

Different Kinds of Kind Terms: A Reply to Sosa and Kim<sup>1</sup>

by  
Geoffrey Sayre-McCord  
University of North Carolina at Chapel Hill

In "'Good' on Twin Earth" I set about defending a semantics for moral terms that builds, in important ways, on recent developments in the semantics of natural kind terms without treating moral terms as natural kind terms. My hope is that the account that emerges can claim certain advantages not available to others even as it avoids seeing moral terms as referring to natural kinds. Putting my hope in this way makes clear, I trust, just why the comments by Professors Sosa and Kim are so relevant. Professor Sosa's main concern is that the primary advantages I claim for the semantic approach I favor are actually features of *any* plausible semantic view.<sup>2</sup> And Professor Kim's main concern is that my view will be metaphysically acceptable only if I give up the idea, which is central to my approach, that moral kinds are distinct from natural kinds.<sup>3</sup> In effect, then, Professor Sosa wonders why one should bother trying to defend the view in the first place, while Professor Kim suggests that the resulting view is either defective or not distinctive. One attacks the view going in, the other coming out. Let me say something about what comes in between, before turning to their worries.

The View Being Defended

The semantic theory I build upon holds that the reference of certain kind terms is determined not by concepts or descriptions that are associated with them by competent speakers but instead by the use of those terms bearing an appropriate causal connection to instances of the kind in question. And it holds that the contribution such terms make to the truth-conditions of the claims in which they appear depends on the nature of the things in their extension. The theory, applied to natural kind terms, is now quite familiar and more than a few

people have suggested that it might be applied directly, without change, to moral terms as well.

Applying the theory directly to moral terms involves holding that they refer not to whatever satisfies some concept or description we associate with them but (if they refer at all) to the natural kinds the instances of which causally regulate, in the appropriate way, our use of those terms. Exactly which natural kinds these might be is not at all clear, nor is it clear what precisely is required in order for some instance of a kind to count as appropriately regulating our use of the terms. Yet, however the details might be filled in, there is good reason to think that treating moral terms as natural kind terms will get things wrong, as Timmons and Horgan (and others) have argued.

The argument from Graceland (as I call it) runs as follows. (i) *If* the natural kind semantics were appropriate to moral terms, then we should be willing to acknowledge that another community's term 'good' (say) cannot be translated by ours if we discover that their use of their term is causally regulated (in whatever counts as the appropriate way) by things that fail to fall within the same natural kind as the things that causally regulate our use of the term 'good'. (ii) But, as a suitably constructed Twin Earth example shows, we are not so willing. Therefore, (iii) the natural kind semantics is not appropriate to moral terms.

This argument turns on the fact that the natural kind semantics commits those who embrace it to the following claim: if the terms used in two communities are causally regulated (in the appropriate way) by things that fail to fall within the same relevant natural kind, the terms cannot rightly be seen as intertranslatable (nor should they be seen making the same contribution to the truth-conditions of the sentences in which they are used). And that claim, I agree, gets things wrong when it comes to moral terms. Under suitable conditions, our term 'good', for instance, might well be correctly translated by another community's (perhaps orthographically identical) term despite the two terms being causally regulated by things that are not instances of a single relevant natural kind.<sup>4</sup> I maintain, though, that the suitable conditions -- the conditions under which we will see the terms as intertranslatable -- are just those in which the terms *are* causally regulated (in the appropriate way) by things that fall within the same *moral* kind, where things count as falling within the same

---

<sup>1</sup>. I am grateful to Nicholas Bull, Jon Harrison, Sean McKeever, Michael Ridge, Michael Smith, and Jon Tresan, for helpful discussion.

<sup>2</sup>. See his "Water, Drink, and 'Moral Kinds'" in this volume.

<sup>3</sup>. See his "Moral Kinds and Natural Kinds: What's the Difference -- for a Naturalist" in this volume.

---

<sup>4</sup>. Another community's use of 'good' might be causally regulated in the appropriate way by pleasure, say, and still we might correctly translate their term by ours even if our term is pretty clearly causally regulated by things other than what is pleasant. And it looks too as if two terms, used in different communities (say 'pleasure' in ours and 'good' in theirs) might be causally regulated by things that fall into the same natural kind (being pleasant) even though one cannot adequately be translated by the other.

moral kind if, but only if, those things are instances of some kind countenanced as significant by the best moral theory. The main purpose of my paper was to argue that the general semantic theory I rely upon could easily, and in an illuminating way, accommodate the intuitions that underwrite the argument offered by Timmons and Horgan, as long as talk of moral kinds is substituted in for talk of natural kinds.

