

Evolutionary Debunking Arguments, Explanatory Structure, and Anti-Realism

Karl Schafer

2/20/16

Any non-skeptical moral theory claims that our moral judgments are not systematically false. Thus, such theories take these judgments to be at least somehow reliable. But how is our reliability about moral questions best explained? One option is to explain the reliability of these judgments by making the moral facts dependent upon our moral reactions or judgments. For example, suppose that the moral facts depend upon our moral sentiments - and that our moral judgments are also sensitive to these sentiments. In this case, it seems relatively easy to explain why our moral judgments and the moral facts generally agree. After all, on this view, both the moral facts and our moral judgments are a function of a single common factor - namely, our moral sentiments.

Explaining the reliability of our moral judgments becomes more difficult, if we think of the fundamental moral facts as wholly mind-independent. For example, consider a more realist version of sentimentalism, which takes our moral sentiments to be our primary means of access to the moral facts, but takes many of these facts to be independent of our sentiments. Unlike the previous view, this second position cannot explain the reliability of our moral sentiments via an explanatory connection that runs from our moral sentiments to the moral facts. And the most obvious alternative to this strategy raises worries of another sort. For suppose our realist sentimentalist tries to appeal here to an explanatory connection in the opposite direction by, say, taking our moral sentiments to be partially caused or constituted by certain moral facts. While this might help explain the reliability of these sentiments, it would also call into question whether they can be understood in purely naturalistic terms. Thus, this sort of "robust realism" about moral facts has seemed to many to conflict with a purely naturalistic understanding of moral psychology.¹

The most fashionable way to bring out this sort of conflict appeals to an evolutionary account of the origins of our moral faculties. In particular, such "evolutionary debunking arguments" claim that evolutionary considerations show us that our moral faculties cannot give us knowledge of

¹ A simpler style of debunking argument can be found in Harman(1977). The debate over this style of argument is complex, but it seems to many to beg the question against the moral realist, by assuming that we ought to form our beliefs *solely* through scientific forms of inference to the best explanation.

mind-independent moral facts. Thus, if such debunkers are right, we must choose between a robust and non-skeptical form of moral realism and a naturalistic picture of human psychology and its origins.²

Not surprisingly, debunkers do not agree on the proper response to these arguments. For some, they apply very broadly - so that an error-theory about morality is the only rational response to them.³ For others, these problems are best dealt with by turning to a non-cognitivist metaethical theory on which questions of reliability are arguably misplaced.⁴ But many debunkers believe that the proper reaction to them is something more modest - namely, a turn away from robust forms of moral realism towards a view on which the moral facts are dependent upon our psychology.⁵ Thus, many debunkers take these arguments to support a cognitivist form of sentimentalism or constructivism on which there are moral facts, but these facts are a function of our moral sentiments or attitudes.

For the most part, I will focus on the last of these alternatives in what follows. I believe that a moral error theory is best avoided if at all possible. And while I have a good deal of sympathy with quasi-realist forms of expressivism, the issues such views raise lie outside the scope of this paper. So while some of what I say will be relevant to these views, I will focus here on debunking arguments that aim to establish that moral facts are mind-dependent, while remaining within the context of cognitivism about moral judgment. Such arguments aim to undermine - not the very existence of moral facts - but rather the "robust realist" claim that these facts are independent of our moral psychology.

But should we find such arguments convincing? To consider this question, I first distinguish two basic forms such arguments can take - focusing on the epistemological assumptions that drive their skeptical conclusions. Then I argue that we have good reason to be skeptical of these assumptions on general epistemological grounds. In doing so, I note that many moral anti-realists also have reason to reject these assumptions - for they are powerful enough to jeopardize moral knowledge even given many forms of anti-realism. Thus, as will become clear, what drives these

² Of course, one may be skeptical of the project of giving a wholly naturalistic or non-normative characterization of the mind. I am not unsympathetic to such views, but for the sake of argument, I will mostly concede to the debunker that a wholly naturalistic account of the nature and development of the human mind is preferable. For an argument against this, see Wedgwood(2007).

³ Joyce(2006).

⁴ Blackburn(1998), Gibbard(2003). Compare Schafer(2013).

⁵ Street(2006), Bedke(2009), Setiya(2012).

"evolutionary debunking" arguments often has less to do with the distinction between realism and anti-realism *per se*, and more to do with the details of the explanatory structure of one's metaethical views.

In this way, the majority of this essay will be dedicated to defending moral realism against the debunker. But in the end, my attitude towards these arguments is more ambivalent. For while I do think that even "very robust" moral realists can resist the hard choice between anti-realism and skepticism the debunker wishes to force upon them, I also agree with the debunker that her arguments point to a legitimate source of dissatisfaction with many forms of moral realism. Thus, even if these arguments fail as arguments from robust moral realism to skepticism, they succeed in raising real questions about the explanatory structure of many forms of moral realism - questions that might very reasonably be taken to motivate a broadly mind-dependent view of morality of a sentimentalist or constructivist sort.

1. Two Varieties of Debunking

Let's begin by defining the kind of moral realism that will be our focus here. Although some debunkers have attempted to debunk *all* forms of moral realism – and, indeed, all forms of quasi-realism – the clearest target of such arguments are “robust” forms of *non-naturalist moral realism*: views on which the fundamental moral facts cannot be reduced to or wholly explained by any combination of facts about our mind or other natural facts:

Robust Realism: There are moral facts; and the *fundamental* moral facts are explanatorily independent of facts about our normative attitudes, as well as other evolutionarily significant natural facts.⁶

Thus, the Robust Realist claims that there is a domain of fundamental moral facts – facts that hold independently of facts about moral psychology - or any other relevant naturalistic facts.⁷

⁶ See Cuneo(2007), Enoch(2011), Fine(2002), Fitzpatrick(2015), Nagel(1986), Parfit(2011), Scanlon(2014), Shafer-Landau(2003/2005), Wedgwood(2007).

