes to be good] is commonplace. Through spiritual or physical tiredness, through accidie, through . . . despair, . . . one may feel less and less motivated to seek what is good."\(^\text{17}\)

Stocker's example shows how strong internalism may fail when there is something psychologically wrong with an agent. Here is another kind of problem.

The Sadists are a group of people who recognize what sorts of actions, states of affairs, characters, their society calls "morally good."


\(^{17}\) Ibid., p. 744.
They recognize them, but they hate them. They find those things repulsive and avoid them at all costs. In general they are not in the slightest motivated to perform those actions. Here is the way they put it: “We despise what is good because it is good. We believe the things you say are good are in fact, and we are never motivated to promote those things. Quite the opposite!” That is what the Sadists say, and it seems as though it might be true.\textsuperscript{18}

We might say, of course, that the Sadists do not really believe the conventionally approved things are good. They are using the word in the inverted commas sense; that is, when they say they believe kindness is good they mean they believe that kindness is what people call “good.”\textsuperscript{19} We might say that, and if it is right then the internalist will have an easier time. But I am afraid that our temptation to explain away the example, insofar as we are tempted, is due to an excessive fondness for the simple internalist theory. We like the theory in its simple and pristine form, and we can think of a consistent way to save it. But internalist though I am, I do not feel comfortable denying this anti-internalist datum. It just does seem to me that in the situations described, strong internalism fails. Still, as with the Stocker counterexamples, I am inclined to say that the failure of internalism is parasitic on a background of its successes. The internalist will say that the exceptions to his principle seem to reside around the edges of logical space, whereas the important cases are safely in the core.

In each of these types of counterexamples the unmotivated believers are in some way abnormal, and the examples use the normal case to contrast with the problematic situation. The politician’s belief that it would be good to help the less fortunate is given to us as originating in his youth, when he did have the relevant motivations, and the Sadists are described as reacting to the conventional morality. I think that it is crucial to the examples that they do present their exceptions by reference to a normal case, that it is necessary that such cases be present even to get the exceptions off the ground. So internalism can be saved by qualifying it so that it posits a necessary connection between the believed good and motivation only in the normal cases.

Of course, Stocker’s example is in one way quite unlike the example of the Sadists. For as Stocker says, “Lack of [the desire to do what one believes to be good] is commonplace.”\textsuperscript{20} Whereas Sadists like those I have described are anything but commonplace! If we are going to try to save internalism by claiming that it normally holds, then we cannot rely on a statistical notion of what is normal. I am going to develop a notion that is not primarily statistical. But let me first convince you that the kind of person Stocker describes is not so common as all that.

Strong internalism, again, holds only that someone who believes that a certain action is right or good thereby acquires some motivation to act.

\textsuperscript{18} I thank Gideon Rosen for this example.
\textsuperscript{19} Hare, pp. 124–25.
\textsuperscript{20} Stocker, p. 744. I thank an anonymous reviewer for reminding me of this contrast.
It does not require that the person does in fact act on her motivation. It may happen that to do what she thinks is right would mean sacrificing a great deal—her personal projects, as Bernard Williams says—21—or ignoring the ties of family or friendship. That she fails to act does not show that she had no motivation. What Stocker wants to describe is a radically disaffected person, one lacking altogether the motivation associated with the moral judgment. Thus, he characterizes his politician as overcome by “tiredness,” “acciédie,” “despair.” In such extreme forms, these conative states are, it seems to me, highly abnormal. We are to imagine a man who would not lift a finger, as it were, to achieve what he believes to be good—for if he would make even some minimal effort, then the relevant motivation is exposed. It is this sort of characterization, I claim, that begins to strain our imaginative power. If we can make sense of the politician’s being so unmoved in the face of the acknowledged good, then we can do so only because he has been given a history that fixes in our sights a normal background.

Normality

The reader may be suspicious of my employment of the idea of a normal case. Describing the Sade and Stocker examples as abnormal and especially as parasitic on our understanding of normal cases seems intuitively right, and I think it can go some distance toward preserving the internalist thought, so I want to place some weight on this idea. But the skeptic has a right to be suspicious: it is easy to imagine an equivocal or illicit appeal to “normal conditions.”

