

# Must Metaethical Realism Make a Semantic Claim?

**Guy Kahane**

Oxford University  
Littlegate House, St. Ebbe's Street,, Oxford, OX1 1PT, UK  
guy.kahane@philosophy.ox.ac.uk

---

## Abstract

Mackie drew attention to the distinct semantic and metaphysical claims made by metaethical realists, arguing that although our evaluative discourse is cognitive and objective, there are no objective evaluative facts. This distinction, however, also opens up a reverse possibility: that our evaluative discourse is antirealist, yet objective values *do* exist. I suggest that this seemingly far-fetched possibility merits serious attention; realism seems committed to its intelligibility, and, despite appearances, it isn't incoherent, ineffable, inherently implausible or impossible to defend. I argue that reflection on this possibility should lead us to revise our understanding of the debate between realists and antirealists. It is not only that the realist's semantic claim is insufficient for realism to be true, as Mackie argued; it's not even necessary. Robust metaethical realism is best understood as making a purely metaphysical claim. It is thus not enough for antirealists to show that our discourse is antirealist. They must directly attack the realist's metaphysical claim.

## Keywords

Metaethical realism, objective value, antirealism, moral semantics, metaphysics, error theory

## 1. Semantics and Metaphysics

What kind of concepts we have is one thing, what reality is like another. This was the lesson that Mackie taught us.<sup>1</sup> It is one thing to make the semantic claim that:

---

<sup>1</sup> J.L. Mackie, *Ethics: Inventing Right and Wrong* (London: Penguin, 1977).

**Realist Semantics (RS).** Our evaluative discourse is cognitive and objective.<sup>2</sup>

And another to make the metaphysical claim that:

**Realist Metaphysics (RM).** There are objective evaluative facts.

It's not enough for moral and evaluative realists to defend RS – for by itself this does nothing to show that there is anything out there that corresponds to our concepts. And those antirealists who assume that if RM is false then our evaluative discourse must be antirealist are making a similar mistake.

Once we draw this distinction between semantic and metaphysical questions, the logical space of metaethics expands, and the position that Mackie called the error theory comes into view. Mackie thought that both realists and antirealists got things wrong. For although our evaluative concepts are indeed cognitive and objective in content (as realists assert and those antirealists deny), there really isn't anything out there that corresponds to them.<sup>3</sup>

As a result of Mackie's insight, metaethical realism is now understood to be the conjunction of *both* RS and RM. Thus Sayre-McCord claims that "realism involves embracing just two theses: (1) the claims in question, when literally construed, are literally true or false (cognitivism), and (2) some are literally true".<sup>4</sup> And Shafer-Landau assumes that "realism is a form of cognitivism", and that "every realist is a cognitivist".<sup>5</sup>

Now there are many brands of metaethical realism. Mackie's semantic claim was actually rather specific. He claimed that our evaluative concepts aim to refer to 'objectively prescriptive' facts, something very close to the response-independent, non-natural and irreducible evaluative and normative facts that have been recently defended by Parfit, Huemer, Shafer-Landau, Enoch and others.<sup>6</sup> To distinguish this view from naturalist

---

<sup>2</sup> Mackie doesn't always distinguish between moral value and value in general, and between value and practical reasons – all of which are typically at stake in disputes between metaethical realists and antirealists. To save words, I'll also largely ignore these distinctions. With Mackie and most contemporary metaethicists I assume that the truth of RS or competing semantic claims depends in familiar ways on facts about our actual linguistic discourse.

<sup>3</sup> The error theory is, in one good sense, also an antirealist view, but for simplicity I will treat it as a distinct view.

<sup>4</sup> G. Sayre-McCord, 'Introduction: The many moral realisms', in G. Sayre-McCord, ed., *Essays on Moral Realism* (Ithaca: Cornell University Press, 1988), p. 5.

<sup>5</sup> R. Shafer-Landau. *Moral Realism: A Defence* (Oxford: Oxford University Press, 2003), p. 17.

<sup>6</sup> See Shafer-Landau, *Moral Realism: A Defence*; D. Parfit, 'Normativity', *Oxford Studies in Meta-Ethics*, 1 (2006), pp. 325-380; *On What Matters* (Oxford: Oxford University Press,

or weaker forms of realism, I shall call it *robust realism*. In what follows, I will assume with Mackie that realism takes such a non-naturalist form, and I will take RS and RM to refer to these stronger claims. On the other hand, although my discussion is meant to apply to metaethical antirealism quite generally, I will for simplicity mostly use a familiar type of response-dependent theory of value as my example of an antirealist view.

What I intend to explore here isn't the error theory, but another way in which Mackie's distinction extends the logical space of metaethics. For there is a further possibility that is not considered by Mackie or, to the best of my knowledge, by other metaethicists.<sup>7</sup>

Suppose that RS is false. It's often assumed that this amounts to victory for antirealism. After all, if realism is understood as the conjunction of RS and RM, then the falsity of RS would also imply the falsity of robust realism. But what if RM is nevertheless *true*: what if although our concept of value is response-dependent (or antirealist in some other way), strongly objective values *do* exist? If reality can fail to meet the conditions set by our discourse, as claimed by the error theory, why can't it meet conditions *not* set by our discourse? I call this possibility the *reverse theory*.<sup>8</sup>

Although this possibility might seem far-fetched, I will argue that it merits serious attention – in fact, I will argue that robust realists are *committed* to its intelligibility. Why then have metaethicists so far completely ignored it? Perhaps they assume that the reverse theory is simply incoherent, or that, even if it is coherent, it is ineffable, or too implausible to be worth

---

2011); M. Huemer, *Ethical Intuitionism* (London: Palgrave Macmillan, 2006); D. Enoch, 'An outline of an argument for robust metanormative realism', *Oxford Studies in Metaethics*, 2 (2007), pp. 21–50. Nagel and Scanlon can also be interpreted as defending such a view. See T. Nagel, *The View From Nowhere* (Oxford: Oxford University Press, 1986); T. Scanlon, *What We Owe To Each Other* (Boston: Belknap Press, 1998). These realists would of course object to the way Mackie understands objective prescriptivity. Some of these robust realists claim that their view is 'non-metaphysical' because although they believe that there are objective non-natural evaluative *truths*, they deny that this commits them to ontological claims about the *existence* of evaluative *properties*. The contrast I draw between semantics and metaphysics, and the claims I will be making about realism, are meant to be compatible even with this view.

<sup>7</sup> Kalderon mentions this possibility in passing in M.E. Kalderon, *Moral Fictionalism* (Oxford: Clarendon Press, 2005), pp. 143, 145–6.

<sup>8</sup> The reverse theory shouldn't be confused with views that attempt to combine realist and antirealist theses, such as views that combine cognitivist and expressivist elements (D. Copp, 'Realist-expressivism: a neglected option for moral realism', *Social Philosophy and Policy* 18 (2001), pp. 1–43), or with the complaint that highly refined antirealist views, such as Gibbard's, are hard to distinguish from straight realism (J. Dreier, 'Meta-ethics and the problem of creeping minimalism', *Philosophical Perspectives*, 18 (2004), pp. 23–44).

discussing. Or they might suspect that that it wouldn't matter whether the reverse theory is true, or that there is no way we could possibly defend it. In what follows I will try to address these worries. We shall see that reflection on the reverse theory should transform the way that we understand metaethical realism, and the debate between realists and antirealists. In particular, and contrary to a common assumption, it will emerge that robust realism is best understood as a purely metaphysical thesis with no commitment to cognitivism or related semantic claims – and that this metaphysical thesis can be defended *whether or not* these semantic claims turn out to be true.

These implications follow if, as I shall argue, the reverse theory describes a genuine possibility. If my argument fails, this shall have other implications. If the intelligibility of the reverse theory follows from basic commitments of robust realism, then, if we cannot make sense of the reverse theory, we should wonder whether we can really make sense of robust realism.

## 2. Is the Reverse Theory Coherent?

Some philosophers believe that robust metaethical realism is incoherent. To the extent that they think so because they believe that RM is incoherent, then their metaphysical doubts would apply with equal force to the reverse theory. But although there are these familiar doubts, robust realism is a metaethical view that has been much discussed, and that is being taken increasingly seriously. The reverse theory is not, in this respect, any *less* coherent than standard robust realism. The question is whether it might be incoherent in some *further* way – not whether RM is incoherent but whether the *conjunction* of RM with the falsity of RS is incoherent and therefore justifies dismissing the reverse theory.