⁷ The latter qualification is important, because in order to block the debunking arguments considered below, it will be sufficient that the fundamental moral facts are explained by features of the human species that also explain our moral judgments and reactions. Thus, in order to resist these arguments, we need not draw a *direct* explanatory connection between the moral facts and our moral judgments or reactions. It will be enough if we can draw an explanatory connection (of the right sort) between each of these elements and some third factor, which is itself amenable to naturalistic explanation. (Kauppinen in this volume)

Robust Realism may seem an implausibly strong metaethical view. But in considering it, it is important to stress that it does not insist that the moral facts are mind-independent or human-independent in the sense of having *nothing to do with* the nature of human beings or their minds. For example, one popular candidate for the status of fundamental moral truth is the fact that pain is bad, which concerns the badness of certain psychological states. Indeed, on such views, the fundamental moral facts may themselves establish *explanatory* connections between, say, the fact that something is painful and the fact that it is bad. What is crucial for the Robust Realist is simply that *these most fundamental* moral facts are not themselves explained in terms of further natural facts.

Thus, on any such view, more determinate moral facts about human beings will be explained in part by natural facts about us, in conjunction with more fundamental moral facts. The important point for the Robust Realist is simply that the fundamental fact that, say, *pain is bad* is a product neither of our evaluative reactions to being in pain (at least naturalistically described) nor any other natural fact. Thus, the Robust Realist need not deny that morality is especially concerned with (say) pleasure and pain. And they need not deny that many moral facts are partially explained by natural facts. All they must insist is that the *fundamental* moral facts are not explained in this way.

One way of understanding the core claim of **Robust Realism** is in terms of the idea that the most fundamental moral facts do not depend on anything contingent or parochial to a specific culture, community, or species. Again, in saying this, the Robust Realist need not deny that the vast majority of moral facts about human beings *do* depend on such facts. Rather, what is important to the Robust Realist is simply the idea that these more determinate moral facts about human beings in particular are explained by more fundamental moral facts that hold independently of contingent natural facts.

How, then, do debunking arguments attack **Robust Realism**? Unsurprisingly, they begin with some sort of empirical claim about the origins of our normative faculties and judgments. Quoting Sharon Street, we might put this claim as follows:

Assumption: We have come to make the normative judgments we do because “the mindless process of evolution by natural selection shaped us that way”.⁸

For our purposes, what matters most about **Assumption** is not the details of how this process of evolution is understood. Rather, what is important is that **Assumption** implies that the full explanation of why we make the normative judgments we do does not appeal to moral or normative facts, at least when these facts are characterized in moral or normative terms.⁹ Of course, establishing this does not require any appeal to evolution by natural selection *in particular*. All that is required is a high degree of confidence that the best explanation of why we make the moral judgments we do is purely naturalistic. But nonetheless evolutionary theory remains relevant here insofar as the success of evolutionary theory increases our confidence in the possibility of this form of explanation.

Given **Assumption**, the basic form of these arguments is quite simple. From **Assumption**, the debunker argues that, if **Robust Realism** were true, any fit between our moral judgments and the moral facts would, at least at the most fundamental level, be an *accident* or *coincidence* in an epistemically problematic sense. From which they conclude that these judgments cannot be regarded as genuine instances of moral knowledge because of the presence of the relevant form of accidentality or luck.

Crucially this last step can be made in two quite different ways. First, the debunker may argue that we should regard accidents of this sort as *highly unlikely*. If she proceeds in this way, she will then conclude that, conditional on the conjunction of **Assumption** and **Robust Realism**, we should regard it as highly unlikely that our moral judgments are reliable.¹⁰ In short, on this version of the argument, we would have to believe something quite *incredible* in order to regard our normative faculties as reliable (given **Assumption** and **Robust Realism**). Thus, even if we do possess an initial entitlement to treat our moral judgments and faculties as reliable, this entitlement will be defeated by the recognition that they are highly unlikely to be reliable, given

⁸ Street(2006),(2009). An interesting question is whether a focus on biological as opposed to cultural evolution matters here. This, of course, will depend greatly on one's views about the nature of cultural explanation. But I will follow Street in assuming that the influence of biological evolution on the development of our faculties and judgments is sufficient on its own to raise these concerns.

⁹ This qualification is important to allow for the possibility that the moral or normative facts are identical with certain naturalistic facts. For the most part, I will focus on the moral case, but similar considerations apply to normative judgments more general, and it will sometimes be useful to include these in the scope of my discussion.

¹⁰ Compare Street's discussion. See also Joyce(2007) and Bedke(2009) among others.

what we know about their origins.¹¹ I'll refer to this style of argument as the **Argument from Incredibility**.

This is perhaps the most familiar way for such arguments to proceed. But there is another way the debunker could appeal to claims about “accidental reliability” to attack moral knowledge. Rather than entering into a debate about our entitlement to believe that our moral faculties are reliable, she might simply appeal to the attractive idea that knowledge requires non-accidentally true belief to argue for this conclusion *directly*. For example, she might claim that a belief can only count as knowledge if it is formed by a method that is non-accidentally reliable:

No Accident: When S knows that P, she knows it by a reliable method, and it is no accident that her method is reliable.¹²

If this is right, and if the sense of “accident” that this condition appeals to is the same as the one that is involved in the debunker’s earlier claims, then the debunker would be in a position to conclude that knowledge of the fundamental mind-independent moral facts is impossible for us (given **Assumption**) simply in virtue of our moral beliefs lacking the relevant sort of non-accidental connection with the moral facts. Thus, while the first of these styles of argument attacks moral knowledge via attacking our entitlement to believe that we are reliable about moral questions – and then infers from this that the relevant sort of moral knowledge is impossible – the second style of argument attacks the possibility of such moral knowledge directly, by arguing that the realist cannot satisfy a basic condition on *what it is* for a belief to count as knowledge. I'll refer to this second argument as the **Argument from Non-Accidental Truth**.¹³

As we will see, there are costs and benefits to each of these strategies. But the upshot of them is quite similar. Given **Assumption**, both aim to leave us with a choice between moral skepticism

¹¹ The initial entitlement at issue here can be understood in at least two ways. First, we might understand it in terms of an *a priori* entitlement to *believe* that our moral faculties are reliable. Or, second, we might understand it in terms of the idea that we are entitled to form moral judgments through the *use* of our faculties, without thereby having any sort of entitlement (prior to the use of our faculties) to *believe* that these faculties are reliable. The first option corresponds to so called “rationalist” responses to skepticism. (See, e.g., White (2005). The second represents a form of “dogmatism” in Jim Pryor’s influential sense of this term. (Pryor(2000)) The difference between these views won’t matter for our discussion– but I am inclined to view both of them as capturing part of the truth about these matters. In particular, the first may accurately characterize our *propositional* justification to believe that our faculties are reliable, and the second may characterize our *doxastic* justification for such beliefs. See Willenken(2011).