How then shall we explain what is to count as abnormal? One way would be to try to list the psychic states, social contexts, and so on, that defeat normality. I do not think much of that approach. For one thing, faced with any proposed list the skeptic can always ask, Why that list? What makes those conditions abnormal? Of course, we could simply insist that as a matter of brute fact these conditions just constitute our conception of the abnormal for the present purpose, and that no further explanation is possible. But the skeptic will suspect that we are packing an ethically loaded conception of normality into the definition of moral terms, that we are passing moral judgment on agents of a certain type instead of constructing a genuine metaethical view. The “brute list” method looks too post hoc; it looks like we are ruling out as abnormal precisely the situations in which a person is not motivated to do what he or she believes to be the right thing. Couldn’t we pull a similar trick with any common but contingent kind of motivation? If so, then internalism tells us something deep about moral terms in particular. Finally, any list we come up with will, I suspect, be open-ended. We should not be generally confident that

21. Williams stresses the importance of personal projects in ethical theory in many places, but I am thinking particularly of “Persons, Character, and Morality,” in Moral Luck, pp. 1–19.
we have identified all of the potential defeaters of crude internalism. We have the feeling that we could recognize the abnormal cases as they came along, so it seems as though there must be something they all have in common. But what?

What the skeptic really suspects is that the only way of capturing the commonality is: “case in which A believes \( x \) is good and A is not motivated to promote \( x \).” If so, then internalism tells us nothing about “good” that is not equally applicable, mutatis mutandis, to say, “expensive.” The upshot of internalism would be to point to a common object of desire, note that there are exceptions, and then dress this fact in a specious logical thesis.

That will not do. So we need to show that we have some independent grip on normality, a conception of the normal that does not depend on a prior notion of “morally good.” I think there is evidence that we do have such a conception. For one thing, if the Stocker and Sade examples are restated so that the background of central internalist cases is entirely absent, we are no longer comfortable describing those examples as abnormal. Suppose the politician to have been born to pioneers, and to have spent his life in the company of rugged individualists. And suppose he now says, “What my friends believe is wrong: not individualism but a life in the service of others is really good.” But the politician has no inclination to serve the less fortunate; instead he advances the cause of self-reliance whenever he can.

Or suppose we discover an isolated culture of mean-spirited folk who go out of their way to cause harm and humiliation at every opportunity, while shunning kindness and fairness and generally avoiding any action that would contribute to the well-being of sentient creatures. Unlike the Sadists in our earlier example, they have never met people who feel differently. These Sadists call the things that attract them “gad” and those that repel them “bood.”

It seems to me that what we want to say about the new politician is that he is using the word “good” either insincerely or incorrectly. We will not take his assertion at face value and attribute to him the belief “Life in the service of others is good.” Similarly, what we want to say about the isolated Sadists is that their “gad” means good and “bood” means bad. And if I am right about our intuitions in these new cases, then a satisfactory account of the difference between them and the original examples is that in each of the originals there was a salient standard of normality against which to contrast the peculiar, whereas in the modified examples there were no such standards. So we can distinguish among cases in which agents are not motivated to do what we think is good, recognizing on the one hand genuine exceptions to strong internalism and on the other insincerity or failure properly to understand or use moral terms. Since we can make that distinction, our conception of normality is not merely, “Those circumstances under which a person is motivated by what she believes to be good.” We must have some independent idea of normality.
I do not know how to specify this independent conception rigorously. It is clear to me that if everyone in a community behaves in a certain way, then that behavior is normal for the community, and if a person has a certain state of character for all of her life, then behavior flowing from that state is normal for her. But that is only to specify a limit. Though I think I have successfully argued that we do have a grip on the conception I need, I cannot now provide an analysis. But let us call modest internalism the principle that in normal contexts a person has some motivation to promote what he believes to be good. Modest internalism is not vacuous, though of course it is weaker than strong internalism.

SPEAKER RELATIVISM EXPLAINS INTERNALISM

I think modest internalism captures the kernel of truth in the philosopher's truism that the good attracts. Now the question is, How are we to explain this internalist fact? Or more sharply, how could even modest internalism possibly be true? It is puzzling how there could be beliefs, the mere having of which would be logically connected to a motivation. It is not that our folk psychology conceives beliefs and desires as completely independent; there is no mystery as to how a belief that by doing A I will do B could lead to my motivation to do A—provided that I am motivated to do B. The worrying feature of internalism is that it seems to posit a categorical imperative: it says that there are beliefs the mere having of which normally establishes the presence of a motivation without depending on any other prior motivation.