The effect of the substitution comes out once the original argument is altered to take account of it:

(i') *If* the moral kind semantics were appropriate to moral terms, then we should be willing to acknowledge that another community's term 'good' (say) cannot be translated by ours if we discover that their use of their term is causally regulated (in whatever counts as the appropriate way) by things that fail to fall within the same *moral* kind as the things that causally regulate our use of the term 'good'. (ii') But, as a suitably constructed Twin Earth example shows, we are not so willing. Therefore, (iii') the moral kind semantics is not appropriate to moral terms.

The second premise here looks much less plausible than it did in the original argument. For, to put things a bit too crudely, the Twin Earth example on offer would have to be one in which the other community's term is causally regulated (in the appropriate way) by things that are not good.<sup>5</sup> This fact about their use of the term will itself (and quite independently of any commitment to a particular semantic view) suggest that what falls within the extension of their term is in relevant respects different in kind, because not good, from what falls within the extension of ours. And if what falls within the extension of their term is not good, then their term and ours must differ in meaning (if they mean anything at all) and we won't have the right sort of Twin Earth example.

This point can only be pushed so far, of course, since someone might insist that other respects in which they use their term as we use ours are sufficient to support the claim that their term and ours have the same meaning (so that what actually falls within the extension of their term is like what falls within the extension of ours in being good). But the plausibility of insisting on this diminishes significantly once we notice that what apparently falls within the extension of their term -- the things they actually identify as 'good' -- are not just different from what falls within the extension of ours, but different precisely in not being good. This suggests that the argument from Graceland tells against

---

<sup>5</sup>. This is too crude, as I say. It might be that some of what causally regulates their use of the term is actually good. The key difference between our community and theirs would be that in their community whether or not the things are actually good seems to be irrelevant to whether they fall within the extension of their term 'good'.

seeing moral terms as natural kind terms but does not seriously threaten the suggestion that moral terms refer to whatever the relevant kind is the instances of which causally regulate the use of those terms. And it is this part of the semantic theory that I set about defending in the face of the argument from Graceland.

### Why Bother?

The natural question to ask at this point is Sosa's: why bother? What advantages come with developing a semantic theory that works with a notion of moral kinds in the way I've suggested?

Sosa observes that, in discussing the suggestion that moral terms rigidly designate natural kinds, I explicitly cite four considerations in its favor. I say that (i) the view allows that "moral properties might be natural properties even though moral terms are not synonymous with non-moral terms," (ii) it "makes sense of the fact that people with dramatically different beliefs concerning value nonetheless often seem to be talking about one and the same property," (iii) it "fits well too with our thinking that as our own moral views evolve and shift they might continue to be views about the same thing," and (iv) it "explains easily the supervenience of the moral on the non-moral, since it identifies the properties referred to by our moral terms with properties referred to by certain non-moral terms." Sosa then argues that *any* plausible semantic theory can similarly claim these advantages. So, he concludes, there is no special reason to be attracted to this view as opposed to others and therefore no special reason to be interested in altering the view in a way that accounts for the argument from Graceland.

Before turning to Sosa's argument, I should note that the moral kinds semantics I advance cannot claim all these advantages, precisely because it rejects the view that moral terms refer to natural kinds. In particular, the view I am advancing does not -- or at least does not inevitably -- identify moral properties with natural properties (even though I do think it is compatible with a respectable form of naturalism) so it cannot claim (i); and because it does not (inevitably) identify the two kinds of properties, it cannot explain supervenience in the way (iv) suggests (even though I do think the view can explain supervenience).

At the same time, I should also note that the view I advance can claim, in addition to (ii) and (iii), a number of other advantages. Most important among these, I believe, are: (v) that it accommodates well the arguments standardly offered for thinking of some terms that their reference is not determined by a description or concept with which they are associated by competent users of the term, when these arguments are applied to moral terms<sup>6</sup>; and (vi) that it explains

---

<sup>6</sup>. These are the arguments I summarize in the paper in the form of tricked up versions of the Open Question Argument, that taken together suggest that satisfying some concept or description

what is plausible in the Graceland argument -- that what falls within the extension of moral terms can be, from the point of view of science, stunningly heterogeneous.