¹² See Yamada(2011) and Setiya(2012).

¹³ This is Setiya's preferred version of the argument.

and the rejection of **Robust Realism**. Thus, those unsympathetic to moral skepticism, and sympathetic to a naturalistic account of the nature and origins of the mind, are left with a powerful argument in favor of some form of moral anti-realism - or least, some "less robust" form of realism or quasi-realism.

2. The Argument from Incredibility

But does either of these arguments really require us to choose between moral skepticism and anti-realism in this way? There is reason to be skeptical in both cases.

To see why, let's begin with the first style of argument. In considering it, it is important to recognize that it is not meant to devolve into a general skeptical argument against our entitlement to trust *all* of our faculties. Rather, the debunker means to offer a targeted skeptical argument – one that (given **Robust Realism**) effectively undermines our entitlement to trust our moral or normative faculties, while leaving our entitlement to trust our faculties for perception or non-moral reasoning intact.¹⁴

Given this, the debunker should allow that we are equally entitled *a priori* to rely on all of our faculties – both moral and non-moral. And she should concede that this *a priori* entitlement is equally robust and resilient in the moral and non-moral cases. Otherwise, her argument will beg the question against the realist who claims that prior to any empirical discoveries about our faculties, our moral and non-moral faculties deserve the same epistemic respect.¹⁵

Accordingly, the debunker should be understood as claiming that what we have discovered about the empirical origins of our moral judgments is sufficient to *undermine* our initial entitlement to treat these faculties as reliable. To put things in Hume's terms, what is at issue here is a form of consequent, not antecedent, skepticism.¹⁶

¹⁴ Plainly this is essential to her aims – otherwise the debunking argument would be self-undermining.

¹⁵ Again, this is the basic problem with Harman's argument. This section draws on the discussion in Schafer(2010). For related discussion see Enoch(2010), Wielenberg(2010), Clarke-Doane(2012), Shafer-Landau(2012), Vavova(2014), Berker(2014), and FitzPatrick(2015).

¹⁶ Hume(1999).

Of course, there is nothing very odd about this idea. For there are many cases in which discoveries about the origins of some belief or faculty or method has this effect. For example, consider the following, due to Dustin Locke:¹⁷

Martian: Jack has not received the training of an ordinary physics student. Rather, Martians brainwashed Jack to believe that certain kinds of streaks are caused by protons. Moreover, they brainwashed Jack to have this disposition, rather than a disposition to believe that something else causes such streaks, not because they themselves had done any physics, but simply because they liked the sound of the word ‘proton’. You can even suppose, if you like, that there is some deep law of Martian psychology that makes them like the sound of the word ‘proton’, and so in what is a certain sense inevitable (given the initial conditions) that Jack would come to have this disposition. Now Jack sees a streak and believes there goes a proton. However, just after coming to believe this, Jack learns that he was brainwashed in just the manner described above, and that he has not received the training of an ordinary physics student.

As Locke notes, it is very plausible that Jack’s discovery about the origins of his disposition to believe that certain kinds of streaks are caused by protons is sufficient to undermine whatever entitlement he might otherwise have to treat this disposition as reliable. But Jack’s case is very different from the debunking arguments we are discussing. In particular, Jack presumably has a great deal of background evidence in favor of the proposition that dispositions for forming beliefs about matters of empirical fact are generally only reliable when they are caused in certain ways. Thus, in Jack’s case, we can explain the debunking effect of his discovery via a straightforward appeal to his background knowledge about when dispositions for the formation of empirical beliefs are reliable – plus his discoveries about the nature of his disposition in this case.

This is a plausible model for many *local* debunking arguments, which target only some of our moral judgments. For many of these do appeal to conflicts between what empirical psychology tells us about some of our moral judgments and our background moral beliefs. But this style of argument does not seem promising for the debunker who aims to undermine *all* our moral judgments and faculties simultaneously. After all, any knowledge we might have of the conditions under which our moral faculties are reliable is itself (at least in part) a species of moral

¹⁷ Locke(forthcoming).

knowledge. And if we are allowed to appeal to background moral beliefs about *this* in the present context, we should also be allowed to appeal to other background moral beliefs as well. But if we are allowed to appeal to our background moral beliefs in evaluating the reliability of our moral faculties, the result of this evaluation will often be vindicating as opposed to debunking. After all, the debunker has given us no reason to think that our moral faculties generally produce results that conflict with our background moral beliefs. Indeed, given that these background beliefs are the product of these very faculties, it would be very surprising if this were the case across the board.

On the other hand, suppose we do bracket all of our background moral beliefs, and consider the significance of the discoveries we've made about the development of our moral faculties. In this case, the problem for the debunker is that, given this, it is unclear that these discoveries give us *any* evidence about whether these faculties are likely to be reliable. Such discoveries do tell us that the development of our moral faculties was sensitive to certain natural facts, but without drawing on *some* background moral beliefs, we seem at this stage to have no reason to regard this influence as *either* distorting *or* truth tracking with respect to the moral facts.

In short, if we rely on our background moral beliefs, many of these influences will (unsurprisingly) appear to have put us in a position to track the moral facts. But if we bracket our background moral beliefs, we seem to have insufficient evidence to judge whether these influences are likely to be distorting. So from neither of these perspectives, do we gain the *positive evidence of unreliability* that the debunker needs to provide us with in order to undermine our initial entitlement to rely on our moral faculties.