Emotivism and the Moral Sense Theory

I do not think the difficulty has been fully appreciated in metaethics, except perhaps by some emotivists and other radical noncognitivists. The question is not, as Mackie puts it, what kind of fact it could be that motivates intrinsically upon its apprehension. It is rather how apprehension of a fact, without any particular prior disposition to act, could itself provide motivation. Noncognitivists answer by explaining that what look on the surface to be indicative sentences really have a disguised form. The emotivist claims that “x is good” has no truth conditions, and that “John believes that x is good” is true, not when John has a belief with a certain content, but just in case John is motivated to attain or promote x.

I have some sympathy for emotivism and other noncognitive metatheoretical theories. Relativism flows naturally from them, and speaker relativism is in a way their child. But the noncognitivist notoriously has grave difficulty explaining why moral sentences are so easily embedable in truth-functional and other contexts in which familiar, contentful

22. Mackie, p. 38: "If there were objective values, then they would be entities or qualities relations of a very strange sort, utterly different from anything else in the universe."
expressions are at home. If moral utterances have only superficial resemblance to indicative sentences, it is quite mysterious why they should substitute readily into such contexts as:

- It is true that p.
- Ronald knew that p.
- If p, then q.
- It is possible that p.

Some noncognitivists have begun to address this difficulty. Simon Blackburn has proposed a logic for a kind of emotivist semantics of evaluative expressions, or more precisely, he has offered two accounts, the first developed in *Spreading the Word*, the second in “Attitudes and Contents.” Each of these accounts is designed to explain how sentences containing evaluative terms embed into (typically) truth-functional contexts. At present, there seem to be some fairly serious obstacles to completing Blackburn’s project. A different approach was advocated by R. M. Hare in chapter 2 of *The Language of Morals*, where Hare subsumes the logic of evaluation under the more general logic of prescriptions. Both Hare and Blackburn supplement their logic with an account of moral truth, that is, with an explanation of what it means to call an evaluative utterance “true.” These projects deserve a more lengthy treatment than I can give them here. It does seem to me that neither addresses the full force of Geach’s point: that an explanation is wanted of how evaluative sentences play a semantic role in propositional contexts in general, including truth functions and the truth predicate, but equally in propositional attitudes and modal contexts.

My view addresses all of these problems at once, or better, the problems never even arise for my view. On my proposed analysis, sentences containing moral terms straightforwardly express propositions—though which propositions they express will vary with the context. So I can simply say that such sentences embed into the relevant contexts just as any other indexical sentences do. I take this to be an advantage. Unless the noncognitivist can give a satisfying response to Geach, his account of moral language must limp along where others fly.

So I want to look at a theory akin to noncognitivism, a kind of subjectivism which still allows moral utterances to be truth-bearing expressions. It is sometimes called the Moral Sense Theory.

23. The problem was first posed, as far as I know, → Peter Geach in “Assertion,” *Philosophical Review* 74 (1965): 449–65.
Michael Smith proposes a contemporary version of a Moral Sense Theory in “Should We Believe in Emotivism?”26 “According to this view 'x is good’ means 'x is such as to be approved of by those who share our moral viewpoint under suitable conditions.”27 Smith claims as a virtue for the theory that it explains “the action-guidingness of morality,” what I am calling the internalist feature of moral terms. “The Moral Sense Theorist can say that, necessarily, ... if I judge that x is such as to be approved of by those who share our moral viewpoint under suitable conditions and believe that conditions are suitable then I approve of x .... It is plausible to suppose that this suffices to capture a version of the action-guidingness principle.”

Smith builds action-guidingness into the semantics of moral terms by putting conative states of the speaker into the truth conditions of moral utterances. I think something like this must be involved in explaining internalism. But I also think there is an essential error in the Moral Sense Theory.

One difficulty is that it is stronger at one end and weaker at another than the modest internalism it should be trying to explain. It is too strong because it implies that the Sadists, who we may presume are not mistaken about their mental states and do, under suitable conditions, judge good the conventional actions of their fellow citizens, must approve of those conventions and be motivated to support them. But the Sadists have no such motivations or attitudes. The theory is too weak because it fails to explain why someone who falsely believes “x is good” should be apparently under the same rational requirement to promote x as someone who truly believes “x is good.” For someone might well falsely believe “x is good.” Then she falsely believes, according to the Moral Sense Theory, that x would be approved of by those who share her moral view under suitable conditions. It is understandable and laudable that someone should in general be guided that way. But, from the point of view of rationality, it is quite optional.

