

Five Kinds of Epistemic Arguments Against Robust Moral Realism*

Joshua Schechter

I. Introduction

There are many objections that have been presented against nonnaturalist moral realism. The objections involve several different kinds of considerations. There are objections from metaphysics: For instance, how can nonnatural moral properties fit into a general scientific picture of the world?¹ How can moral realism explain the supervenience of nonnatural moral properties on natural properties?² There are objections from the philosophy of language and philosophy of mind: If moral properties are nonnatural properties, how can our linguistic and mental representations stand for them?³ There are objections from moral psychology: If the role of moral beliefs is to represent facts about objective moral properties, what explains the tie between moral belief and moral motivation?⁴ There are also objections based on normative considerations: If moral properties are nonnatural properties more akin to abstract mathematical properties than anything tied to the natural world, how can we explain the significance of morality and the place it does and should have in our lives?⁵

The purpose of this chapter is to examine a different group of objections to nonnaturalist moral realism, objections that are epistemic or at least have an epistemic flavor. These objections involve the nature of knowledge, justification, rationality, and reliable belief. The aim of this chapter is not to answer these objections or to argue that they are successful. Rather, the goal of this chapter is to discuss several families of epistemic objections to nonnaturalist moral realism and to examine which ones are more pressing and which ones may safely be dismissed. In particular, I will discuss five families of epistemic objections to nonnaturalist moral realism: (i) one involving necessary conditions on knowledge, (ii) one involving the idea that the causal history of our moral beliefs reflects the significant impact of irrelevant influences, (iii) one relying on the idea that moral truths as the realist conceives of them do not play a role in explaining our moral beliefs, (iv) one involving the claim that if moral realism is true then our moral beliefs are unlikely to be reliable, and finally, (v) one involving the claim that moral realism is incompatible with there being a plausible explanation of our reliability about morality. My overall conclusion is that most of these objections are not very persuasive. The final

* This is the penultimate draft of a paper in *The Oxford Handbook of Moral Realism* (2023), Paul Bloomfield and David Copp (eds.), Oxford: Oxford University Press, pp. 345-69.

<https://doi.org/10.1093/oxfordhb/9780190068226.013.16>

¹ This is one strand in Mackie's (1977, ch. 1) "queerness argument".

² Blackburn (1984, pp. 182-4).

³ See Wedgwood (2001) for discussion of this kind of objection, and a response to it.

⁴ See Shafer-Landau (2003, p. 121) for a presentation of this kind of objection, which he attributes to several earlier philosophers, going back to Stevenson (1937).

⁵ It is tempting to reply to this sort of objection by saying that there can't be an issue here since, of course, morality has significance – it has moral significance. But this response is too quick. If a metaethical view tells us that morality concerns such-and-such, and such-and-such looks to be insignificant, then we have reason to worry either about the significance of morality or (more likely) the truth of the metaethical theory.

objection, concerning the difficulty of explaining our reliability, is by far the most pressing.⁶ (The third objection, concerning the apparent failure of moral truths to explain our moral beliefs, does show that moral realism should not be taken to be the default metaethical view.)

Before I discuss these families of objections, it will first be useful to briefly present a strong form of nonnaturalist moral realism to serve as the target of these objections. We may call this view “robust moral realism”.⁷ I present this view for two reasons. First, robust moral realism is the metaethical view that I am most tempted by. Second, treating the objections as targeted against a strong form of moral realism will help us to avoid various complexities that arise when additional targets are considered. For instance, we won’t have to determine the exact range of realist views that each argument is best targeted against. (That’s an interesting question, but it would distract attention away from the focus of this chapter.) So that is the task to which I now turn.

II. Robust Moral Realism

Robust moral realism, as I use the term here, has seven theses. The first thesis concerns moral language and moral thought:

Cognitivism. Certain sentences and mental representations purport to represent moral facts. They are both meaningful and truth-apt (that is, capable of being true or false).

This thesis rules out views on which, for instance, the role of moral language is not to represent but rather is to express emotions or other non-cognitive attitudes. It is, however, compatible with the view that moral language has multiple roles simultaneously, such as both expressing emotions and representing moral facts.

The second thesis is the denial of moral error theory:

Non-Error Theory. Some basic moral properties and relations are exemplified.