Few deny the coherence of the *error theory*. If the error theory is coherent, then RS doesn't imply RM. For the reverse theory to be coherent, the converse claim needs to be true: that the falsity of RS doesn't imply the falsity of RM. But if anything, this direction of logical independence is more plausible. RM is a claim about facts out there, and their independence from us. If there are such robustly objective evaluative facts, how could their very existence depend on what kind of discourse we happen to have?

Robust realists must, I believe, accept both directions of logical independence. It is familiar that robust realists take value to be mind

independent – to exist independently of our desires, beliefs and other mental states.<sup>9</sup> It is less often noted that, with strong realist views in other domains, robust metaethical realists are also committed to holding that:

**Discourse Independence.** Robustly objective facts and properties are metaphysically independent of our practice, discourse or conceptual scheme.<sup>10</sup>

It is not only pressure by error theorists such as Mackie that should lead realists to accept that RS doesn't imply RM. Robustly objective values just *aren't* the kind of thing whose existence could be guaranteed by semantic fiat.

From Discourse Independence it is but a small step to the conclusion that:

- (1) Even if objective value exists, it is still possible that agents exist whose evaluative discourse is antirealist.

Even if RS and RM are true, surely we can conceive of *other* cultures whose evaluative discourse is nevertheless antirealist. This seems undeniable. But what is true of others could also be true of us, and our own discourse:

- (2) If objective value exists, it would still exist even if our discourse *had been* antirealist.

Even if RS *is* in fact true, surely it *might* have been false. And even if RS and RM are true, surely we might in the future mistakenly come to accept Mackie's error theory, and, in response, replace (as Mackie seems to recommend) our realist evaluative concepts with alternative response-dependent ones. This again seems undeniable.

The step from Discourse Independence to (1) and (2) would be blocked only if RM implied RS or, even less plausibly, if RS was (in a non-trivial sense) a necessary claim. Yet surely it is a contingent matter whether our evaluative discourse is cognitive and objective, or has some other metaethical character. The scenarios I have just described are easily conceivable.<sup>11</sup>

---

<sup>9</sup> See Shafer-Landau, *Moral Realism: A Defence* for a more precise statement of mind (or 'stance') independence.

<sup>10</sup> It is hard to see how a view that rejected this thesis could count as a robust form of realism; but it seems to anyway follow from mind independence.

<sup>11</sup> The only way I can see of arguing that if RM is true then RS must be true is to appeal to something like Lewisian 'reference magnetism' (D.K. Lewis, 'New work for a theory of

We might, in the future, come to adopt an antirealist evaluative discourse. But if something might happen in the future, it could also have happened in the *past* – indeed, for all we know, it might have *actually* happened in the past. In fact, it is not implausible that the increasing acceptance of sceptical relativist views by many educated people has already amounted to such a semantic shift. If this is the case, then the reverse theory might actually hold.<sup>12</sup> This again seems perfectly intelligible. This scenario would amount to the more radical possibility that:

- (3) Objective value might exist even if our *actual* evaluative discourse is antirealist.

But no past semantic shift is needed for (3) to hold. There needn't have been any 'fall' from objectivity. It's at least possible that our evaluative discourse was *always* antirealist – despite the fact that objective values were also always out there. After all, if some antirealist view gives the correct account of our discourse, this by itself says nothing about what does or doesn't exist. It's entirely compatible with the existence of facts about objective value.<sup>13</sup>

Most metaethicists assume that, before we can ask whether RM is true, we must first establish the truth of RS. But we saw that even if RS is true, we might, at some future point, turn our evaluative discourse antirealist, as some error theorists recommend. Presumably, if RM *is* true, this semantic shift could not make it any *less* true. Objective values won't suddenly vanish into thin air. (Compare the more extreme example of a community taken over by fanatic verificationists who understand statements about the past to be about presently available evidence. Would this semantic revolution cut a hole in space-time and make the past disappear?)

---

universals', *Australasian Journal of Philosophy*, 61 (1983), pp. 343–377), if it can be extended to cover non-natural properties. Such a view seems to me implausible. But although it would rule out the reverse theory, its upshot would be anyway identical to the upshot of this paper: that what matters is RM, regardless of standard semantic criteria for the truth of RS.

<sup>12</sup> A. MacIntyre, *After Virtue* (Notre Dame: Notre Dame University Press, 1981) can be interpreted as arguing for this, or something close enough (I owe this example to Jussi Suikkanen). Or it might be argued that at least *some* of the many who espouse antirealist metaethical views are in fact correctly describing *their own* evaluative concepts, if not ours. This appears to be Parfit's view in *On What Matters*.

<sup>13</sup> If RM can be true only if RS is true then, given the contingency of RS, this absurdly implies that the existence of objective value is also contingent, indeed contingent on what kind of discourse we happen to have, which itself presumably depends on a range of contingent facts about us.

This, however, should also be true when we consider a past semantic shift to antirealism. If we discover that, although our evaluative discourse is presently antirealist, it nevertheless *used* to be realist, it would be absurd to conclude from this that RM *might* have been true in the past, but, because RS is no longer true, we can now simply dismiss it. But once all this is granted, it is hard to see why the contingent fact that RS is now true or was true (or, for that matter, could be true in the future) makes *any* difference to whether RM is true. It seems, then, that the metaethicists' assumption is mistaken.

This conclusion may be too quick. Although (1) and (2) are hard to deny, (3) is a puzzling claim, and it is only (3) that directly challenges the metaethicists' assumption. To be sure, the step from (1) and (2) to (3) can seem simple. How could the reverse theory describe perfectly intelligible scenarios about others or our counterfactual or future selves, yet nevertheless could never be actually true of us *now*? (Or put differently, if we can truly assert (1) and (2), why can't those we refer to in these claims truly assert the proposition in (3)?) But (3) also raises great difficulties and, as we shall now see, it is not clear what it even means, and whether we can intelligibly assert it.

### 3. Is the Reverse Theory Ineffable?

The possibilities in (1) and (2) describe counterfactual, or future or third-person contexts in which RS is false yet RM still true. When stating these possibilities, I *assumed* that RS is true, and could thus employ our supposed *actual* objective evaluative discourse to state RM. But (3) is a very different claim. For if our actual discourse is *response-dependent*, then how can we even state RM?

Indeed, it might be objected that I formulated RS in a misleading way. If the semantic claim in RS explicates our concept of value, then the subsequent metaphysical claim should simply read:

There are evaluative facts.

But if it's not realism but some form of antirealism that gives the correct account of our normative concepts, then *this* metaphysical statement could *still* be true. It would now, of course, express a different proposition, and refer to very different properties and facts. It might be true, for example, because (to pick one common response-dependent view),

There are facts about what we would desire in ideal conditions if we were procedurally rational and fully informed.

But if so, what space is left to say that *objective* values exist? What concept of value would we be using in stating RM? It won't be *our* concept of value, it seems, given that, by assumption, our evaluative discourse refers to psychological facts of the kind cited above. If the reverse theory can't even be stated, it's no wonder it's never discussed.

The worry is that if RS is false, we cannot intelligibly assert that the reverse theory actually holds. This wouldn't yet show that the reverse theory is not a genuine possibility. To repeat: objective value wouldn't suddenly fade away if we changed our discourse. All it would show is that if the reverse theory *actually* held, it would be *ineffable*.

But even this weaker worry is implausible. After all, the entire discussion so far seems to falsify this worry, since it precisely involved reference to the alleged ineffable possibility. Worse, since we can't claim that the reverse theory is ineffable without *referring* to it, this claim is self-defeating.

What this worry does bring out is that we need some vocabulary to state the reverse theory, and if the reverse theory holds (if RS is false) then this vocabulary couldn't be our everyday evaluative discourse. Thus, even if RM doesn't imply anything about our semantics, *discussing* it does presuppose the availability of a realist discourse that would allow us to refer to the relevant facts.