A deeper problem is hinted at by the second half of the mismatch, the failure to explain the necessary motivation of the mistaken believer. The reason the Moral Sense Theory failed there is that it makes moral beliefs about our motivations, so it should allow that where the belief is false we lose the connection to motive. Whereas the emotivist, who holds that the sincere utterance of “x is good” expresses a conative state, can explain why we can infer the motivation from the utterance even when the judgment is based on false beliefs. On this count the emotivist is right and the Moral Sense Theorist is wrong. A view that equates sentences involving moral terms with sentences about the speaker’s actual or counterfactual attitudes, dispositions, and so on, will run into difficulty explaining

27. Ibid., p. 301.
how belief about a motivation translates into the motivation itself—especially when the belief is false. Furthermore, the Moral Sense Theory is committed to an intuitively mistaken conflation of states whose contents are propositions about such a component.

Let me say a bit more about this. The analysis given by the Moral Sense Theory has it that, when Marsha utters,

1. “Cutting taxes is good,”

she is saying that cutting taxes is the kind of thing that she and others like her would approve if conditions were suitable. Is Marsha issuing a prediction when she utters 1? If the Moral Sense Theory is correct, then she is predicting that she would approve of cutting taxes in a certain situation. But we take Marsha’s avowal to be, not a prediction, but itself an expression of approval. The distinction is parallel, I think, to the difference between describing one’s beliefs and stating them. Of course, one can express the belief that New York is the largest city in the United States by saying,

2. “I believe that New York is the largest city in the United States.”

Conversationally, that sentence is taken to be equivalent to:

3. “New York is the largest city in the United States”

(though 3 usually indicates a greater degree of certainty). But the utterance of 2 could be true even if 3 is false, because the first is about the speaker’s belief, while the second expresses the belief.

When Marsha utters 1 we do not have to wonder whether she is correct in order to infer something about her attitude. That is because her statement is not about her attitude of approval. If she utters 3 we don’t have to wonder whether she is correct in order to infer something about her beliefs. That is because her statement is not about her beliefs. Rather, each of her claims expresses an attitude—the one an attitude of approval, the other the (propositional) attitude of belief. Whereas according to the Moral Sense Theory, when Marsha utters 1 and is mistaken, then she has falsely said that she would approve of cutting taxes under suitable conditions. The Moral Sense Theory tells us that in order to conclude something about her attitudes, we must determine whether what Marsha said is true (notice: not sincere, but true). So the Moral Sense Theory mistakes a sentence expressive of an attitude with one about an attitude.

Indexicals and Behavior-explaining Beliefs

It is one thing to lay out strict requirements for an adequate semantics of moral vocabulary, another to say how those requirements might be met. My strategy will be to draw a loose analogy between internalism in morals and the problem of the essential indexical, and suggest that the solution to the first ought to parallel the solution to the second. I then
spell out the indexical account of the internalist feature offered by speaker relativism.

Let me begin by quoting at some length from the opening of John Perry’s “The Problem of the Essential Indexical.”

I once followed a trail of sugar on a supermarket floor, pushing my cart down the aisle . . . , seeking the shopper with the torn sack to tell him he was making a mess. But I seemed unable to catch up. Finally it dawned on me. I was the shopper I was trying to catch.

I believed at the outset that the shopper with a torn sack was making a mess. And I was right. But I didn’t believe that I was making a mess. That seems to be something I came to believe. And when I came to believe that, I stopped following the trail around the counter, and rearranged the torn sack in my cart. My change in beliefs seem to explain my change in behavior. . . .

At first characterizing the change seems easy. My beliefs changed, didn’t they, in that I came to have a new one, namely, that I am making a mess? But things are not so simple.

The reason they are not is the importance of the word “I” in my expression of what I came to believe. When we replace it with other designations of me, we no longer have an explanation of my behavior and so, it seems, no longer an attribution of the same belief. It seems to be an essential indexical. But without such a replacement, all we have to identify the belief is the sentence “I am making a mess.” But that sentence by itself doesn’t seem to identify the crucial belief, for if someone else had said it, they would have expressed a different belief, a false one.28

Perry’s ultimate view of the change of belief is that the new belief, identified by the sentence, “I am making a mess,” does play a special role in the explanation of behavior but not because of the content that it has. For there might be a belief identified by a sentence with the same content, say, “Perry is making a mess,” which failed to play that role (because Perry has, or might have, forgotten his name). What he came to believe was not something new; he came to believe the same old thing in a new way. He came to appreciate the same fact by accepting a sentence with a special character. I want to say that sentences involving moral terms have an analogous feature.