This thesis rules out error theoretic views on which basic moral properties and relations are never exemplified. Different moral views will take different moral properties and relations to be basic, but natural candidates include *being morally right*, *being morally permissible*, *having moral value*, *being a moral reason for*, and so forth.⁸

The third thesis is the claim that the moral facts are independent of our attitudes, linguistic practices, and social practices.

⁶ This chapter is thus in agreement with Enoch (2011, ch. 7) and Schechter (2017).

⁷ I borrow this term from Enoch (2011).

⁸ We might want to generalize this thesis to also rule out simple patterns of exemplification of basic moral properties and relations. For instance, we might want to rule out a view on which *being morally permissible* is the sole basic moral property and every action is morally permissible.

Independence. The fundamental moral facts do not depend on us. In particular, they do not depend on facts about our minds, language, or social practices.⁹

Of course, some moral facts do depend on our minds, language, and social practices. The fact that someone has a strong preference may be relevant to what I morally ought to do. The fact that a word has a certain meaning in my community may be relevant to what I morally ought to say. And so forth. But these are not the fundamental moral facts – the ones that (in concert with the non-normative facts) determine the rest. According to this thesis, what is independent of our minds, language, and practices are the fundamental moral facts. For example, perhaps it is a fundamental moral truth that it is wrong to cause a gratuitous harm. If so, then the thesis entails that this truth does not depend on our minds, language, or social practices. But it is compatible with the thesis that it does depend on our practices that making a particular utterance would cause a gratuitous harm, and thus be morally wrong.

The fourth thesis is nonnaturalism about the moral:

Nonnaturalism. Moral properties and relations are not identical with, reducible to, or fully grounded in natural properties and relations.

It's controversial exactly how to understand the natural/nonnatural distinction. For instance, perhaps the natural properties should be understood to be the properties that play a role in the natural sciences. Alternatively, perhaps the natural properties should be understood to be the descriptive properties. We don't need to decide between these conceptions. What's important for some of the arguments considered here is a consequence of nonnaturalism, namely that moral facts are not causes of natural facts. And that is plausible on any reasonable characterization of the natural/nonnatural distinction.

The fifth thesis concerns the possibility of moral error:

Intelligibility of Error. It is intelligible and genuinely possible to have moral beliefs that are largely wrong.

This thesis rules out certain kinds of morally plenitudinous views. Consider a view on which there are the properties of (moral) rightness and wrongness as well as the properties of rightness* and wrongness*, the properties of rightness** and wrongness**, and so forth. Suppose that for every coherent practice of assessing actions as “right” or “wrong”, there is a corresponding pair of properties that exactly fits that practice.¹⁰ Finally, suppose that moral semantics works so that the terms “right” and “wrong” as used in a community, and the corresponding concepts of *right* and *wrong*, have their semantic values assigned so as to stand for the pair of properties that most

⁹ The thesis is distinct from what Shafer-Landau (2003, p. 15) calls “stance independence”. Stance independence is the thesis that the moral facts are not true in virtue of being ratified by some privileged stance (or privileged set of stances). That rules out a certain kind of dependence of the moral facts on our practices. By contrast, this thesis rules out any kind of dependence. The reason for the generalization is that realism is also intuitively incompatible with a view on which the moral facts fully depend on our practices but not in virtue of the acceptance of any stance.

¹⁰ We might also want to add that the situation is perfectly symmetric. For instance, there is no pair of properties that is somehow metaphysically privileged.

closely fits the community's usage.¹¹ On such a view, it couldn't have been the case that our moral beliefs were grossly wrong.¹² Ruling out such plenitudinous views is important because they fit the prior theses but are intuitively non-realist (or, at least, non-objectivist) views. The thesis is also important because it is part of what generates some of the intuitive worries facing moral realism. If gross moral error were not intelligible, then (at least *prima facie*) it would not be puzzling how it is that we have moral knowledge, how it is our moral beliefs are reliable, and so forth.