Such a vocabulary is readily available. When realists assert RS, they are making a claim about our actual discourse. But in doing so they are also describing a *possible* discourse – a possible set of realist concepts. And when antirealists defend their proposed semantics, they typically do so by contrasting it with RS, or less directly, by contrasting their semantic account with paradigmatic instances of realist discourse, such as talk about physical objects. So when antirealists deny that our evaluative discourse is cognitive, or response-independent, they are also gesturing at a possible realist alternative, and thus at a sense for 'objective value'. The vocabulary implicit in this discussion allows us to discuss (and assert) RM, *whether or not* RS is true.<sup>14</sup> It seems easy enough to ask: 'What would evaluative statements refer to, *if* we take them to refer to robustly objective facts?'<sup>15</sup>

---

<sup>14</sup> This is compatible with thinking that in order to understand this quasi-technical vocabulary, we must first grasp everyday evaluative concepts.

<sup>15</sup> This question is even easier to ask if, as many hold, the surface of our evaluative discourse at least *appears* cognitivist and objective. It would be especially straightforward if our evaluative discourse is fictionalist: the relation between pretending something to exist and asserting its existence is obvious enough. See Kalderon, *Moral Fictionalism*. (I am grateful here to an anonymous reviewer.)



In any case, antirealists of different stripes often explicitly deny, and argue against, RM. Few think that, because these antirealists all also reject RS, these denials and arguments aren't even intelligible. But if these denials are intelligible, then so is the reverse theory, which employs these very same propositions. If it's intelligible to deny *both* RS and RM (something antirealists repeatedly do), then it should also be intelligible to deny RS and assert RM.

In order to avoid confusion, and to make explicit what is already implicit in much metaethical discussion, I will explicitly stipulate that in what follows the phrase 'objective value' refers to:

The objective facts which, according to robust realists and error theorists, our actual evaluative discourse aims to refer to.

The exact content of this phrase is given by common realist attempts to elucidate RS, to distinguish it from alternatives, and to identify criteria that would establish its truth.

#### 4. How Could 'Objective Value' Refer to *Value*, if Antirealists Are Right about the Semantics of Our Actual Discourse?

Antirealists might complain that this is not enough. If the reverse theory obtains, then the terms 'value' and 'normative force' have a response-dependent meaning. So, by definition, these objective facts *couldn't* literally instantiate value, or have normative force, if these phrases are used in their everyday sense. They could only possess value in *some other sense*, and the realist still owes us an explanation of what that sense is.

My reply to the worry about ineffability was that these phrases have senses that transcend our actual semantics, senses already implicit in current metaethical discussion. But that reply doesn't fully address the present worry. The realist needs to explain why these facts *deserve* to be described using terms drawn from our actual evaluative discourse. It's hardly enough that we can single out some technical sense of the phrase 'objective value' that is independent of the actual semantics of 'value'. If we stipulate that the word 'value' can be used to refer to the starry skies, then, speaking in this technical sense, we could trivially truly assert the sentence 'objective values exist', but this would hardly be victory for realism. This alternative vocabulary must bear some appropriate relation to our actual evaluative discourse; the two must, so to speak, belong to a single family. The objection is that if this cannot be done, then it remains doubtful whether the reverse theory describes a genuine possibility.

Let me repeat first that if this worry has force, then it spells trouble not only for the reverse theory, but also for common claims in current metaethics, including claims made by many antirealists – indeed claims that it seems they *need* to make to explain what it is exactly they are trying to defend. If we can't address this worry, then it would follow that if some antirealist view gives the correct semantics of our evaluative discourse, then both RS and RM are not simply false but unintelligible. It would mean that metaethicists can't intelligibly debate questions about objective and subjective value, because one of these phrases is an empty tautology, the other meaningless or utterly irrelevant. This seems wrong.<sup>16</sup>

But I think we can do better. We can *explain* the relation between our actual evaluative discourse, if it is antirealist, and the metatheoretical sense of 'objective value' which we need in order to state the reverse theory. Here, in broad outline, is how I think this can be done.

Our evaluative discourse plays a certain role in our practical lives. Call this *the value-role*. What is this role? It's not easy to spell it out in entirely neutral terms, but the basic idea is simple: it's the role of setting a standard by which attitude and action can be made intelligible and justified, and in light of which we deliberate (in the first-person), and give advice or criticize (in the second- and third-person).

The value-role seems to be 'metaethically thin': we can describe it without making reference to metaphysical questions about response-dependence or independence, not even to semantic ones about cognitivism. Thus although it is a fact about our actual evaluative discourse that it plays the value-role, other forms of discourse that vary along these metaethical dimensions could also potentially play it, *even if their metaphysical commitments are different from those of our actual discourse*.

The metaethical thinness of the value-role is reflected by the point that we could identify the evaluative discourse of some utterly foreign culture without needing (or even being able) to determine whether its semantics is, say, response-dependent or not (this had better be true, given that we have trouble answering this semantic question about our *own* discourse!).

Our criteria for identifying evaluative discourse are metaethically neutral. Such identification wouldn't be defeated if it later turns out that the deeper metaethical commitments of this foreign discourse are different from ours. This point is brought out most clearly, I believe, when we

---

<sup>16</sup> A parallel worry would also afflict other philosophical debates, rendering, for example, the common contrast between compatibilist and libertarian freedom problematic.

consider the possibility mentioned earlier that our current evaluative discourse, although presently antirealist in character, is the result of gradual evolution (or mutation) from a past evaluative discourse that was robustly realist. This, I said, is a perfectly intelligible story – a story that some believe to be *true*. But to accept this story is *already* to accept that we possess a conception of ‘evaluative discourse’ that is neutral with respect to the semantic claims that typically divide realists from antirealists. We can understand both how, like our supposed actual discourse, such a past discourse guided deliberation, advice and criticism (this is what makes it *evaluative*) as well as how, *unlike* our actual discourse, it aimed to refer to certain objective irreducible properties (this is what makes it *realist*).

The metaethical thinness of what I called the value-role is also implicit in discussion of the error theory. When Mackie and other error theorists recommend that we replace our faulty realist evaluative discourse with some antirealist alternative, their suggestion seems perfectly intelligible because we can see how these antirealist forms of discourse could occupy the vacant value-role. When these error theorists recommend that we revise our evaluative discourse, they are not merely suggesting that we stipulate some arbitrary new meaning for the word ‘value’.<sup>17</sup>

In this way, different metaethical views (robustly realist, response-dependent, non-cognitivist, etc.) can be understood to not only offer competing semantic accounts of our actual discourse, but to also describe competing *conceptions of value* – forms of discourse that *could* play the value-role, whether or not they in fact do. (Compare: different theological views, realist and irrealist, can play the God-role, although many of these consciously depart from the everyday meaning of ‘God’.<sup>18</sup>) Not every kind of discourse could play the value-role – it’s doubtful that concepts referring to the starry skies could do the job, just as many doubt that using ‘God’ to refer to love amounts to a genuine conception of God.

It is easy enough to see how we could work our way to a realist conception of value even if our common starting point is antirealist just as we could, I believe, similarly work our way to, say, concepts of libertarian

---

<sup>17</sup> This point is also echoed by Mackie’s claims about the independence of second-order and first-order morality.

<sup>18</sup> See J. Bishop, ‘Can There Be Alternative Concepts of God?’, *Noûs*. 32 (1998), pp. 174–188 for discussion of the God-role. Bishop is responding to a parallel worry in the philosophy of religion: he claims that “the [traditional] concept of God as omniGod can have competitors” and he wants to clarify “what these competing concepts may be understood as competing for” (p. 177), if they aren’t competing accounts of the meaning of ‘God’.

freedom or qualia that transcend everyday notions of freedom of consciousness.<sup>19</sup> Once we possess such concepts, we can ask, for example, 'Does anything possess such libertarian freedom? Do we?' If it turns out that the answer to that latter question is 'Yes', and (let us assume) that what governs our choices does not conform to our common compatibilist conception, it would be absurd to conclude that we have discovered that we are *not* really free!<sup>20</sup>

This is how, even if our evaluative discourse is in fact response-dependent, it can still make sense for us to ask what it would be like for concepts referring to response-dependent facts to play the value-role – it can still make sense for us to consider the possibility that objective values exist, even though the word 'value' in that phrase partly transcends its everyday meaning.<sup>21</sup>

Antirealists could respond by trying to argue that non-natural objective facts *couldn't* play the value-role.<sup>22</sup> This, however, could *not* be the claim that, even if RM is true, the non-natural facts it refers to are unable to guide action. As I said, RM could not be true yet turn out to refer to the starry skies, or, say, to irreducible non-natural facts about prime numbers. For RM to be true *just is* for the facts it refers to be capable of playing the value-role. In any case, since my aim here is not to defend RM, I do not need to resist this possible line of argument against it. What is important for my purposes is that such an argument can proceed *without* having to first answer questions about RS. And, of course, robust realists could level similar complaints against antirealist alternatives.<sup>23</sup>

---

<sup>19</sup> It's common for theoretical concepts to transcend the content of everyday terms. Scientific conceptions of space and time radically depart from folk concepts, but this doesn't render them unintelligible, and it's clearly appropriate to use 'space' and 'time' in both contexts despite this divergence in meaning. Conceptual revision doesn't always imply a change in subject matter.