On a given occasion of use, “good” has a content equivalent to a certain descriptive, naturalistic predicate. Suppose Max thinks, “to do A would be to do an act which is such and such,” where “such and such” is coextensive (and cointensive) with “good” relative to the given context. That he has that thought has no logical connection to any action; for any belief with that sort of content, reason for action seems to come only with an accompanying desire. Suppose Max now comes to believe “to do A would be to do an act which is good.” According to internalism,

Max now has a reason to act. We might think, he must have acquired a belief with a new content, that A is good. But then we are stuck with the puzzle: what could this new content possibly be, given that believing it is intrinsically connected to action? Instead, we should think that Max may have come to believe no new proposition, that the natural property named by “such and such” is the property named by “good” on this occasion of use, that Max came to believe the proposition in a new way.

Perry distinguishes “the sense entertained” by a person in a certain psychological state from “the thought apprehended.” The latter designates a proposition, an object of belief. The former designates a way of believing, a mode of presentation. “It is the sense entertained, and not the thought apprehended, that is tied to human action. When you and I entertain the sense of ‘A bear is about to attack me,’ we behave similarly. . . . When you and I both entertain the thought that I am about to be attacked by a bear, we behave differently. I roll up into a ball, you run to get help.”

When Max comes to believe, “A is good,” he is in a psychological state which explains why he decides to pursue A. But “A is good” denotes no new object of belief, no new thought apprehended. Rather, Max comes to “entertain a new sense,” in Perry’s terms; he has a thought with a new character, in Kaplan’s.

**Speaker Relativism**

I will now explain how speaker relativism takes advantage of its indexical semantics to account for internalism. Keep in mind that the main idea is the analogy with Perry’s conclusions about indexical thoughts. Moral terms, and normative terms in general, are “essential” in the same way that indexicals are: replacing them with terms with the same content destroys the explanatory force of belief attributions.

To start with, let me revert to crude internalism and offer a correspondingly crude version of speaker relativism. This crude version says that the content of a moral term is given on an occasion of use by the moral system of the speaker, and the speaker’s moral system is constituted by her attitudes of approval and other motivational states. The character of “good,” for example, is a function from contexts to the property of being highly rated by the moral system of the speaker in the context.

Suppose Margaret sincerely says, “Public education is morally good.” According to crude internalism, it follows that Margaret is motivated to promote public education. Suppose we replace “morally good” with a coextensive term, one that picks out the property endorsed by Margaret’s moral system. Abbreviate that term by “P.” Margaret’s sincere utterance of “Public education is P” implies no motivation on her part. Yet in each case she believes the same thing. The difference is in the character of the terms.

The character of the sentence “Public education is good” is not what someone believes when she believes that public education is good. What she believes, the object of her belief, is a proposition. A proposition is the content of a sentence. The character of the sentence is a rule that determines what content it has in a given context. If Margaret knows the meaning of “good,” she appreciates its character, and then the sentence containing that term expresses her motivation. For if she understands the meaning of “good,” then she knows that when she uses the word she describes public education as meeting certain standards, which standards are determined by her motivations. When Perry, pushing his shopping cart, suddenly says, “I am making a mess!” the character of the belief he expresses figures integrally in explaining his new dispositions to act. When Margaret says, “Public education is good,” the character of the belief she expresses figures in explaining the motivations predicted by crude internalism.

Notice: I did not say that she is telling us something about her motivations. According to speaker relativism, what she is doing is more like expressing one’s beliefs by asserting them than it is like talking about one’s beliefs. And this is as it should be if internalism is true. Christine Korsgaard puts it nicely: according to internalism, she says, a person does not in general do something because it is right; “an internalist believes that the reasons why an action is right and the reasons why you do it are the same.”

I am pressing hard on the analogy between Perry’s examples of how an indexical belief can explain behavior which cannot be explained by any nonindexical belief, and the internalist thesis that accepting a moral claim has an intrinsic connection to motivation that accepting a naturalistic claim lacks. I do not want to press too hard. Let me point out a difference between the two kinds of cases and then explain why that difference does not matter to my argument.