Against this one might insist, as Locke does, that we can *generalize* from conclusions about the conditions under which empirical faculties are reliable to conclusions about conditions under which moral faculties are reliable. Thus, the debunker might argue that we do not need to rely on our moral faculties or any background moral knowledge in order to arrive at conclusions about the conditions under which moral faculties are likely to be reliable. But any such analogy is likely to be contested in the present dialectical context. For example, the realist is likely to insist that the proper epistemology for the moral domain differs from the proper epistemology of sense perception in ways that make such analogies misleading.¹⁸

¹⁸ Compare Parfit(2011) and Scanlon(2014).

At this point, the "incredibility" debunker may turn from cases like Locke's to cases in which it is intuitive that the origins of our *moral* beliefs has a dramatic debunking effect. For example, suppose I know that my basic moral beliefs were implanted in me via a random process that has a fixed 10% chance of giving me true moral beliefs. Plainly in this case, I should lose whatever confidence I might have had in these beliefs. But this case is importantly different from the actual developmental history of our moral beliefs in at least two crucial respects. First, in this case we *know* that our moral beliefs are the product of a random process, and second, we *know* that this process has a low chance of producing true beliefs.

Crucially, neither of these is true of the evolutionary origins of our moral beliefs. For while the evolutionary processes that produced our moral beliefs may not have been sensitive to the moral facts *as such*, these processes were hardly random. Indeed, the results of contemporary evolutionary theorizing suggest that these processes are likely to lead creatures like us to develop importantly similar moral faculties and judgments, even beginning from a fairly wide range of initial conditions.¹⁹ Thus, rather than being a random process, evolution tends to produce creatures who treat certain sorts of naturalistic properties as morally significant. For this reason, the comparison of evolutionary processes to truly random processes like coin flips or dice rolls is quite misleading.²⁰ In particular, given this, the question facing us is whether the *non-random* outputs of these evolutionary processes are likely to be reliable at tracking the moral facts. And we have yet to see a way of tackling this question without drawing upon some sort of background moral (at least normative) knowledge.

3. The Argument from Incredibility Reconsidered

In response to this, the debunker may concede that we do in fact have to draw on *some* background moral (or at least normative) claims to make the **Argument from Incredibility** work. But they may go on to insist that these claims are so minimal and uncontroversial that this need not call the argument into question.²¹

¹⁹ Enoch(2010), Schafer(2010), and Wielenberg(2010).

²⁰ Against this, one might argue that the *diversity* of possible moral systems gives us sufficient reason for skepticism. But given that evolutionary processes are not random, this style of argument has very little to do with the evolutionary origins of our moral faculties. Rather it relies on claims about the epistemic significance of *possible moral disagreement*. I leave the skeptical significance of *that* to the side, but see Schafer(2015a) for some relevant discussion.

²¹ Street.

To see how some debunkers have tried to take this step, it is helpful to ask how the realist can explain the proposition *that our normative judgments are generally reliable*, given **Assumption** and **Robust Realism**.²² Now, the realist may have a good deal to say about why this proposition is true.²³ But, at least according to the debunker, if **Assumption** is true, whatever explanatory story they tell will end up having something like the following structure:

- (1) A naturalistic explanation of why our moral faculties treat certain naturalistic properties as having normative significance.
- (2) Certain further moral claims about the moral significance of the properties these faculties track – claims that at some stage bottom out in certain fundamental moral facts for which no further explanation can be given.

Thus, if the debunker is right, the realist's explanation of our moral reliability will ultimately rest on a *basic and unexplained connection* between the domain of moral facts and domain of evolutionarily significant natural facts - one that delivers the happy result that our faculties are reliable at tracking the moral facts. This *is* a sense in which the realist's explanation of our reliability about moral questions must rest on an "accident" or "coincidence". Thus, if this were right, there *would* be a sense in which our reliability about moral questions is ultimately the product of a happy coincidence.

One way to challenge these claims would be to question whether it is possible to give a purely naturalistic characterization of the nature and development of our moral attitudes and judgments.²⁴ I have some sympathy with this sort of partially normative conception of the mental. But taking this line in the present context calls into question the truth of **Assumption**, which we are taking for granted. Thus, I'll leave this response to the side.

Instead, what I want to argue is that, even if this is the shape the realist's account must take, we should be skeptical of the epistemological conclusions the debunker wants us to draw from this fact. For even if the debunker is right about the structure of the realist's view, what we have discovered about this view is simply that it rests on a basic and unexplained connection between certain basic moral facts and certain evolutionarily significant natural facts. Thus, whatever

²² This is exactly what debunkers like Street have done.

²³ See Enoch(2010), Schafer(2010) Wielenberg(2010), and Berker(2014).

²⁴ Fitzpatrick(2015).

debunking force this argument has must derive from the idea that we should assign a low prior probability to possibilities that involve such “coincidences”.

Now, there is some plausibility to the thought that we ought to assign a relatively low prior probability to “coincidences” of this sort, all things considered. Indeed, a commitment to this idea appears to follow from the rationality of inference to the best explanation as a method of belief formation.²⁵ But the debunking arguments under consideration greatly overstate the force of such considerations. To see why, it is helpful to consider the following points:

- (1) Any non-theistic theory of reality (other than perhaps Spinoza’s) will in the end appeal to a number of basic, unexplained connections of this sort.
- (2) While these unexplained facts may not be relevant to the explanation of the reliability of our moral attitudes on every such theory, they will be relevant to the explanation of *some* facts. Thus, in comparing such theories, it is important not to focus *exclusively* on their ability to explain *one* fact in particular.
- (3) Furthermore, while the non-skeptical moral realist may have to posit certain basic connections of this sort between the evolutionarily significant naturalistic facts and the moral facts, it is far from obvious that he will have to posit *very many* – or that these connections will be at all *implausible*.