<sup>20</sup> The philosophers' GOD also notoriously transcends the everyday notion, but it would be similarly absurd to conclude that most theist philosophers and theologians are really atheists.

<sup>21</sup> *Objection:* But if in stating RM we are employing a non-standard notion of value, then it isn't really true that the metaphysical claim of robust realists is true, as the reverse theory requires, but some *other* metaphysical claim. *Reply:* I understand RM to be *exactly* the metaphysical claim typically asserted by realists. This claim refers to facts that, if they exist, exist whether or not RS holds. I have been trying to explain how realists could still refer to these facts even if RS happens to be false.

<sup>22</sup> This is a plausible reading of some actual non-cognitivist complaints. I am grateful here to an anonymous referee.

<sup>23</sup> See Enoch, 'An outline of an argument for robust metanormative realism'. What the above discussion does bring out, however, is that we should distinguish claims that some

*Objection: Discourse Dependence?*

I have argued that we can make sense of the possibility that RM might be true even if RS isn't, and that this shows that the metaphysical claim of robust realism doesn't require the semantic claims commonly associated with it. But to make sense of this possibility, I had to argue that objective value must be capable of filling the value-role. And this, it might be objected, really amounts to saying that objective value has to satisfy *our concept* of the value-role, a concept that, presumably, reflects core truisms about value. This can seem self-defeating. For doesn't this mean that objective value *does* ultimately depend on our discourse, and that realists *are* committed to making semantic claims?<sup>24</sup>

Let me point out first that this is not an objection to the reverse theory, or to my main claims. The reverse theory claims that objective value could exist even if our evaluative discourse was antirealist, and that therefore realism does not require standard semantic claims such as cognitivism. All that this objection could show is that objective value is dependent on *other* aspects of our evaluative discourse, and that realism needs to make *different* semantic claims than commonly assumed.

We can now turn to the objection. It involves two distinct complaints: one about discourse dependence, the other about the semantic commitments of realism. If correct, the first complaint would be troubling. After all, it was robust realism's commitment to Discourse Independence that motivated reverse theory. If this complaint is correct, then this is a commitment that cannot be fully met.

I have tried to explain how we can refer to objective value, even if our discourse is antirealist. We can refer to it using the concept of objective value, and therefore, trivially, objective value satisfies this concept. This, of course, hardly shows that objective value is in any way metaphysically dependent on our discourse; stones and protons also satisfy our corresponding concepts.

Since the concept of objective value is partly constituted by reference to the value-role, objective value also satisfies minimal criteria for occupying that role. Again, this couldn't be ground for complaint; it could hardly be objected that the reverse theory doesn't leave it open that objective value might turn out to be the starry skies, or prime numbers.

---

discourse *can't* play the value-role, from claims that it doesn't accurately describe what *actually* plays the value-role. The *latter* kind of antirealist argument would be an argument *only* against RS.

<sup>24</sup> An anonymous reviewer pressed this objection.

That certain facts satisfy certain concepts would suggest discourse dependence only if the concepts are in some way *metaphysically prior* to the facts. But nothing I have said thus far suggests such priority. Perhaps the worry is that objective value must satisfy *contingent* features of our *actual* evaluative discourse, even though this discourse happens to be ultimately antirealist. Realists must thus either concede that our discourse does have metaphysical priority, or they must explain this fortuitous fit.

There is, however, no fortuitous fit, and objective value doesn't need to satisfy any contingent criteria set by our actual discourse. It unsurprisingly has to satisfy *necessary* conditions that *all* possible evaluative discourses must meet, and which, naturally, our actual evaluative discourse also meets. It is contingent that we exist, and perhaps it is even contingent that we possess an evaluative discourse (and have a value-role), but it is of course not contingent that our evaluative discourse is an *evaluative discourse*.<sup>25</sup>

It is hardly mysterious that, unlike facts about the starry skies or prime numbers, there is this fit between objective value and what I called the value-role. If objective value exists, then there are irreducibly normative facts – facts about objective reasons for attitude and action. It is in the very nature of such facts that they relate in the relevant way to the value-role. This would be true of such facts even if, in fact, no agents actually exist, and thus no value-role these facts could occupy.

We can reject the first complaint. But the second complaint is more successful. The above remarks do mean that the answer to my titular question cannot be an unqualified 'No'. There is a minimal sense in which realism does need to make some semantic claims: it needs to accept certain core truisms about value. These, however, are truisms that *all* plausible metaethical views must accept. What realism does not need to do is to put forward any *distinctive* semantic claim. And, to repeat, realism isn't committed to RS, the semantic claim that it is assumed to require.

Realism is committed only to core truisms that should be accepted by all plausible metaethical views. I concede, however, that metaethicists might disagree even about these truisms – that they might disagree about the nature of the value-role, and what could fill it. As I have pointed out above, antirealists could try to argue that the facts postulated by realists cannot fill the value-role, and that realists might return the favour. Nothing I said rules

---

<sup>25</sup> If the falsity of RS implied the falsity of RM, then this would suggest that RM isn't metaphysically independent of RS. But what further possibility is being asserted by the current objection? That objective value won't exist if we don't? That it won't exist if we exist but don't possess any kind of evaluative discourse? Robust realists can easily deny both claims.

out such arguments. Such arguments might in part revolve around questions about the core of our concept of value, leaving an important, if residual, role for semantics in the debate about realism.

*Objection: A Suffocating Multitude of Ersatz 'Values'?*

I have argued that objective values might exist even if our discourse is anti-realist, and that we could refer to these objective values by drawing on the evaluative concepts already presupposed by standard realist claims. The facts to which these concepts refer, I argued, count as genuine *values* because they can play the value-role, a role reflected in the metaethically neutral core of our evaluative practice.

It might be objected, however, that this achieves too much and thus very little. After all, there is a *multitude* of metaethical views which describe properties and facts which could play the value-role. There is thus a vast multitude of possible 'values' out there. These facts could not all matter, or make a claim on our practical deliberation. So what's so special about RM?<sup>26</sup>

Antirealist views often ground value in natural facts about what, for example, we would desire in some ideal conditions, or what an ideal observer would want us to do, or the conventions of our culture. Such natural facts would be there even if RS and RM are true. This is not surprising, given that antirealists appeal to these facts precisely *because* they are metaphysically uncontroversial – because their existence is not under dispute in metaethical debate. No one, however, thinks that, because there are these natural facts, then these antirealist views are vindicated at the metaphysical level, even if our discourse *is* in fact realist, let alone if objective value does exist. It would be even more absurd to argue that, even if objective value exists, we should abandon our realist discourse and begin to guide our actions in light of, say, what an ideal observer would want us to do. This would be as absurd as suggesting to 'realist' theists who believe that the traditional God exists to revise their religious discourse so that, when worshipping God, they would instead merely express their affective allegiance to certain rituals and practices.

These points seem to challenge the reverse theory, but they actually support it. They draw our attention to what *is* special about objective values. Of course one unhelpful way in which objective values are special is that

---

<sup>26</sup> I am grateful to an anonymous reviewer for pressing this question.

they are far from being metaphysically uncontroversial. Even if RS is false, it certainly *would* be surprising to discover that RM is true – discovering this might, for example, be in tension with metaphysical naturalism.