These points noted, let me return to Sosa's argument. He is surely right that a feature or set of features shared by all plausible semantic views cannot be cited as a distinct advantage of one view as opposed to the other plausible views. But it is important to register two cautions. First, if few other views have the features "any plausible view will have," then the fact that a view has those features reasonably counts as an advantage. Second, since different plausible views might accomplish various tasks more or less well, that one view does the job especially nicely, should count as an advantage as well.

As it happens, I think the features I mention are not available to the broad range of semantic views that see the reference of a term as being determined by a concept or description associated with that term by competent users. And in the paper I try to highlight the sort of arguments that, when applied to moral terms, would support this view, although I do not spend a lot of time applying those arguments to moral terms. The force of these arguments turns on mobilizing two intuitions:

(a) that people who fail to associate the same concept or description with a moral term -- either because what they do associate with it is so different (as would be the case with 'good' if one is a devout hedonist and the other not) or because they don't associate much at all with it (as is the case, I think, with many people who are competent with the term 'good') - - may nonetheless be using the term with the same meaning, and

(b) that people who associate the same concept or description with a particular moral term may nonetheless fail to be using it with the same meaning if the context in which they use it is such that what falls within the extension of one person's term is different in moral kind from what falls within the extension of the other's term.<sup>7</sup>

The first intuition is, I think, readily mobilized and gains support immediately from the fact that we regularly see people with wildly divergent

views concerning the nature of value as disagreeing with, rather than talking past, each other. The second intuition is harder to bring out, not least of all because there is no consensus among us as to the nature of value, rightness, etc., so no consensus as to what would have to be different as between the extensions of the term as used by different people. Still, the intuition emerges, I believe, with the help of suitably described scenarios. I will not here present even a sample of the scenarios that would be needed. In light of the variety of moral views people actually hold. But I do want to underscore what such scenarios would bring out: that, if there were a community in which the term 'good' was consistently applied, by those competent with the language, to things that are not good, we would be inclined to say that their term must mean something different from ours, whatever concept or description they happen to associate with it. No doubt, the inclination would give way under many circumstances, and in particular if we could explain *as mistakes* the consistent application of their term to things that are not good. Even so, imagining cases where it would seem wrong to see them as making a mistake with their own term is not all that hard. And these are the cases in which the difference in moral kind, as between the extension of our term and the extension of theirs, motivates seeing their term as differing in meaning from ours.

As it happens, too, I think the moral kind semantics does an especially nice job of securing the advantages I mention. It is not simply compatible with the various observations the advantages highlight, it helps to explain them. Moreover, the account need not attribute to people concepts they show no overt signs of possessing just to maintain that their terms are actually associated by them with the same concepts we might associate with ours. Nor does it need to see what look to be very deep disagreements about the nature of morality as cases of people talking past each other. Nor does it have to hold that the reference of our moral terms is, in the end, to be settled by an appeal to science. And it makes good sense as well out of the ways in which we alter our own views about what falls within the extension of our terms as we change our views about the nature of morality even as we continue to think of ourselves as having views about one and the same thing. Thus, it makes good sense out of the stability of reference in the face of often quite drastic changes in theory and belief.

So I think the view I advance can legitimately claim some distinct advantages over views that see the reference of a moral term as being determined by some concept or description associated with that term by competent users of it. Yet I purposely did not argue that the view was better than, let alone a competitor against, either a Fregean theory or a meaning-as-use theory (the two alternatives Sosa mentions). For all I argue, my preferred view might in some way be assimilated by either. Given a suitably sophisticated view of Frege's theory, for instance, a case might be made for thinking that a moral kinds

---

associated with a term by those competent with it is neither necessary nor sufficient for something to count as a referent of the term.

<sup>7</sup>. In using the Open Question format in the way I did, I did not intend to suggest that we need the semantic view I favor in order to take account of successful co-reference. Rather, we need it for an adequate account of disagreement (in contexts where people hold substantially different views about, say, the nature of value) and to accommodate semantic intuitions to the effect that satisfying some concept that is associated with a given term is neither necessary nor sufficient for falling within its extension in a given world.

semantics articulates (in perhaps surprising ways) the Fregean sense(s) of moral terms.<sup>8</sup> What is important, for my purposes, is that, *if* sense determines reference, then the sense of a kind term should not be identified with a concept or description associated with the term by competent users, since the sense would have to implicate, as at least partially determinative of reference, facts about the true nature of what falls in the extension -- facts that might not be available to competent users of the terms. Similarly, someone might defend a meaning-as-use theory, while embracing the claim crucial to my view, by holding that what is distinctive about the use of moral terms -- the aspect of their use that is central to their meaning -- is that they are used to refer to whatever it is in virtue of which what falls within the extension of the term count as instances of the same moral kind of thing.