Thus, while the non-skeptical moral realist may have to posit certain basic connections between the evolutionarily significant naturalistic facts and the moral facts, it is far from obvious that he will have to posit *very many* or *very costly* connections of this sort. So the costs of **Robust Realism** along this dimension may not be very great. Indeed, it is not clear that the realist will have to posit *more* connections of this sort than debunker’s preferred form of anti-realism does. After all, many anti-realist views require us to make a number of basic, unexplained claims about the manner in which our moral attitudes *determine or ground* the moral facts. And it is possible for the realist to conceive of the fundamental moral facts as facts about how certain natural facts ground or determine certain non-fundamental moral facts. In which case, the basic structure of these two views on this score will be quite similar. Thus, it is not clear that the combination of

²⁵ It’s important to distinguish this concern for explanation from related concerns about ontological parsimony. Such concerns also raise questions about **Robust Realism**, but their relationship to these explanatory questions is complicated.

anti-realism and an evolutionary account of our views scores better in this regard than the combination of non-skeptical moral realism and an evolutionary account does.

Of course, the force of this point might be mitigated by treating these fundamental connections between the moral and the natural as identities or reductions - e.g. by seeing them as essential to the relevant moral properties.²⁶ But such a move is equally open to *both* realists and anti-realists – *provided* they are willing to give up on the fully non-reductive ambitions associated with **Robust Realism**. Thus, this possibility doesn't undermine the basic symmetry between realist and anti-realist views here - although it may put pressure on *both* to move towards a more reductive understanding of the relationship between the moral and the natural facts.

Moreover, and more importantly, while I agree that the existence of such connections *is* a potential cost of the non-skeptical realist's view, it is not nearly as great a cost as the debunker's argument suggests. Most theories of reality will in the end appeal to some number of basic, unexplained connections of this sort. So while we should aspire to a complete theory of everything that involves as few such connections as possible, this concern for explanatory simplicity must be treated as *defeasible* in the face of our total evidence. Even if such "coincidences" are *a priori* somewhat unlikely, they are not *so* unlikely that it is always unreasonable to believe that they have occurred. For we can, and often do, acquire evidence that just this is the case.²⁷

Thus, even if these considerations mean that a naturalistic explanation of our moral faculties and views should *somewhat* lower the probability that we assign to these faculties reliably tracking mind-independent moral facts, no reason has been given why it should lower this probability to such a degree that our moral faculties lose *all* evidential significance. To think otherwise is precisely to make the presence of a "coincidence" or unexplained connection of this sort a decisive theoretical failing. And this would give considerations of explanatory simplicity an importance they don't have.

Thus, the realist should agree with the debunker that "coincidences" of the sort at issue are *relatively* unlikely *a priori*. But he can nonetheless maintain that our total moral and non-moral evidence supports the proposition that we have developed faculties that are generally reliable at

²⁶ For the significance of essentialism in this context, see Rosen(2010) and DasGupta(2015).

²⁷ Setiya(2012).

tracking the mind-independent moral truth. After all, debunking arguments aside, it appears that evolutionary factors often *have* produced moral faculties that track properties that we *take* to have the relevant sort of moral significance. And while the debunker's argument should perhaps diminish our confidence in the reliability of such judgments to some degree, we have yet to see any reason to believe that it should do so to a degree that would completely undermine our reliance on them.

Against this, the debunker might complain that we have no *independent* source of evidence that such a coincidence has occurred. After all, our only source of evidence for this proposition rests on a reliance on the very moral faculties that are under attack. But while this challenge does raise deep epistemological issues – these issues are perfectly general, and have very little to do with the causal origins of our moral beliefs *in particular*. After all, it is just as difficult to give evidence for the reliability of our perceptual faculties that is wholly independent of *any* reliance on those faculties. So while there are genuine questions here, I don't think they support a *targeted* debunking argument of the sort the **Argument from Incredibility** aims to be.

3. The Argument from Non-Accidental Truth

With this in mind, let's turn to the second, more direct, manner the debunker might appeal to claims about “accidental reliability” to attack the possibility of moral knowledge?

Once again, this style of argument relies – not on claims about the probability we should assign to hypotheses with a certain sort of explanatory structure – but rather on the idea that a *condition* on knowledge is that a belief can only count as knowledge if it is formed by a method that is *non-accidentally reliable* in some sense:

No Accident: When S knows that P, she knows it by a reliable method, and it is no accident that her method is reliable.

This condition will only be relevant in the present context, if we understand the notion of “accidental reliability” it appeals to so that it makes contact with the explanatory structure of **Robust Realism**, given **Assumption**'s claims about the origins of our moral faculties. In order for this to be the case, we must read **No Accident** so that the developmental history of our moral faculties is relevant to whether their reliability is an accident in the intended sense. As a result, in

order for **No Accident** to have the “desired” debunking effect, it is normally understood to require an explanatory connection between two facts: first, the fact that some method M is reliable and, second, the fact that the believer is disposed to use M:

Strong No Accident: When S knows that P, she knows it by a reliable method, and there is an explanatory connection between (i) the fact that this method is reliable and (ii) the fact that she uses it.²⁸

Such an explanatory connection can take many forms. For instance, it might involve a variety of different forms of explanation – including, not just causal explanations, but also constitutive or teleological explanations. And more importantly, the explanatory connection of interest to us may be the product of a variety of different *directions* of explanation. So, for example, the reliability of some method might (in part) explain why we use this method – as would be true if the reliability of our moral methods helped to explain why use them. Or this method’s reliability might be explained by our use of it – as would be the case if the moral facts themselves were mind-dependent. Finally, there might be some more complicated story about how the two facts in **Strong No Accident** are explained by a common factor.

Having said all that, the debunker is ready to object to *any* story the **Robust Realist** offers to explain the connection between our moral methods and the reliability of that method. For given **Assumption**, the **Robust Realist** faces a dilemma in trying to connect together the two facts at issue in **Strong No Accident**: (i) the fact that we use method M to form moral beliefs and (ii) the fact that M is reliable at tracking the fundamental moral facts:

On one hand, the Robust Realist cannot claim that (i) is explained by (ii). For (ii) is itself a moral fact. Thus, by **Assumption**, neither it, nor the fundamental moral facts that explain it, plays any role in the explanation of (i).