Here is the difference: it is easy to understand how Perry, pushing his cart and trailing sugar, could think, “That guy is making a mess” without realizing “I am making a mess,” that is, without realizing that he is himself making a mess. But how can Margaret sincerely assert, “Public education is P,” where “P” abbreviates a naturalistic description of good-making characteristics, without realizing “Public education is good”? She would have to be strangely unaware of her own moral system—perhaps no one could be so ignorant. Suppose, anyhow, that Margaret simply cannot recognize that public education has all those characteristics without realizing that it is by her lights good. So her case is not in that respect analogous to Perry’s. Is the disanalogy a crucial one for my argument?

This content downloaded from 77.80.49.240 on Thu, 28 Nov 2013 10:20:58 AM
All use subject to JSTOR Terms and Conditions
No. For the important point of analogy is as follows. In “The Essential Indexical,” Perry shows why we may infer something about the psychology of a person who sincerely asserts “I am making a mess” but not about the psychology of that person when he sincerely asserts “He is making a mess,” even if the two sentences express the same proposition in the context. The inference is licensed by the character of the first sentence, absent in the second. The question posed by the internalist is why we may infer something about the psychology of a person who sincerely asserts, “Public education is good,” but not about the psychology of that person when she sincerely asserts, “Public education is P,” even if we stipulate that “P” is a naturalistic predicate with the same content as “good” in that context. Our answer is that the inference is licensed by the character of the first sentence, absent in the second.32

The crude form of speaker relativism can give a good account of crude internalism. We say that having a belief with the character of “x is good” involves seeing x as instantiating or facilitating one’s own goals. On the crude picture, the content of “x is good” is of the form “x is approved by moral system M,” and when someone has a belief with the moral character, her M is near enough to the surface that the desires are registering.

But we might worry that the more sophisticated internalism will not support such an individualistic form of relativism. According to modest internalism, the logical connection with motivation is mediated by the “normal case,” and normality seemed to be a partially social rather than a wholly individualistic concept. Thus, our handling of the Sade example. Crude internalism fails in that example because the Sadists are described against the backdrop of a normal society and are portrayed as abnormal. We do not want our metaethical theory to predict that someone who is abnormal in that way will have his moral judgments intrinsically connected to his motivations. It looks like only “socially normal” people will experience such a connection. Why, then, should we not endorse what is sometimes called “speaker’s group relativism,” or a kind of cultural relativism? According to those forms of relativism, the moral system that fills in the content of a moral term on an occasion of use is that of the society, not that of the speaker. Shouldn’t speaker’s group relativism fit better than speaker relativism with modest internalism?

No, speaker relativism can allow that what a person’s moral system is may depend on factors extrinsic to that person. In central cases, the ones that have to be in the background to provide a standard of normality, an agent is motivated by what her moral system prescribes. Defeaters like accidie are absent in these normal cases. And a person’s moral system is (defined as) that set of rules, meeting to a suitable degree certain

requirements of content or subject matter, form, and function, which do motivate the person normally.

Since what is normal may depend on social context as well as individual psychology, it may look as if the content of moral terms is really relative to the speaker’s group and not simply the speaker individually, notwithstanding the technical fix I advocate to save the individualistic version. I grant the thrust of this point. But let me make a couple of remarks in favor of running the semantics of moral terms on individualistic lines.

First, speaker relativism is the more general case of which speaker’s group relativism is a species. One way for something to be relative to a person is for it to be relative to that person’s society. If, as speaker’s group relativism would have it, the content of a sentence containing a moral term is fixed relative to a social group then that content is fixed relative to any member of the group. By using the speaker rather than the group we do not lose any relevant information—but we do lose information that might be relevant if we take the group as the relevant feature of the context and drop the individual. So by endorsing speaker relativism we hedge our bets.

Now the Sade case showed that we want to be able to allow the moral system that gives content to an utterance to be one accepted by the speaker’s society. Speaker relativism allows that; speaker’s group relativism insists on it. We should prefer speaker’s group relativism only if we want to insist in general that the content giver is the society’s moral system. And that would mean that we could not disagree with our society’s values by using moral language. Maybe the counterintuitiveness of that consequence can be explained away—though I doubt it. In any case, it is wise at least to leave it open, by using the more inclusive speaker relativism.

FURTHER COMPLEXITIES
I want to give a more satisfactory account of the relativist treatment of the Sade counterexample to internalism. In doing so I will complicate the theory of the character of moral terms and explain in a more intuitively appealing way what lies behind the indexical account I have given. I will show how, despite the complications, the connection between internalism and relativism can be maintained.