The sixth thesis is that we are not, in fact, largely wrong about morality:

Reliability. The moral claims we believe (at least upon reflection and discussion) are true considerably more often than chance alone would predict.¹³

Despite the intelligibility and possibility of gross moral error, our beliefs are reliable in the sense that a much greater proportion of them are true than would be predicted by chance alone. Notice that this thesis does not have modal force. To be reliable requires only that our actual beliefs are true more often than chance predicts. It is not also required that our beliefs tend to be true in various counterfactual scenarios.

Notice that the Reliability thesis is compatible with our moral view departing from the true moral view in important ways. Our moral view may contain a considerable number of falsehoods or omit a considerable number of truths. One way to think about the thesis is as follows: Consider the set of all possible coherent moral views (or the set of moral views that are at least as coherent as our own moral view). Our moral view is much closer to the true moral view than a random moral view chosen from the set is likely to be. That's because we're broadly on the right track about many moral matters – for instance, suffering is morally bad – and a randomly selected moral view will likely be highly off track.¹⁴

The seventh and final thesis is that many of our moral beliefs have a positive epistemic status:

Justification. Many of the moral claims we believe (at least upon reflection and discussion) are epistemically justified.

¹¹ This view is broadly analogous to Balaguer's (1998) full-blooded Platonism about mathematics.

¹² The Intelligibility of Error thesis should presumably be strengthened to rule out applications of what Lewis (1989, p. 132) calls "the trick of rigidifying". Consider the view that is just like the plenitudinous view with the addition that when members of a community are considering what would be the case in a counterfactual scenario, they apply their actual standards of rightness and wrongness, not their counterfactual standards. On such a view, gross moral error is possible – we can imagine a situation in which we had very different moral beliefs and correctly judge that in that situation we would get it wrong. But this sort of plenitudinous view presumably should be ruled out, too, since it is intuitively a non-realist (or non-objectivist) view.

¹³ For present purposes, we can take "the moral claims we believe" to be the general consensus of presently living people. It is plausible that the Reliability thesis is true on this understanding. But readers who are worried about this claim are welcome to take the relevant group of people to be smaller.

¹⁴ It could be that there are some limits on how far off track a community's beliefs can be on moral matters – if they are too far off track, they are no longer thinking about morality but (at best) something else. See Cuneo and Shafer-Landau (2014) for a view in this ballpark. But even if that's true, the limits are pretty wide – a community can have moral beliefs that are widely off the mark. (Indeed, there are examples of past and present communities with that feature.) See Plunkett (2020) for relevant discussion.

Taken together, these seven theses describe a strong form of nonnaturalist moral realism. This package of theses will serve as the official target of the epistemic arguments discussed below.

III. Argument 1: A Necessary Condition on Knowledge

The first family of arguments I will consider is based on the idea that if robust moral realism is true, then our moral beliefs fail to satisfy some necessary condition on knowledge. More explicitly, arguments in this family have the following schematic form:

- (i) For a belief to count as knowledge requires that it have feature F;
- (ii) If robust moral realism is true, then our moral beliefs don't have feature F;
- (iii) So either robust moral realism is not true or we don't have moral knowledge.

This kind of argument is sometimes extended as follows: (iv) We do have moral knowledge, and so (v) Robust moral realism is false.

Arguments in this family differ on what feature F is. The most familiar version of this argument – albeit concerning a different domain – is what is often called “the Benacerraf problem” for mathematical Platonism.¹⁵ In Benacerraf's argument, the relevant feature F is that the believer be in some causal relation with what the belief is about. More specifically, Benacerraf claims that for a belief to count as knowledge requires the believer to stand in a causal relation with the referents of the names, predicates, and quantifiers featuring in the belief. But, if mathematical Platonism is true, mathematical entities are acausal, so the referents of mathematical terms – e.g., numbers – cannot stand in any kind of causal relation. So either mathematical Platonism is not true or we don't have mathematical knowledge. An analogous argument can be made against robust moral realism.

It is widely recognized that Benacerraf's argument fails against mathematical Platonism and that the analogous argument fails against robust moral realism. The trouble is with the claim that there is a causal constraint on knowledge. Benacerraf made this claim because it was a then-popular response to the Gettier problem for the justified true belief analysis of knowledge.¹⁶ The idea was that the reason that subjects in Gettier cases don't count as knowing the relevant claim is that their belief failed to satisfy a causal constraint.¹⁷ However, proponents of the causal theory, such as Goldman, explicitly exempted a priori knowledge from requiring the satisfaction of a causal constraint, since they recognized that otherwise the constraint would rule out pretty much all a priori knowledge. The causal theory is also no longer popular, and for good reason. There are Gettier-style cases in which the relevant belief satisfies the causal constraint.¹⁸ More

¹⁵ Benacerraf (1973).