On the other hand, the Robust Realist cannot claim that (ii) is explained by (i). After all, (ii) is explained by the fundamental moral facts. And, given **Robust Realism**, these facts are not explained by (i) or by any other naturalistic fact relevant to the explanation of (i).

²⁸ For a full development of this argument, see Setiya(2012).

Thus, given these assumptions, there seems to be no way of satisfying **Strong No Accident** – at least with respect to the fundamental moral facts.²⁹

This is the heart of the argument. But it is worth noting that **Strong No Accident** also raises similar worries for forms of moral *anti*-realism that make positive metaethical claims about how the moral facts are determined by the psychological facts.³⁰ For example, consider an anti-realist view that takes as basic the some version of the claim that *the facts about value hold in virtue of our attitudes of approval or disapproval*. Call this metaethical fact **Approval**.

Then we can ask whether knowledge of **Approval** is possible, given **Strong No Accident**. Once again, to explain how this is possible, we have to show that there is an explanatory connection between the fact that (i) we use method M to form metaethical beliefs and (ii) the fact that M is reliable at tracking the basic metaethical facts like **Approval**. And this raises the same issues that **Strong No Accident** generated for the realist. In particular:

On the one hand, we cannot claim that (i) is explained by (ii). After all, provided that irreducible metaethical facts like **Approval** are not part of a naturalistic account of the human mind, then given **Assumption**, (ii) will play no role in the explanation of (i).

On the other hand, we also cannot claim that (ii) is explained by (i). After all, (ii) is explained by **Approval** – a fact that is explained neither by (i) nor by any other naturalistic fact relevant to the explanation of (i).

Thus, once again, on this form of anti-realism, **Strong No Accident** seems to make the fundamental facts about how our attitudes determine the moral facts unknowable. More generally, while a view that accepted **Approval** *could* use **Approval** to explain how we can satisfy **Strong No Accident** with respect to the first-order moral facts, this only pushes the basic challenge that **Strong No Accident** raises with respect to our knowledge of *fundamental* facts back a step.

²⁹ Importantly, the same is not true of the *non-fundamental* moral facts. So one option is to concede to the debunker that knowledge of the fundamental moral facts is impossible, but to insist that this result is not overly skeptical so long as knowledge of the non-fundamental moral facts *is* possible.

³⁰ Of course, neither quasi-realists nor error-theorists will be tempted to make such claims. So these comments only apply to one form of metaethical anti-realism.

In order to avoid this, the anti-realist might dig even deeper – by offering a further explanation of facts like **Approval** – one that allows her to establish the explanatory connection **Strong No Accident** requires of us in *that* case. But this will only invite an infinite regress of such explanatory claims.³¹

In this way, principles like **Strong No Accident** naturally invite an explanatory regress, on pain of making whatever facts they take to be fundamental unknowable simply in virtue of the non-accidentality of knowledge.³² I think this should make us suspicious of **No Accident**'s interpretation of the connection between knowledge and non-accidental truth. In saying this, I don't mean to suggest that there is *no* sense in which knowledge is incompatible with accidental truth or accidental reliability. Indeed, I think there are at least two ways in which this is true – having to do with the non-accidental connection that the truth of a piece of knowledge must have with both the *subject* who possesses the knowledge and the *objects* that are known. These two connections can be roughly glossed as follows:³³

Reliability: If a belief is to count as knowledge, that belief must satisfy some sort of intuitive “reliability” condition – be that a matter of safety or reliability or something less reductive. In this sense, the belief's truth must have a non-accidental connection with the manner in which it was formed.

Attributability: If a belief is to count as knowledge, it must be reasonable to attribute the truth of that belief, as well as the positive epistemic features that are relevant to the **Reliability** condition, to the believer so that these features are regarded as relevant to one's *epistemic evaluation of the believer*.

³¹ A better possibility might be to understand the relationship between the basic facts about value and our attitudes of approval or disapproval as one of *identity* as opposed to grounding or determination. But it is unclear that **Strong No Accident** will not generate the same problems for our ability to know the relevant identities. In any case, this move would be equally available to a realist, provided that they were willing to give up on the non-reductive ambitions of **Robust Realism**.

³² As Berker(2014) notes, such a regress may not be totally unnatural with respect to some forms of anti-realism. For example, someone who accepts **Approval** might also find it natural to see the truth of **Approval** as *itself* grounded in facts about our attitudes—and so on *ad infinitum*. Another option for the proponent of **Strong No Accident** might be to think of **Approval** as constitutive in some sense of the meaning of the relevant concepts. In this case, these facts would have an explanatory connection with beliefs involving these concepts. But then we could ask, as above, how **Strong No Accident** allows for knowledge of *these* facts about meaning.

³³ Schafer(2014). Compare Pritchard(2012).

In the end, I believe that these two forms of non-accidentality can be traced to the role played by our practice of knowledge attribution. But what is crucial for our purposes is that neither of them suggests that anything as strong as **Strong No Accident** is a condition on knowledge. Thus, we can accept that knowledge requires non-accidental truth in both senses without accepting the principle that drives the second debunking argument.

To see why, suppose that I have a basic, innate faculty that provides me with intuitions about some subject matter – say, morality – in a highly reliable, safe and sensitive manner. And suppose I form a belief through normal reliance on this faculty. Plainly such a belief will satisfy **Reliability** – since we can simply stipulate that the faculty in question is reliable in whatever sense is relevant. And whatever else might be true of such a case, it seems reasonable to attribute the truth of this belief to me as a believer, if in fact it is true. After all, it seems plausible that whenever I form a belief through normal reliance on one of my basic epistemic faculties, the positive epistemic features of this belief are relevant to the epistemic evaluation of me.

Crucially, this seems to be the case however the reliability of this faculty is *ultimately* explained. For instance, it seems reasonable to give me this weak sort of "credit" for the truth of this belief whether or not the reliability of this faculty is *ultimately* the product of some grand cosmic accident deep in the developmental history of creatures like me. Indeed, it is hard to see how such a fact *could* be relevant to the attributability of these features of my beliefs to me in the sense at issue.