But this isn't the only way in which RM is special. If objective values *do* exist, then intuitively this *does* matter. Similarly, although it would be crazy for theists who believe that the traditional God exists to go antirealist, a move in the reverse direction seems far from absurd. If the robustly real God revealed Himself to an antirealist theist, could the latter reply: 'It's nice to know you exist, but, if you don't mind, I was in the middle of my prayer'?

The natural facts to which antirealists' views appeal would exist whether or not these views truly describe our evaluative discourse. These facts, then, are discourse independent – *qua natural* facts. Can we also say that they are discourse independent *qua values*? If our evaluative discourse did refer to the desires of an ideal observer, then facts about these desires would play the value-role. In this sense, these facts *can* play the value-role. But, by definition, these natural facts *aren't* irreducibly normative. They get their normative significance only when plugged into some evaluative practice. They relate to the value-role *only* when conjoined with an appropriate evaluative discourse. Otherwise, they are just inert natural facts. This is why their existence would be irrelevant if our evaluative discourse points in another direction.

This could not be true of the irreducibly normative facts postulated by robust realists. Unlike these natural facts, we cannot even conceive of irreducibly normative facts as anything other than normative. To deny their normative significance is to deny their very existence; there is nothing else for them to be. They are, to borrow from Kant, uniquely 'normative in themselves', regardless of what kind of discourse we have.

To hold that such irreducibly normative facts exist is to hold, not just that facts exist that can in principle occupy the value-role, but precisely that facts exist that can occupy the value-role *independently* of what kind of discourse we have. This just is what makes these facts robustly objective. (We can say that objective values are not just potential *occupiers* of the value-role, but that, if they exist, they *already compete for it*.)

It might be complained that I am answering the objection by simply asserting that irreducibly normative facts have this superlative character. But I am not simply asserting this. I am simply drawing attention to what would follow from the existence of irreducibly normative facts. These implications may seem surprising, but this is only because robust realists have not paid enough attention to the relation between RS and RM.



## 5. Would it Matter if the Reverse Theory was True?

Richard Hare famously thought that the metaethical debate about realism is empty. He wrote:

Think of one world into whose fabric values are objectively built; and think of another in which those values have been annihilated. And remember that in both worlds the people in them go on being concerned about the same things – there is no difference in the ‘subjective’ concern which people have for things, only in their ‘objective’ value. Now I ask, “What is the difference between the states of affairs in these two worlds?”<sup>27</sup>

Hare thought that we should answer: ‘None whatever’. But Mackie’s error theory offers a better answer. As Mackie pointed out, if our discourse is realist, then if we inhabit a world where such objective properties are absent, then this *does* make a difference: it would mean that all of our (positive) evaluative beliefs are untrue, and that our evaluative concepts fail to refer to anything.<sup>28</sup> Still, in a sense, Mackie agrees with Hare that this discovery wouldn’t (or shouldn’t) really make a difference to our subjective concerns, and to moral practice. Mackie thought that we would have reasons to go on just as before, even if our evaluative beliefs are in error. But this isn’t a mandatory part of the error theory. Some moral error theorists dissent: they think that we should abandon moral practice, if we discover that it refers to merely mythical facts.<sup>29</sup>

But once the semantic/metaphysical distinction is in place, we can also read Hare’s question in the *other* direction. What if, as Hare implies, our evaluative discourse and subjective concerns really were completely independent of, and made no reference to, objective properties. What difference would it make if the reverse theory obtained and such properties nevertheless *did* exist? Should we answer, with Hare, ‘None whatever’?

If Hare is right, then if RS is false little would be achieved by showing that there are objective values. For it might not *matter* that (and whether) there are such values. If so, then RS matters not because it’s *required* for RM to be true, but because RM *without* RS won’t *matter*.

We are asking what difference the discovery of objective values should make to our subjective concerns and evaluative practice, if our evaluative

---

<sup>27</sup> R.M. Hare, ‘Nothing matters’, in his *Applications of Moral Philosophy* (London: Macmillan, 1972).

<sup>28</sup> Mackie, *Ethics: Inventing Right and Wrong*, pp. 20–25.

<sup>29</sup> The independence of the value-role from metaphysical debates about realism thus doesn’t vindicate Hare’s view that these debates are empty or don’t matter.

discourse is antirealist. Should we value the things that possess such objective value? Should we respond to the objective reasons provided by these values, desiring and acting in their light?

Antirealists sometimes complain that robust realists are unable to explain why the objective facts that they describe should have any normative hold on us. These antirealists insist that we can always ask *why* we should guide our actions in light of these objective facts, a normative question which, they allege, robust realists cannot answer.<sup>30</sup> Robust realists reject this challenge. They point out that, on their view, when we ask normative questions – when we ask what we have reason to do or care about – we just are asking what objective reasons we have. Once we are told what these are, there is no remaining gap that needs to be bridged. This seems to me right.

But this normative question does seem to have force in the case of the reverse theory. For why *should* we care about these objective facts? Suppose that our evaluative discourse is in fact response-dependent. Thus, when we ask normative questions about what we should do, we are asking for *response-dependent* reasons, for reasons based in our desires.<sup>31</sup> It seems to follow from this that if we *don't* happen to care about these objective facts, if we don't desire to follow them, then we have *no* reason to care about them – they would have no normative hold on us. Indeed, even if we somehow did happen to desire to follow these objective reasons, their normative force would still ultimately derive from our subjective attitudes.

Now if the reverse theory is correct, then there exist certain response-independent facts on which our discourse is silent. The present objection is that discovering the existence of these queer facts would be no more than a surprising discovery about the universe that needn't make any practical difference. If this is correct, then it *would* be enough for antirealists to show that the realist's semantic claim is false for them to declare victory. The further question whether objective values might nevertheless exist would be of merely theoretical interest – like questions about the existence of the Higgs boson particle.

---

<sup>30</sup> C. Korsgaard, *The Sources of Normativity* (Cambridge: Cambridge University Press, 1996), p. 240.

<sup>31</sup> I shall be using 'response-dependent reasons' to refer to what others call *internal* or *desire-based* reasons. It is irrelevant here that some antirealist views might be compatible with externalist views of reasons.

But could this be right? Suppose again that our discourse was realist but that, because we mistakenly concluded that the error theory is correct, we shifted to a response-dependent semantics. On the current suggestion, objective reasons would immediately lose all of their normative force, and objective values would suddenly cease to matter. Worse, even if we later discovered the mistake, this would give us *no* reason to return to objectivity. This seems absurd. If morality imposes objective demands, these won't just vanish the day a community of psychopaths rewrites the dictionary.

To accept RM is to accept the existence of objective values – to accept the existence, not just of certain non-natural facts, such as, say, sets or prime numbers, but of non-natural fact that have intrinsic response-independent significance. *If* they exist, then there *are* objective reasons for us to do this or that, whatever our subjective concerns. This *just is* what it is for objective reasons to exist. There is nothing else, queer or not, for these facts to be. To deny that we have these reasons is to simply deny RM.

So we can again reject the antirealist's normative question. It's incoherent to both assert that objective values exist and to deny them any normative force: to believe that these facts exist just is to believe that we should act in certain ways (where 'should' is already understood objectively), and there is no space left for asking *why* we should.

Antirealists are likely to respond that the very idea of values and reasons that transcend our attitudes and discourse is no more than a florid myth. Robust realism's commitment to radical discourse independence can make it seem even more metaphysically extravagant and thus, to many, even less plausible. And notice that this commitment would bear on realism's plausibility *even if RS is true*. But, again, to respond in this way is no longer to ask whether, *if* such objective values exist, they *would* matter. It is rather to ask whether objective facts that matter in this superlative way *could* exist. This is a different and familiar metaphysical doubt, a doubt that can also be raised about standard forms of robust realism. Whether these doubts are justified is not a question I need to answer here. What I have been trying to show is that antirealists *need* to press these doubts about RM even if RS is false.

## 6. Revolutionary Realism?

I said that the natural facts to which antirealists appeal have no inherent interest if RS and RM are true. These facts matter only if our discourse is antirealist and they actually occupy the value-role. Things might be

different, however, if the error theory is true – if it turns out that only RS is true. Here the value-role would be vacant, and we might need something to occupy it. *Revolutionary antirealists* propose that in response to this predicament, we should adopt some form of antirealist discourse, and bring those inert natural facts to life.