¹⁶ Gettier (1963).

¹⁷ Goldman (1967).

¹⁸ Goldman's (1976) fake barn case is such a case. So is the following variant of Chisholm's (1966) sheep in the field case: Suppose a thinker sees something that resembles a sheep in a field in the distance and comes to believe that there is a sheep in the field. If there really is a sheep in the field, but what the thinker saw was a sculpture modeled on the sheep in question, the thinker's belief would be a justified true belief that satisfies the causal constraint but isn't knowledge. (I don't know who first modified Chisholm's example in this way.)

importantly for present purposes, there are cases of knowledge where the relevant belief does not satisfy a causal constraint. These include cases of knowledge of the future: I can know that something will exist tomorrow, say tomorrow's newspaper, without my standing in a causal relation with the future existent. So there are good reasons to reject the causal constraint as playing a role in the analysis of knowledge, or even as being a necessary condition on knowledge.

More recent arguments in this family make use of different (purported) necessary conditions on knowledge. Like Benacerraf's argument, the conditions are supposed to be necessary conditions on knowledge, but not on justified true belief. For instance, arguments of this sort can be presented where feature F is that the belief in question satisfies an explanatory rather than a causal constraint – the truth of the belief is somehow explanatorily connected with the believer's believing it.¹⁹ Or the feature F could be that the belief in question is non-accidentally true,²⁰ or that it satisfies a sensitivity constraint (roughly: if the proposition believed were false, the thinker would not believe it), or that it satisfies a safety constraint (roughly: the thinker could not easily have had a false belief on the matter).²¹

Given such a (purported) necessary condition on knowledge, one could try to argue that robust moral realism precludes our moral beliefs from satisfying the condition. For instance, one could argue that if robust moral realism is true, moral facts do not cause or ground our moral beliefs, so the moral facts cannot explain our moral beliefs. And if robust moral realism is true, by the mind-independence of morality, the moral facts do not depend on our moral beliefs. So there is no explanatory connection between our moral beliefs and the moral facts. Similarly, one could argue that if robust moral realism is true, the truth of our moral beliefs would have to be accidental, since there is no obvious reason why our moral beliefs would correlate with the moral facts. Or, again, one could argue that if robust moral realism is true, our moral beliefs don't satisfy a sensitivity condition because if the nonnatural moral properties were differently distributed, we would not have different moral beliefs, so it is not the case that if the contents of our moral beliefs were false, we would not believe them.²² Or, finally, one could argue that if robust moral realism is true, our moral beliefs don't satisfy a safety condition, since we can easily imagine having had false moral beliefs – indeed, there are plenty of past and present groups with false moral beliefs.

What should we make of such arguments? It is important to notice that they are not dialectically very effective. This is for three main reasons, in order from least to most important. The first reason is that it is not obvious that if robust moral realism is in fact true, the constraints fail to be

¹⁹ For an example of an argument of this sort, see Lutz (2020).

²⁰ For an example of an argument of this sort, see Bengson (2015).

²¹ See Nozick (1981) for a sensitivity constraint on knowledge and Sosa (1999) for a safety constraint on knowledge. The "roughly" hedges are present because there are many ways to formulate such constraints. They can be formulated using counterfactual conditionals, by a direct appeal to possible worlds, or in terms of whether something could easily happen. Some formulations also add a requirement that the counterfactual belief be formed using the same belief-forming method so, for instance, a belief counts as sensitive if the following obtains: If the proposition believed were false, the thinker would not believe it using the same method.