I would not myself want to defend either of these suggestions, but the arguments I offered in "Good' on Twin Earth" were not directed against them. So the virtues I cite are not offered as virtues that they could not share. My aim, instead, was to argue that, if they are to share the virtues, they must in some way accommodate the fact that the reference of our moral terms is not determined by some concept or description competent speakers associate with the terms but instead, at least in part, by what turns out to be the true nature of what falls within the extension of those terms -- where moral theory, not science, plays the roles of determining their nature.

#### Can a Naturalist Love it?

Clearly, a crucial part of my view turns on distinguishing moral kinds from natural kinds, on emphasizing the difference between moral sortals and the sortals of natural science. Assuming, as a naturalist will, that moral theory and scientific theory cover the same domain, one cannot avoid Kim's question: how are the taxonomies of moral theory and science related? "Are they," Kim wonders, "mutually concordant in the sense that the boundaries drawn by one are respected by those drawn by the other, there being no crisscrossings between the two sets of boundaries? Or do the natural and the moral boundaries randomly crisscross each other?" Kim suspects that facing these questions will reveal an unresolved tension in my view.

In answering the questions, I need to distinguish the commitments of the semantic theory *simpliciter* from the commitments that emerge once the

---

<sup>8</sup> I am skeptical about securing a satisfying reconciliation, but this is not the place to explore my doubts. In any case, see David Wiggins' "Putnam's Doctrine of Natural Kind Words and Frege's Doctrines of Sense, Reference, and Extension: Can They Cohere?", and Michael Dummett's "Frege's Distinction Between Sense and Reference," both in Meaning and Reference, edited by A. W. Moore (Oxford University Press, 1993) pages 192-207 and 228-256, respectively.

semantic theory is combined with other commitments, say to supervenience and naturalism. That I insist on distinguishing these may help to explain why Kim detects an ambiguous message in my paper, in some places it seeming as if I think of the two taxonomies as independent, in others it seeming as if I think the boundaries among moral kinds must, in certain ways, respect the boundaries among natural kinds.

In any case, the semantic view I am defending -- considered in the absence of a commitment to naturalism -- is silent about how moral kinds line up with natural kinds. Whether two things count as being of the same moral kind depends only upon whether they are morally homogeneous. Two things may well be morally homogeneous even if they are heterogeneous from the point of view of science; and two things that are heterogeneous from the point of view of moral theory may well be homogeneous (even indistinguishable) from the point of view of science. The two taxonomies are independent... so far. Adding a commitment to supervenience into the mix doesn't change things. Supervenience does bring in the requirement that a difference in moral kind cannot be an insulated difference -- there must be some other difference that explains it. Prescinding from naturalism, though, the commitment to supervenience allows that things that differ morally might not differ in their natural features, even as it rules out the possibility that they differ only in that one moral respect. But add in some version of naturalism according to which any two things that differ at all must differ in some way capturable, at least in principle, by science, and a constraint on moral kinds emerges. For while moral homogeneity will still admit natural heterogeneity, moral heterogeneity will entail some natural difference (although the difference might be represented as a difference in degree rather than in kind within the taxonomy of science).<sup>9</sup> This much, I think, follows in the wake of a commitment to naturalism.

Kim suggests, however, that I may be committed to more -- to holding that moral kinds are natural kinds -- by my holding that moral kinds causally regulate (in some appropriate way) our use of moral terms. As he points out, if moral kinds play this role they "must have, or be, causal powers." And he wonders, "how can this be unless moral kinds are causal kinds?" If they are, then the apparent, and apparently significant, distinction between moral kinds and natural kinds disappears.

---

<sup>9</sup> Kim was puzzled by the distinction I draw here between differences in degree and differences in kind. I suspect the puzzlement comes from his switching, in effect, from talk of natural *kinds* to talk of natural properties (where any difference that can be registered by science counts as a difference in natural properties). I think the switch does not raise any serious difficulties as long as one keeps in mind that some sortals do admit of degrees.