The reverse theory, however, presents us with a further possibility. What happens if RS is false but RM is true – if objective values exist but something else *already occupies* the value-role?

I earlier asked whether objective value would matter if our discourse was antirealist. We can now see that this question has not yet been fully answered. I argued that if objective value exists, then many things matter objectively, and give us objective reasons. What is still not clear, however, is whether these things matter *period*. For if our evaluative discourse is antirealist, then we also possess a standpoint from which many things matter *subjectively*, and give us *response-dependent* reasons. It would be an incredible coincidence if these two sets of values and reasons always aligned. Suppose, for example, that we have objective reasons to shout, and response-dependent reasons to shut up. Should we shout or shut up? It is pointless to reply: objectively, we should shout; subjectively, we should shut up. We cannot do both. So what are we to do? It is not as if we are facing opposing reasons of different *strengths*, which we can simply weigh against each other. Nor would it help to ask *which* set of reasons we should follow, or which evaluative concepts we should adopt, or what should occupy the value-role. Since these questions can be given both objective and response-dependent readings, our answers to them cannot solve our predicament.<sup>32</sup> We cannot adjudicate between these opposing sources of normativity by appealing to reasons arising from one of them.<sup>33</sup>

---

<sup>32</sup> Kalderon writes: “[C]onsider the hypothetical community whose acceptance of moral sentences is governed by noncognitive norms despite the existence of moral facts. By the norms internal to their moral practice, a competent speaker may be justified in accepting a moral sentence ... Nevertheless, the fact that moral acceptance is fixed independently of the moral facts is just grounds for criticism. Noncognitivism might be a correct description of the standards of acceptance internal to moral practice, but it is a further question whether moral acceptance ... is legitimate or justified.” (*Moral Fictionalism*, pp. 145–6). The problem is that normative terms such as ‘just grounds’, ‘legitimate’ and ‘justified’ can themselves be understood not only by reference to objective value but also in terms of standards internal to our (supposedly antirealist) moral practice.

<sup>33</sup> This predicament might not arise for antirealist views that don’t recognize reasons for ends, since, on these views, the reasons for ends provided by objective values will face no opposition.

Antirealists might deny that there is any such predicament. From within our actual evaluative discourse, certain acts appear as what we have *conclusive* reason to do. This is still how they would appear, of course, if our discourse turns out to be antirealist. But this, again, seems to deprive the idea that there might nevertheless be competing *objective* reasons to do otherwise of its very sense. Perhaps we can think of these reasons in the way we think of the elaborate norms of some foreign culture. These objective facts may describe possible standards for attitude and action – even *universal* standards for attitude and action – but to us these standards must be devoid of any normative force.

My response to this suggestion is not likely to surprise: it makes no sense to both assert that objective values exist and deny them normative import, or to allow them import only relative to some discourse. Yet the objection is correct in pointing out that the value-role is *exclusive*: we cannot claim our allegiance to competing sources of normativity. This might mean that, if RM is true but RS is false, we face an irresolvable predicament. Such a practical dualism would be unfortunate, but it is not clearly incoherent.<sup>34</sup>

Consider next that the very same claims about conclusive reasons could also be said in the opposite direction. For suppose that RS and RM *are* true, but we encounter people whose evaluative discourse is response-dependent. They could similarly say that they (indeed that also we) have conclusive reason to, for example, permit torture, and that our protestations to the contrary represent inert standards of conduct which have no normative import.

This example shows that the key questions raised by the reverse theory can be asked *even* if RS is true, and do not arise only when we try to transcend our actual evaluative discourse. Even if RS is true, we can still ask, about a foreign antirealist culture (or about our counterfactual or future selves): ‘Would it be possible for them to refer to objective value? Would they be able to know that it exists? Would it matter, and have normative force for them?’

Our disagreement with those imagined antirealists would not be disagreement in belief. Indeed, the beliefs of each side might be impeccable relative to its own evaluative discourse. Yet we are not simply talking past each other. We disagree, you might say, about what to do. At another level, we could be said to also disagree about what to place in the value-role. But

---

<sup>34</sup> See H. Sidgwick, H. *The Methods of Ethics*, first edition (London: Macmillan, 1874), p. 473 for a familiar precedent.

again, it is useless for each side to just assert what we *should* (objectively or subjectively) do, or what *should* (objectively or subjectively) occupy that role. Both sides can make these assertions, but they seem to lack any real critical bite.

Such symmetry is intolerable to robust realism. It introduces a peculiar kind of second-order relativism, and clearly violates discourse independence. Realists need a metaethically neutral way of showing that these antirealists shouldn't permit torture, *period*. They need to show that, if objective value exists, then antirealist forms of evaluative discourse are in some way defective, or inferior. Can such *revolutionary realism* be defended?

Realists are accustomed to appealing to intuition without embarrassment. They should not be embarrassed to appeal to the intuition that these imagined antirealists *are* in the dark about something crucial, just like, say, verificationist conceptions of the past, or behaviourist conceptions of mental states (the deepest problem with such conceptions *isn't* that they are mistaken about our actual concepts). And consider again the example of a community that adopted an antirealist understanding of religious discourse on which talk about God is merely expressive of certain deep attitudes and commitments, but where it turns out that God in the traditional realist sense *does* in fact exist. Is it really in question which discourse should give way?<sup>35</sup>

To me at least the answer seems clear. In a way that is hard to explain or defend by argument, the robustly real has an intuitive primacy.<sup>36</sup> It is not by accident that antirealist views are typically fallback positions, something we learn to grudgingly accept when we realize that nothing better is forthcoming. As I noted above, the fact that no one recommends that we should adopt an antirealist conception of value if RM is in fact true, far from being a problem for the reverse theory, is rather evidence in its favour.

It would be nice to be able to say more. Perhaps the very idea of a value-role imposes a functional standard of betterness that is metaethically neutral, and certain kinds of discourse, certain kinds of facts, can play that role better than others – Perhaps the irreducibility and discourse independence of objective values makes them superior, in this respect, to their antirealist competitors in a way that even these competitors must acknowledge. Or perhaps we can adjudicate between these opposing sources of

---

<sup>35</sup> It seems absurd to hold that such antirealist theists could dismiss the robustly real God, yet true atheists must succumb!

<sup>36</sup> For a related discussion, see also my 'The value question in metaphysics', forthcoming in *Philosophy and Phenomenological Research*.

normativity by appealing to *theoretical* reasons. We strive to represent the world as it really is in itself, as free as possible from the contingencies of our subjective standpoint. Robustly objective values, if they exist, are part of the very fabric of the universe (though not of course of the *natural* world). This is not true of their antirealist competitors, which are merely built up from more fundamental natural facts and properties, and which get their normativity only through their relation to our discourse and psychological states.<sup>37</sup>

The strongest argument for revolutionary realism is, however, rather simple. I said that if the reverse theory actually holds, then we face a practical predicament. We have to act in some way, but objective and response-dependent reasons point in opposing directions. What are we to do? Objective reasons are discourse independent. They would exert their normative force whatever discourse we happen to have. Response-dependent reasons matter *only* if we have an antirealist evaluative discourse. Objective reasons are inescapable, response-dependent reasons merely optional. So the solution to our practical predicament is obvious. Objective reasons win.<sup>38</sup>

## 7. If the Reverse Theory Actually Held, How Could We Know This?

Mackie's insight was that even if RS can be shown to be true, the realist is at best only half way home. The realist still needs to show that RM is also true to avoid the error theory. It's nevertheless widely assumed that to establish antirealism, it's sufficient to show that RS is false. This just follows from taking realism to encompass both RM and RS. What I have been urging, however, is that if the reverse theory is coherent – if RM could be true even if RS is false – then this raises a reverse challenge for *antirealism*. The idea is that even if antirealists can show that our discourse is antirealist, they still need to rule out the possibility that objective values nevertheless exist.

Antirealists might concede that the falsity of RS leaves RM as a coherent logical possibility. But they can complain that there is no need to rule out the existence of facts whose mere possibility came into view only because

---

<sup>37</sup> Compare: there is no flaw in not possessing a naturalist or expressivist religious discourse. But it's an understatement that there is something important that we would be missing if God existed in the robust sense, but we failed to believe that. And to believe that is *already* to accept a certain *exclusive* practical authority.