This fact about **Attributability** in the case of *basic* faculties or methods for belief formation is easily obscured by a focus on *derivative* methods for belief formation – that is, methods whose use is explained by the use of some other method. In the case of derivative methods, it is plausible that the positive epistemic features of some belief can only be attributed to me if the method that led me to use the derivative method in question was itself reliable.³⁴ It is this idea that drives the intuitions behind **Strong No Accident**. But while this thought is plausible in the case in which

³⁴ For example, suppose that John accepts some highly complex but ultimately truth-conducive *non-basic* method M, but does so on obviously irrational grounds – such as a recommendation of M by the *National Inquirer*. Intuitively the beliefs that John forms via M do not count as knowledge. And this seems to be because it would be unreasonable to attribute the truth-conduciveness of this method to John in the sense discussed above. This shows that the causal history of our methods is sometimes relevant to the conditions above – my point is simply that this is not always the case in the manner **Strong No Accident** requires. For a view on this topic that does not support the **Argument from Non-Accidental Truth**, see Bernecker(2011).

our use of some method depends on the use of a further method, it seems to me unmotivated and, indeed, counterintuitive in the case of basic methods.³⁵

In this sense, at least in the case of basic methods or faculties, **Reliability** and **Attributability** are relatively “local” conditions on belief formation – neither is sensitive to the deep explanatory history of my beliefs or faculties. Thus, contrary to what **Strong No Accident** implies, the *ultimate* explanation of why such basic method or faculties does not seem to play a role in either of these two conditions on knowledge.

That having been said, I know that my intuitions about such cases are not universal. But I can see no reason to think that the two non-accidentality conditions noted above will imply anything as strong as **Strong No Accident**. Thus, it seems to me that **Strong No Accident** should *at best* be regarded as a highly controversial epistemological principle – one that goes well beyond the established sense(s) in which knowledge requires non-accidental truth. Given this, it is hard to see why the realist should feel compelled to accept it – or any debunking argument that relies on it.

4. Dissatisfaction with the Explanatory Structure of Robust Realism as a Motivation for Anti-Realism

Thus, each of the two debunking arguments ultimately rests on epistemological claims that are highly controversial. As a result, a non-skeptical moral realist can resist these arguments by taking issue with the epistemological presuppositions at their heart. Of course, in doing so, she will take on epistemological commitments of her own. But these commitments are not implausible. Indeed, they are far more modest than the epistemological commitments on which the debunking arguments rely. And it is hardly surprising that the moral realist will have to take a stand on certain epistemological issues in resisting moral skepticism.

Unfortunately for the realist, this is hardly the end of the story. For whatever we think of these epistemological questions, these arguments point to something about the explanatory structure of **Robust Realism** that we *should* find unsatisfying – even if this dissatisfaction is not nearly as dramatic as the debunker suggests. In particular, I believe that the debunker is correct in her

³⁵ For example, consideration of basic methods makes clear that it must sometimes be possible for the positive epistemic features of a method to be attributable to someone even though they have no independent evidence of its reliability. See Enoch and Schechter(2008).

insistence that the very structure of many robust forms of moral realism limits the degree to which such views can offer us a *satisfying* explanation of our reliability about moral questions. After all, so long as we think that the most fundamental explanation of why we think what we do about moral questions is naturalistic, any attempt by Robust Realists to explain our reliability in the moral domain will, at least at the fundamental level, rest on a conjunction of this naturalistic story with a moral explanation of why the relevant natural properties are morally significant. Of course, there are various moves the Robust Realist can make to minimize the explanatory costs of this aspect of his view. But in the end, any form of moral realism with this structure will be limited in its ability to offer us a truly satisfying explanation of our reliability about moral questions.³⁶

This is a real cost of forms of realism with this structure, one that it would be best to avoid or at least to minimize – something which helps to explain the perennial appeal of debunking arguments. For while these arguments may fail at establishing that moral realism leads inevitably to skepticism, they nonetheless point to a real cost of many forms of robust moral realism – and a real potential advantage of their competitors.

In particular, even if a mind-dependent account of normative facts does come to rest on certain basic, unexplained claims about how our attitudes ground the normative facts, the structure of such accounts seems to offer us *more* of an explanation than their realist competitors of *why these attitudes are reliable*. After all, this sort of anti-realism makes the reliability of our normative faculties *essential* to the nature of the normative facts and attitudes themselves, which seems to answer that question in as satisfying a manner as may be possible.

Moreover, while this is only one explanatory task among many, it is one with *considerable* significance for the success or failure of a metaethical theory. Thus, it is no surprise that even realists feel drawn to the sorts of intuitions that drive mind-dependent or constructivist views of the normative:

³⁶ An important question is whether moving to a reductive form of moral naturalism is helpful in this regard. In thinking about this, it is important to stress that a reduction of the moral to the natural will only be helpful insofar as it establishes a genuine *explanatory* connection between the moral facts and the relevant natural facts. Thus, it is not sufficient to simply claim that the moral facts are identical with or reducible to certain natural facts on the metaphysical level, these claims must also provide us with an explanation of why the moral facts are as they are claimed to be.

The connection between objectivity and truth is therefore closer in ethics than it is in science. I do not believe that the truth about how we should live could extend radically beyond any capacity we might have to discover it (apart from its dependence on nonevaluative facts we might be unable to discover).³⁷

In short, these explanatory considerations do provide considerable support for the idea that there should be some sort of *constitutive connection* between the nature of our normative attitudes and the normative facts themselves.³⁸ Thus, even if these arguments fail to reduce **Robust Realism** to skepticism, they do help to demonstrate what is so attractive about less robustly realist accounts of the normative domain.

Indeed, this way of understanding the dialectic may actually make anti-realism more attractive than the understanding of the dialectic many debunkers prefer. Given an understanding of these arguments like Street's, we must either reject these arguments or be pushed, like Street, to a very extreme form of anti-realism. But if we understand these arguments in terms of *how good* an explanation of our reliability a theory makes possible, we are free to reject this choice between extremes in evaluating the metaethical significance of these considerations. Doing so is important because it allows us to see the nuanced costs and benefits of different forms of realism and anti-realism with respect to these explanatory questions. Thus, it opens the door to a variety of views that are intermediate between the most "robust" forms of realism and the most "extreme" forms of anti-realism. Of course, to properly evaluate these different views, we would need to carefully consider the costs and benefits associated with their explanatory structure. But it is there, as opposed to simple attempts to show that some form of realism implies skepticism, that we should focus our attention in these debates.