On the view I am advancing, the reference of a kind term is established, initially, by its being associated (*via*, say, a suitable referential intention) with what is taken to be an instance of the kind, and the subsequent success of others who use the term to refer to that kind with the term is due to their use of the term bearing some appropriate (again, usually, causal) connection to instances of that kind. Thus the view does require that (some of) the instances of the kind in question figure in causal explanations of our use of the term. Yet it does not require that their being of that kind is causally relevant to that use; nor does it require that everything that falls within the kind be such as to share any particular causal powers or scientific significance at all.

Of course, when the terms in question are natural kind terms, they will succeed in referring (as this view would have it) only if what falls within their extensions do share some particular causal powers or scientific significance. But this is a reflection of the kind of kind terms at issue, not a reflection of the general semantics of kind terms. When the terms in question are moral terms, the general semantic theory allows that they may succeed in referring even if what falls within their extensions fail to share any particular causal powers or scientific significance. Thus the view does not require holding that the fact that what falls within the extension of our term 'good' is good is part of the causal explanation of our use of the term. Their being good is part of the explanation of why they fall within the extension of the term as we use it, but not necessarily part of the explanation of why we use the term 'good' in response to them.<sup>10</sup>

How is it, then, that we succeed regularly in using our moral terms to refer to the same kind of thing, if our use of the term is not causally explained by those things being of that kind? The answer lies in recognizing that once a kind term is in use, its role is to refer to the kind in virtue of which what falls within the extension of the term count as bearing the relevant (theoretically determined) same-kind relation to one another. Whatever actual things there are that bear the relevant same-kind relation to the instances of the kind fall within the term's extension even when those who use the term have no causal interaction with them. We succeed in referring to the same kind not because that kind causally explains our interactions with the world but because it characterizes accurately what the things we interact with have in common. When it comes to moral terms, moral theorizing reveals which, if any, such characterization is accurate.

Does such a view square with commitments to naturalism and supervenience? I think so, though neither commitment will directly follow from

---

<sup>10</sup>. So while my commitment to naturalism leads me to hold, in light of my semantic view, that the things in the world that fall into the extension of the term 'good' differ in some natural respect from those that do not, that they are good need not be part of a causal explanation of our use of the term.

the semantic view. Instead, such commitments get swept up by the semantic view as one succeeds in defending the idea that no moral theory could be acceptable unless it took account of them. Thus, if we have reason to think naturalism is true, then ipso facto we have reason to think an acceptable moral theory could countenance as morally significant only differences that are capturable by science. And the semantic view will have it that our moral terms refer to kinds that respect this constraint. Similarly, if we have reason to think supervenience is necessarily true, we have reason to think an acceptable moral theory must respect it. And, again, the semantic view will have it that our moral terms refer to kinds that respect this constraint.<sup>11</sup>

To accept this view is to see the reference of our moral terms as being hostage not to the results of science but to those of moral theorizing. For all the semantic view has to say, it may turn out either that the best moral theory reveals that some of our terms have no reference or that none of them do (if no moral theory is in the end defensible). What comes with a sincere deployment of the terms is a commitment to thinking there is a difference in kind to be discovered as between those things to which the terms apply and those to which they do not, a difference that itself can be seen as normatively significant.

---

<sup>11</sup>. If one holds that moral terms refers to natural kinds the supervenience of the moral on the natural follows trivially. It comes easily too if instead one embraces a functionalist *definition* of moral terms according to which they refer to whatever fills a certain causal role (say of promoting happiness, prompting admiration, or whatever). I reject both ways of securing supervenience, the first because it makes the reference of our moral terms inappropriately dependent upon the results of science, the second because it holds that our moral terms refer to whatever happens to satisfy the relevant functional description. I am myself tempted by the suggestion that what our moral terms refer to in fact satisfy functionalist characterizations, although I think (i) that the characterizations can not be defended as capturing what competent users associate with the term and (ii) that the plausible functionalist characterizations will themselves have to be specified in non-causal terms. In any case, appealing to these semantic proposals seems to me to be the wrong place to look for a defense of supervenience. Instead, the main argument for supervenience is to be found in normative theory, in the observation that distinctions that failed to respect supervenience would be arbitrary – they would be distinctions for which one could offer no grounds.