<sup>38</sup> Ironically, this form of argument mirrors an influential type of *antirealist* argument. See Korsgaard, *The Sources of Normativity*.

of a realist misinterpretation of our evaluative discourse. This is why although some antirealists offer metaphysical arguments for thinking that objective values don't (or even couldn't) exist, others see this as unnecessary. The possibility that such values exist is, for them, no different than the possibility that fairies exist at the bottom of the garden.<sup>39</sup>

These antirealists couldn't be claiming that:

If our discourse is antirealist, then RM is false, or very unlikely to be true.

Since the antirealist's semantic claim is metaphysically silent, its truth should also say nothing about the truth, or even *probability* that such facts exist. So it seems that the mere falsity of RS couldn't make RM *less* plausible.

Still, that the reverse theory is coherent doesn't yet show that we have, or even could have, grounds for thinking it might be true. After all, robust realists typically defend their view by asking us to take our *actual* evaluative discourse, beliefs and phenomenology at face value. Their implicit argument for the existence of objective value typically takes the following form:

*The Face Value Argument*

- (A) Our evaluative discourse is cognitive and objective [RS].
  - (B) We have at least some evaluative beliefs that survive critical reflection.
  - (C) There are no positive grounds for thinking that RM is false or incoherent.
- Therefore:
- (D) Objective values exist [RM].<sup>40</sup>

In other words, these robust realists don't offer *positive* metaphysical arguments to show that RM is true, they think that it's sufficient to successfully *resist* negative arguments that aim to show it to be false or incoherent, so long as RS is true and at least some evaluative propositions appear to be true.<sup>41</sup> (Contrast the positive arguments for the existence of moral

<sup>39</sup> See D. Gauthier, *Morals by Agreement* (Oxford: Oxford University Press, 1986), ch. 2.

<sup>40</sup> Robust realists typically appeal to something like this argument when they try to address *epistemic* worries about realism. But presumably if we are justified in believing in at least one objective evaluative proposition, then we are justified in believing that objective values exist.

<sup>41</sup> Korsgaard, *The Sources of Normativity*, p. 31, accuses realists of offering nothing besides criticism of alternatives, a point that Nagel and Parfit concede (Nagel, *The View From Nowhere*, pp. 143–4; Parfit, 'Normativity'), and which is implicit in Shafer-Landau, *Moral Realism: A Defence*. One exception is Enoch, 'An outline of an argument for robust



properties sometimes offered by *naturalist* evaluative realists when they try to show that such properties play an indispensable role in common causal explanations, or the *a priori* and evidential arguments for God's existence offered by theists.<sup>42</sup>)

Some will doubt whether the premises of this argument are sufficient to establish its conclusion. And these premises are themselves controversial. Although few doubt (B), metaethicists are by no means close to reaching agreement about the truth of (A). And to defend (C), realists need to resist the arguments against RM put forward by Mackie, Blackburn and others.<sup>43</sup>

It's not my aim here to assess the merits of the Face Value Argument. What is clear, however, is that if this is all that robust realists can offer as a defence of RM, then the reverse theory is in trouble. Although the fact that our discourse is antirealist gives us no reason to think objective values *don't* exist, it does undermine common grounds for thinking they *might* exist. More precisely:

If our discourse is antirealist, then, given that the main positive argument for RM presupposes RS, there is no (and couldn't be) reason to think that RM is true.

And if there is no way to defend the reverse theory, then this might also justify refusing to see it as a serious metaethical possibility.

Now all that this would establish is the epistemic claim that although RM *might* be true, if RS is false we would have no means of finding out whether this is the case. Still, this is an outcome that should make robust realists uncomfortable even if they are confident that RS is in fact true. After all, the truth of RS would only be a contingent fact. So it seems that it's only by a stroke of luck that we know that there are objective values; if we had taken the wrong semantic turn, we would be in complete darkness, hopelessly out of touch with what really matters. Worse, if RM is not epistemically independent of RS, one might begin to worry whether it is really metaphysically independent. Perhaps there is a necessary epistemic tie

---

metanormative realism', which sketches a positive argument for robust realism by arguing for the deliberative indispensability of non-natural objective norms.

<sup>42</sup> Notice that common realist arguments that appeal to our parallel normative commitments in the *epistemic* realm are *not* positive arguments for believing in the existence of practical norms (for accepting D/RM). They are just ways of resisting arguments against their coherence – they are arguments for (C), and in themselves they only establish that such norms *could* exist, not that they do.

<sup>43</sup> See, e.g. Mackie, *Ethics: Inventing Right and Wrong*; S. Blackburn, *Essays on Quasi-Realism* (Oxford: Oxford University Press, 1994). As noted earlier, another way of arguing against RM is to show that objective facts can't occupy the value-role.

between RS and RM because value is really *constituted* by our evaluative discourse and practices – which wouldn't be an especially robust form of realism, if it even counts as realism.<sup>44</sup>

The reverse theorist needs to find a way of defending RM even if RS is false, and it would be good news for robust realists if this could be done. But it seems that such a defence would require offering *positive* metaphysical grounds for thinking that objective value exists, and it's hard to see what these grounds might be.

Consider the scenario described earlier on which our evaluative discourse was initially realist, and we had adopted an alternative response-dependent discourse because we became convinced that the error theory is true. If this was the case, then we might later have reason to change our minds. We might later discover that the arguments for the error theory were mistaken. But here we would only be finding grounds to doubt what seemed to be positive grounds for thinking that objective values *don't* exist (so we can reaffirm (C) and appeal again to the Face Value Argument). These are still not positive grounds for thinking that they *do* exist, and won't help us in the pure case, where our discourse is antirealist from the start.

And in that pure case, even if we could reject the positive metaphysical arguments against the existence of objective values (the arguments from queerness, supervenience, etc.), that would merely mean that we don't have reasons to think such facts *don't* exist, not that we have reason to think they do exist. The failure of such arguments is merely a necessary condition for the truth of RM.

We cannot infer (D) from (C) alone. But we must find a way of getting to (D) that doesn't presuppose (A).

I think that there is a way. Robust realists often claim that we come to know about objective evaluative and normative facts simply by exercising our general capacity for rational reflection.<sup>45</sup> If, as seems plausible, we would have this capacity even if our evaluative discourse was antirealist, then realists might (perhaps must) claim that we should still be able to find out that objective values exist (and what they are) by exercising that capacity. So the reverse theorist can argue that:

---

<sup>44</sup> I'm *not* suggesting that we could have epistemic access to objective value without possessing the relevant *concepts*. What's at question is whether we could have such access if our *discourse* doesn't include these concepts. In other domains epistemic progress often involves acquiring new concepts, and we saw earlier that the relevant concepts are already available.

<sup>45</sup> See, e.g. Parfit, 'Normativity'; Huemer, *Ethical Intuitionism*.

If we are able to form some beliefs about objective value that survive critical reflection, then, so long as we have not been given positive grounds to think that objective evaluative facts don't exist, then we are justified in taking these beliefs at face value as referring to a domain of genuine facts.

In other words, *if* something like the Face Value Argument is valid, then to support (D) it should be enough to defend (B) and (C), so long as the relevant evaluative beliefs are realist in content.

As an illustration, consider the kind of intuitions that robust realists often appeal to when they try to argue against response-dependent views. Take for example Parfit's discussion of Future Tuesday Indifference. For someone to know what pain is like but not to care about feeling agony if it falls on a Tuesday seems deeply irrational in a way that seems incompatible with response-dependent accounts of value and desire-based accounts of practical reason.<sup>46</sup>

What are realists doing when they appeal to this and similar examples? One thing that they are doing is eliciting *substantive* intuitions about what we have reason to do or care about in certain situations. They are identifying strong intuitions that lead to firm evaluative or normative convictions – examples that support (B) in the Face Value Argument. But in doing so, they are also eliciting *semantic* intuitions – intuitions suggesting that our concept of a reason is objective. In fact, Parfit's example could still retain its *semantic* force even if we thought that such indifference *is* rational, so long as we agree that Parfit's contrary intuition expresses an *intelligible* normative view that is compatible with our concept of a reason. This would be sufficient for the example to support (A/RS).<sup>47</sup>

What happens to such examples if it turns out that our evaluative discourse is really response-dependent? Must we conclude that these objectivist intuitions are merely illusory? If such intuitions cannot be accommodated by response-dependent views, then these objectivist intuitions are clearly *not* employing our actual response-dependent concept of a reason. And this means that if they are true, what makes them true are *not* the kinds of psychological facts that (we are now assuming) our response-dependent evaluative discourse in fact refers to. So realists should

---

<sup>46</sup> D. Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1986), pp. 125–126.