³⁷ Nagel(1986), 139.

³⁸ There are views that might reasonably be regarded as forms of "realism" that provide us with a connection of this sort. Compare Schafer(2015b).

References

- M. Bedke. "Intuitive Non-Naturalism Meets Cosmic Coincidence." *Pacific Philosophical Quarterly* 90 (2009): 188–209.
- P. Benacerraf. "Mathematical Truth." *Journal of Philosophy* 70 (1973): 661–79.
- S. Berker. "Does Evolutionary Psychology Show That Normativity Is Mind-Dependent?" in J. D'Arms and D. Jacobson (eds.), *Moral Psychology and Human Agency: Philosophical Essays on the Science of Ethics*. Oxford: Oxford University Press, 2014.
- S. Bernecker. "Keeping Track of the Gettier Problem." *Pacific Philosophical Quarterly* 92:2 (2011): 127-152.
- S. Blackburn. *Ruling Passions*. Oxford: Oxford University Press, 1998.
- J. Clarke-Doane. "Morality and Mathematics: The Evolutionary Challenge." *Ethics* 122(2) (2012): 313–40.
- T. Cuneo. "Recent Faces of Moral Nonnaturalism." *Philosophy Compass* 2 (6) (2007):850-879.
- S. Dasgupta. "The Possibility of Physicalism." *The Journal of Philosophy* 2015.
- D. Enoch. 'The Epistemological Challenge to Metanormative Realism: How Best to Understand It, and How to Cope with It.' *Philosophical Studies* 148 (3) (2010):.413 - 438.
- *Taking Morality Seriously: A Defense of Robust Realism*. Oxford: Oxford University Press, 2011.
- D. Enoch and J. Schechter. "How are basic belief-forming methods justified?" *Philosophy and Phenomenological Research* 76 (3) (2008): 547-579.
- K. Fine. "Varieties of Necessity." in T. Gendler & J. Hawthorne (eds.), *Conceivability and Possibility*. Oxford: Oxford University Press, 2002: 253-281.
- W. Fitzpatrick. "Debunking Evolutionary Debunking of Ethical Realism." *Philosophical Studies* (2015).
- A. Gibbard. *Thinking How to Live*. Cambridge: Harvard University Press, 2003.
- J. Hampton. *The Authority of Reason*. Cambridge: Cambridge University Press, 1998.
- G. Harman. *The Nature of Morality*. New York: Oxford University Press, 1977.
- D. Hume. *An Enquiry concerning Human Understanding*. T. Beauchamp (ed.), Oxford: Oxford University Press, 1999.
- R. Joyce. *The Evolution of Morality*. Cambridge, MA: MIT Press, 2007.
- D. Locke. "Darwinian Normative Skepticism." forthcoming in M. Bergmann and P. Kain (eds.), *Challenges to Moral and Religious Belief: Disagreement and Evolution*.

- T. Nagel. *The View from Nowhere*, Oxford: Oxford University Press, 1986.
- D. Parfit. *On What Matters*. Oxford: Oxford University Press, 2011.
- D. Pritchard. "Anti-Luck Virtue Epistemology." *Journal of Philosophy* 109:3 (2012):247-279.
- J. Pryor. "The Skeptic and the Dogmatist." *Noûs* 34.4 (2000): 517–49.
- G. Rosen. "Metaphysical Dependence: Grounding and Reduction." in B. Hale & A. Hoffmann (eds.), *Modality: Metaphysics, Logic, and Epistemology*. Oxford: Oxford University Press, 2010: 109-36.
- K. Schafer. "Evolution and Normative Scepticism." *Australasian Journal of Philosophy* 88.3 (2010): 471–88.
- "Quasi-Realism, Projectivism, and the Explanatory Challenge." in G. Hubbs & D. Lind (eds.), *Pragmatism, Law, and Language*. Oxford: Routledge, 2013: 136-151.
- "Knowledge and Two Faces of Non-Accidental Truth." *Philosophy and Phenomenological Research* 89:2 (2014): 373-393.
- "How Common is Peer Disagreement? On Self-Trust and Rational Symmetry." *Philosophy and Phenomenological Research* (2015a).
- "Constructivism and Realism in Kantian Metaethics (1 & 2)." *Philosophy Compass* (2015b).
- T. M. Scanlon. *Being Realistic About Reasons*. Oxford: Oxford University Press, 2014.
- R. Shafer-Landau. *Moral Realism: A Defence*. Oxford: Oxford University Press, 2003/2005.
- "Evolutionary Debunking, Moral Realism, and Moral Knowledge." *Journal of Ethics and Social Philosophy* 7 (1) (2012).
- K. Setiya. *Knowing Right From Wrong*. Oxford: Oxford University Press, 2012.
- S. Street. "A Darwinian Dilemma for Realist Theories of Value." *Philosophical Studies* 127 (2006): 109-66
- "Evolution and the Normativity of Epistemic Reasons." *Canadian Journal of Philosophy* 39, supplement 1 (2009): 213–48.
- R. Wedgwood. *The Nature of Normativity*. Oxford: Oxford University Press, 2007.
- R. White. "Problems for dogmatism." *Philosophical Studies*, 131 (2005).
- E. Wielenberg. "On the Evolutionary Debunking of Morality." *Ethics* 120.3 (2010): 441–64
- T. Willenken. "Moorean Responses to Skepticism: A Defense." *Philosophical Studies* 154 (2011): 1-25.

K. Vavova. "Debunking Evolutionary Debunking." in R. Shafer-Landau (ed.), *Oxford Studies in Metaethics*. Vol. 9. Oxford: Oxford University Press, 2014: 76–101

M. Yamada. "Getting it right by accident." *Philosophy and Phenomenological Research* 83.1 (2011): 72-105