The character of an ethical term is not so simple as I have made it out to be. It is not exhausted by building the internalist feature into the semantics. The character, a function determining content given context, is a complicated function whose argument is the context as a whole, not merely the speaker. Before describing the complications, let me say a word or two about the semantics of uncontrovertially indexical terms. In Kaplan’s semantics, the character of “here” is a function from a context of utterance to the position in the context. A context is an ordered quadruple of an agent, a time, a position, and a world. Why a position? Given the world, agent, and time, isn’t the position determined? For

33. Kaplan, pp. 505–6, and clause 3 of the definition of LD structures, p. 543.
someone can be in only one position at a time. Why not let “here” refer always to the position of the speaker at the time and world of utterance? But things are not so simple.

Imagine that we are looking through a pile of papers for a train schedule. “It just isn’t here,” I say. “It’s here, it’s here,” you reply, “we just aren’t looking carefully.” Now, if the content of “here” were the location of the speaker, then, since we obviously have different locations, the theory of indexicals would have to tell us that we are not really disagreeing. But it is clear that there is a disagreement. So indexicals must function in a more subtle and complicated way than the simplest indexical account would allow. In many instances “here” may refer to an area of common interest to the parties engaged in a discussion or cooperative activity, say, the volume between them or the intersection of the volumes they are facing. Precisely which volume it is must be determined in a complicated way by the conversational context, and it might change from sentence to sentence or even from phrase to phrase.

Consider next “nearby.” Surely it is an indexical, for if I now say “Jaegwon is nearby” I express a different fact from the one expressed by John Perry if he utters the same sentence in Palo Alto. But what is the character of “nearby”? Is it a function from a context to the property of being sufficiently close to the position of the context?

Claude and Patricia are coming to dinner, but they have run out of gas on Route 95. Claude calls from his car phone. “We ran out of gas half a mile past exit 17,” he says. “Don’t worry,” I reply, “You passed a gas station a few hundred yards back.” Claude tells Patricia, “Jamie says there is a gas station nearby.” In Claude’s sentence, “nearby” refers to the property of being within a few hundred yards of the speaker.

But now consider a different version. This time Claude calls and reports, “We ran out of gas, but don’t worry, because we passed a gas station a few hundred yards back.” I turn to you and say, “They have run out of gas, but Claude says not to worry because there is a gas station nearby.” In my sentence, “nearby” refers to the property of being within a few hundred yards from the agent mentioned in the sentence.

So the character of “nearby” must be a function from conversational context to the property of being interestingly close to a salient position. Standards for being interesting and salient are complicated, but surely they must enter into the character of the word. The point is that it is not merely ad hoc to introduce complexities into indexical semantics for moral terms. Natural language is not designed by logicians, and we should not suppose it will be simple and well behaved.

The general form of the analysis of moral terms will still be this:

“x is good” means “x is approved of by M,” where M is filled in contextually.

It seems to me that, given a context of use, we determine the content of the moral term by looking for the most suitable candidate for a moral system in that context.
What makes a moral system suitable? Probably many factors. But I think the two most influential factors are (1) the content and subject matter of a rule or principle (or standard, etc.) and (2) the motivational and affective connections with the speaker.

In this article I have been stressing the second of these. I started by allowing the speaker's actual motivational states to determine completely the relevant moral system. The modification in response to counterexamples to crude internalism allowed the moral system to be picked out by the speaker's motivations under normal conditions. The sophisticated relativist will look to what Westermarck called the "retributive emotions." So elements of the system will be rules which are such that if the speaker violates them he will tend to feel guilty; if others violate them he will tend to feel indignation, and so forth. These affective states are generally thought to belong to the conative as opposed to the cognitive powers of the mind, but it is not obvious that they are invariably motivating. It is the potential gap between having the states and having a motivation that Stocker exploits in questioning crude internalism.

On the other hand, it is the first of the most important factors, that of content or subject matter, that comes into play in our thoughts about the Sade example. The idea is that what can count as a moral principle is settled in part by what it is about and what it says. It was perhaps this part that noncognitivists neglected, as Foot points out in "Moral Arguments" and "Moral Beliefs." She argues persuasively that for a principle to be recognizably moral it is insufficient that its proponent display the relevant moral attitudes. The principle must be concerned with, and be concerned to advance, the well-being of persons (or, we may add, sentient beings in general). Now if such a content restriction were substantial enough to fix entirely the content of a moral system, there would be no room for relativism. We would have to say, there can be only one genuinely moral system: that which correctly specifies the rules a person must follow in order to advance the well of sentient creatures.