<sup>47</sup> This semantic intuition could be endorsed even by an error theorist – someone who accepts (A) but rejects (C) and therefore (D). Strictly speaking, the robust realist probably only needs such examples to support (A), given that once it's shown that our evaluative discourse is realist, there are plenty of other firm beliefs that could support (B).

now construe these intuitions as *purely substantive* – and as employing a conception of value that transcends our discourse. They should claim that, through reflection, they have arrived at true evaluative conclusions that go beyond what is recognized by our current conceptual scheme. But that is just to say that objective values do exist – and that the reverse theory is true. (One consequence of this possibility is that it might be hazardous to use intuitions about such far-fetched cases as *semantic evidence*.)

There is one way in which this form of argument might be weaker than the standard Face Value Argument. This isn't because the truth of RS in itself makes much of a difference here. Rather, *if* RS is true, then the body of evaluative belief which the argument tells us we should take at face value encompasses the *whole* of ordinary evaluative beliefs – the numerous everyday evaluative beliefs of numerous people today and in the past, at least some of which have survived sustained critical scrutiny. It's plausible that this vast body of belief has greater epistemic force than the objectivist intuitions of a few philosophers. So the truth of RS could, in this indirect way, increase the overall plausibility of RM. Still, the realist intuitions would still win the day if they survive further reflection and turn out to be widely shared. So the main point stands: realism's prospects could remain very much open, even if RS was conclusively refuted.

Needless to say, those who reject the Face Value Argument can also reject this parallel route. But my aim here is only to show that if robust realism deserves serious metaethical attention, then so does the reverse theory. And if the Face Value Argument can support robust realism, then there is also a way to defend the reverse theory.

## 8. Conclusion: Robust Realism as a Purely Metaphysical Claim

During the heyday of the linguistic turn, metaethicists used to assume that the debate about metaethical realism is a debate about the nature of our evaluative discourse. Realism can be vindicated, they thought, if we can show that evaluative discourse is cognitive. Mackie has shown why this isn't sufficient. In response, realism is now often formulated as involving a conjunction of claims: a claim about the nature of our concepts, and a further claim about the nature of reality. This assumption shapes the way the debate in metaethics is typically set out as starting with the question of cognitivism, followed by the question of whether, if evaluative discourse is cognitivist, there really are evaluative facts, and finally whether, if there are,

these facts are natural or non-natural.<sup>48</sup> It's therefore assumed that if some antirealist view can succeed in establishing its semantic claim, the dispute is over and realists lose.

But once we fully absorb Mackie's insight, we should see that this is mistaken.

If objective values exist, then realism *is* true. It's not as if the truth of RM gives us only *half* of realism. This metaphysical claim isn't only necessary for the truth of robust realism. It's sufficient. Realism doesn't require the truth of cognitivism, or of any similar semantic claim – not if 'cognitivism' is understood in its standard meaning, as a claim about our discourse. Realism, at least in its purest sense, should be understood as the claim that:

**RM.** There are objective evaluative facts,

or, if you wish, truths, about value and practical reasons. But not that our *own* discourse is *truth-apt*, or refers to these truths.<sup>49</sup> Of course realists also typically assume that our actual evaluative beliefs and discourse are realist in character. But that is a further claim,<sup>50</sup> just as the common realist assumption that we possess evaluative knowledge is a further claim. After all, the truth of RS and RM is perfectly compatible with evaluative scepticism.

The upshot of this paper is that the debate between realists and antirealists can proceed in pretty much the same way *whether or not* RS is true. But this needn't be a last resort for realists, something to hang on to if they lose the semantic debate. What realists *really* want to defend is RM. Now we did see that it would be easier to defend RM if RS is also true. Still, if realists can defend RM by appealing to substantive realist intuitions, then they can simply *sidestep* the semantic debate. This is an important result, given how little substantive and methodological agreement there is on the semantics

---

<sup>48</sup> See, e.g., A. Miller, *An Introduction to Contemporary Metaethics* (Cambridge: Polity Press, 2003), p. 8.

<sup>49</sup> Discussing realism about the physical world, Devitt has similarly argued that "[r]ealism says *nothing semantic at all* beyond ... making the negative point that our semantic capacities do *not* constitute the world" (M. Devitt, *Realism and Truth* (Princeton: Princeton University Press, 2nd Edition, 1991), p. 39). Devitt however is arguing against a purely semantic interpretation of debates about realism.

<sup>50</sup> Actually, if robust realists want to defend RS (let alone if they *must*), then, since RS is contingent, they face a challenge. It's familiar that realists face an *epistemic* challenge: they need to explain how we could ever know about non-natural facts with which we have no causal contact. But there is an overlooked prior *conceptual* challenge: realists need to explain how, even if RM is true, we ever came to acquire an objective evaluative discourse. Is this merely a fortuitous accident?

issues. It would allow realists to avoid difficult questions in the philosophy of language, or to have to hold their breath as the empirical evidence from experimental philosophy comes in.<sup>51</sup>

All that the realist case requires are firm and persistent realist evaluative intuitions.<sup>52</sup> So long as such realist intuitions are on the table, the antirealists' work is not done even after they have established the falsity of RS. Indeed, when realism is understood as a purely metaphysical thesis, it's *not* incompatible with response-dependent and non-cognitivist views, if these are understood only as offering competing semantic accounts of our evaluative discourse. These views are in conflict with realism only when they also involve a further negative metaphysical claim, denying the existence of objective evaluative facts. Another upshot of this paper is that antirealists *need* to defend this negative metaphysical claim. Just as the truth of RS isn't sufficient for realism, its falsity isn't sufficient for antirealism.

Much past and current metaethical discussion has focused on questions about the nature of evaluative discourse, questions that have so far largely resisted resolution. Although I have no intention to dismiss these semantic questions and their interest, I've argued that they are less important than they might have seemed.<sup>53</sup> At least so long as robust realism remains a contender, the key question is metaphysical, not semantic.

In these ways, the possibility of the reverse theory helps clarify what is really at stake in the debate about metaethical realism. By releasing robust realism from its present commitment to RS, it gives realism a dialectical advantage. But some realists will not be cheered by this because,

---

<sup>51</sup> This is why my aim here wasn't to show that the reverse theory is *true*, but that the truth of RS isn't necessary, or even especially important, for a defence of robust realism. There have been ingenious attempts to defend different forms of antirealist semantics. I have nothing to add to that literature, but it's clear enough that we are not fast approaching a consensus on this semantic question. It's not yet clear whether, and how, empirical work would advance us towards such a consensus. Indeed, one weakness of current empirical work on folk metaethics is that it simply overlooks the distinction between semantic and substantive (and potentially discourse-transcendent) evaluative intuitions.

<sup>52</sup> Ironically, this is also the conclusion of R. Dworkin, 'Objectivity and truth: You'd better believe it', *Philosophy and Public Affairs* 25 (1996), pp. 87–139, which denies that we can even make sense of semantic and metaphysical disputes about claims such as RS and RM.

<sup>53</sup> Kalderon finds the semantic focus of current metaethics similarly puzzling, though he draws from this the rather different lesson that realism is best understood as an 'epistemological posture' (see Kalderon, *Moral Fictionalism*, pp. 94–95, 142). As noted earlier, semantics would still be needed to clarify the nature of the value-role and what could occupy it.

by exposing realism's radical discourse independence, the possibility of the reverse theory might also make realism seem less plausible. Indeed, if realism is committed to this possibility, this worry would have force even if RS *is* true. My defence of the intelligibility of the reverse theory is, in part, also an attempt to address this worry.<sup>54</sup>

---

<sup>54</sup> I am very grateful to several anonymous referees for extremely helpful comments. Work on this paper was partly supported by a grant from the Wellcome Trust [WT087208MF].