²² Assuming that basic moral truths are necessarily true, this line of thought requires that some counterfactual conditionals with necessarily false antecedents are false, contrary to the orthodox Lewis/Stalnaker semantics.

met. Consider an explanatory constraint. Is it obvious that robust moral realism precludes there being an explanatory connection between our moral beliefs and the moral facts? If we permit common-explanation (i.e., third factor) accounts, then there may be an explanatory connection after all. For instance, on an account suggested by Enoch, our moral beliefs – or, better, the motivational tendencies that ultimately give rise to our moral beliefs – are the products of evolution by natural selection. Evolution “aims at” survival and reproductive success. Survival is, proposes Enoch, of positive moral value. So there is a kind of explanatory connection between our moral beliefs and the moral truths.²³

Similarly, it is not obvious that given robust moral realism, our moral beliefs fail to be sensitive or safe. If the moral truths are necessary, and if counterfactuals with necessarily false antecedents are trivially true, it will be trivially true that if the moral facts were different, the thinker would not still believe them. So the sensitivity constraint will be met. If there is a good causal evolutionary/historical/sociological explanation of how we came to have our moral beliefs, then perhaps it is not the case we could easily have had deeply mistaken moral beliefs, so many of our moral beliefs will in fact be safe. And so on for other purported constraints.

The second, and more important, reason that these arguments are not dialectically very effective is that the purported conditions on knowledge are controversial. They may not be as obviously wrong as Benacerraf’s original causal constraint. But there is nevertheless strong reason to worry about each of them.

It is plausible that there is some kind of non-accidentality condition on knowledge. But the precise formulation of such a condition is a vexed matter. To give an illustration, consider the case of someone who wins a raffle whose prize is a well-known encyclopedia of birds. Suppose the thinker randomly flips to a page and reads that the average airspeed velocity of an unladen European swallow is approximately 24 miles per hour, and comes to believe this. Such a belief (assuming it’s true, etc.) would seem to count as knowledge even though it was highly accidental that the thinker came to believe it. So, plausibly, we have to distinguish between kinds of accidentally true belief and restrict the condition on knowledge to require that the belief not have the wrong kind of accidentality, whatever that is.²⁴

There are worries that face an explanatory constraint on knowledge. For instance, to accommodate knowledge of the future, we should allow common-explanation stories to count as an explanatory connection – my belief that tomorrow’s newspaper will exist and the fact that tomorrow’s newspaper will exist are both explained by various facts about the present and past, e.g., concerning the newspaper publisher. Even permitting such explanatory connections, there are also apparent counterexamples to an explanatory requirement. Consider, for instance, the following case: Suppose that Edgar knows that Allan has taken a fatal dose of poison and that enough time has now passed that Allan is dead. Suppose Allan is in fact dead. However, unbeknownst to Edgar, Allan died because he ran into the street after taking the poison and was

²³ Enoch (2011, ch. 7).

²⁴ Pritchard (2005) distinguishes five varieties of epistemic luck and argues that knowledge is compatible with three of them.

promptly run over by a bus.²⁵ In such a case, Edgar knows that Allan is dead despite there not being an explanatory connection between Edgar's belief and Allan's death, at least not the way it actually happened.

There are also numerous purported counterexamples to sensitivity and safety constraints on knowledge. Against sensitivity, it might be pointed out that anti-skeptical beliefs are not sensitive. Take the belief that I am not a brain in a vat being fed the perceptual experiences of a normal non-envatted person. If this belief were false, I would still presumably have it. So the belief is not sensitive. But it is plausibly an instance of knowledge. There are also more mundane purported counterexamples, such as this one due to Sosa: Suppose I throw a bag of trash down the garbage chute of my apartment. A little while later I believe, and presumably know, that the trash bag is in the basement. However, the nearest world in which my belief is false is one where unbeknownst to me, the bag got stuck somewhere along the chute and in which I still believe it is in the basement.²⁶ Similarly, there are many purported counterexamples to safety, although to be fair, they tend to be a bit more complicated than the counterexamples to sensitivity.²⁷