I think we have reason to doubt that content restrictions on moral systems are so substantial. There could be different conceptions of what constitutes well-being for a person or other sentient creature, different weights assigned to various aspects of welfare, and different views about

34. This is as good a time as any to mention the dangers of the "conditional fallacy." You might think that the relevant affective attitudes are the ones the speaker would have if conditions were normal. But that cannot be quite right. For it might be that the judgment in question is one the speaker would have no reason to make if conditions were normal; perhaps she judges that it would be a morally better thing if conditions were closer to being normal than they in fact are. So we must be careful to say that the attitudes in question must be the ones the speaker would have under normal conditions, toward the situation as it in fact is.
35. Westermarck, p. 63.
the importance attaching to the thriving of family and friends as opposed to strangers and members of other species. So it seems to me that the details, as it were, must still be filled in by the context and usually most importantly by the affective state of the speaker.

Now most examples of moral discourse are “normal” ones. Speakers do not suffer from affective abnormalities, and content restrictions on the system determined by their affects are easily met. The two influential factors for determining a moral system from a context can work together, and the simple form of speaker relativism is adequate to explain the internalist connection. Complications arise in abnormal examples. Affective failures pull apart the attitudinal elements that normally converge on a single system. Crude internalism can fail, and speaker relativism looks to the idealized version of the speaker to find a moral system uniformly supported by the emotive reactions. This explains the peculiarity of the Stocker-style counterexamples.

In the Sade-style counterexamples to crude internalism, the simple determination of a moral system by the character of moral terms breaks down for a different reason. It is the content factor that is problematic. For the system determined by the Sadists’ motivational states simply fails to meet those requirements. We cannot see the Sadists’ practical principles as advancing the welfare of sentient beings; of course not, since that is precisely what those principles are not designed to do. So the content factor is at war with the affect factor.

Because the Sadists’ motivations are so drastically out of sync with our requirements on the content of a moral system, the character of the moral terms tells us to reject those motivations as a candidate for giving content to their judgments. Instead, we search the context for a more suitable moral system. And in the Sade example, there is a more suitable candidate: it is the system constructed from the attitudes of the larger community. It helps that the thoughts and judgments of the Sadists themselves are focused around the conventional morality, so that our attention is drawn directly thereto.

That said, one may complain that I still have not specified precisely the character of moral terms. I say that the character takes the elements of the context and balances content and emotive considerations to choose a moral system or fill out the content, “x is highly evaluated by M.” But how does the character balance those factors? Precisely what function is it? I cannot say. I think it is a vague matter what weight is assigned to the two considerations, or to others if others turn out to be important. It is plausible to think that David Lewis’s “conversational scoreboard” may play a role in setting one or another system as the one named by “M” in my analysis. Vagueness will explain why intuitions may vary as to exactly what counts as a counterexample to crude internalism. But an

indexical account of moral terms retains some relativistic punch as long as there is a role for considerations other than purely content restrictions. Likewise, internalism remains compelling (in its modest formulation) to the extent that affective, motivational factors are important elements of the character of moral terms. So the connection between internalism and relativism is a deep connection, one which survives complication of an indexical account of ethical terms.

CONCLUSION

In this article I have argued that the most plausible kind of internalism, what I called “modest internalism,” is explained by speaker relativism. It explains modest internalism better than the Moral Sense Theory, because it does not conflate the motivation mentioned by internalism with a belief about motivation, and the Moral Sense Theory does. It is preferable to noncognitivist views, themselves likely candidates for explaining internalism, because it accounts for the straightforward functioning of moral sentences in contexts in which we expect to find truth-bearing expressions. I pointed out an analogy between the problem of explaining internalism and the problem of the essential indexical, and suggested that the solutions might naturally be thought of as parallel if we adopt speaker relativism as a metaethical theory. Finally, I defended my choice of speaker relativism as against a more socially oriented version.

Of course, this does not amount to anything like a proof. For all I have said, there may be other ways of explaining internalism without relativistic commitments. What’s more, there may be objections to even the modest form of internalism I presented, and if those objections are decisive there may be nothing left of internalism to explain. On the other hand, if you take internalism seriously as a thesis about moral assertions and beliefs, whether or not you find it compelling as stated here, then I have given you some reason to take speaker relativism seriously, too. That has been the main object of this article.