The third reason that these arguments are not dialectically very effective is the most important one. Namely, despite appearances, the conclusion of the argument is not very worrisome for the robust moral realist. Recall that the conclusion of the argument is that either robust moral realism is false or we do not have moral knowledge. The latter disjunct, at least initially, may sound highly unpalatable. But this is a mistake. The various purported constraints on knowledge are all intended to be constraints on knowledge rather than on justified true belief. It is compatible with the conclusion of the argument that robust moral realism is true, and that while our moral beliefs do not strictly speaking count as knowledge, they do count as justified true beliefs.²⁸ It seems to me that if robust moral realists are forced to endorse this view, they should not find that very concerning. Robust moral realists can still maintain that their moral beliefs are true. They can still maintain that their beliefs are well-supported by the evidence or are justified in whatever way moral beliefs are justified. They can also maintain that it is perfectly rational to hold on to robust moral realism and to their moral beliefs. Insofar as what we do (or should) epistemically care about is to get at the truth, and perhaps also to have justified or rational doxastic states, the robust moral realist needn't feel very concerned by the conclusion of the argument.

There would be a bigger problem for the robust moral realist if the conclusion of this family of arguments could be parlayed into a claim about justification. Lutz suggests a principle of this sort. His principle says, in effect, that if a thinker learns that one of her beliefs lacks some necessary condition on knowledge, this defeats the thinker's justification for the belief.²⁹ For the purpose of extending the arguments, we need a principle that is stronger still, something like: If a thinker learns that, assuming some theory is true, one of her beliefs lacks some necessary

²⁵ This is a variant of a case in Feldman (2003, p. 85).

²⁶ Sosa (1999, p. 145).

²⁷ See Grundmann (2018) for a useful survey of purported counterexamples to safety.

²⁸ This response was originally made to the Benacerraf problem for mathematical Platonism in Burgess and Rosen (1997, p. 37). I think this point has been underappreciated in the literature.

²⁹ See Lutz (2020, p. 293), loosely inspired by Enoch (2011, pp. 161-1).

condition on knowledge, this defeats the justification of the conjunction of the theory with the belief. But in what follows, for simplicity, I'll focus on the weaker principle.

There is reason to think that Lutz's principle is false. If the principle is true, presumably so is the principle that says that having strong reason to believe that one does not know a claim defeats the justification one has for the claim. (Indeed, this latter principle, if true, is plausibly what explains the truth of Lutz's principle.) But it is plausible that the latter principle is not true. Indeed, it is plausible that it is not even true that having strong reason to believe that one does not know a claim is incompatible with knowing the claim. Consider, for example, the case of someone who has strong reason to believe a false theory of knowledge. Perhaps, for instance, the thinker is a college student taking an epistemology course who rationally comes to believe, based on the say-so of their eminent (but mistaken) professor, that knowledge requires absolute certainty. Given that they are not absolutely certain of some mundane proposition (e.g., their car is parked in their driveway), they have strong reason to believe that they don't know the mundane proposition. But surely, they still count as knowing it.³⁰

In response to this sort of case, Lutz might point to the fact that his principle differs in two ways from the latter principle. First, his principle is put in terms of "learning", which is factive, unlike "having strong reason to believe". Second, his principle concerns learning that a belief does not satisfy a genuine constraint on knowledge, rather than something that one justifiably takes to be a constraint on knowledge. But these two differences seem unlikely to be relevant – it is plausible that if Lutz's principle is true, so is the latter one. Lutz also might also reject the claim that the college student in the case has strong reason to think knowledge requires certainty, or that the student counts as knowing the mundane claim. But there are other kinds of cases, too, in which a thinker can know a claim while having strong evidence that they don't know.³¹

There is another problem with Lutz's principle. Suppose I'm told that I'm in a group of people (say, the people in the room I'm presently in) who have the following feature: Anyone in the group who has a justified true belief about some specific topic (e.g., the weather) is in a Gettier situation and so doesn't have knowledge on that topic. Lutz's principle predicts that this should defeat my justification for my beliefs about the topic. But that is not intuitively correct. Learning that if my belief on the topic is a justified true belief, then it is not knowledge, doesn't give me any reason to think that my belief is untrue or unjustified or otherwise defective. So there doesn't seem to be any rational pressure on me at all to give up the belief or to reduce my confidence in it. So, it seems, Lutz's principle is false.³²

What this suggests is that arguments against moral realism that appeal to a necessary condition on knowledge are not dialectically effective. Even if it were to turn out that one of them can be used to show that robust moral realism is incompatible with moral knowledge, this wouldn't be

³⁰ This is a variant of a case due to Barnett (2021, p. 659).

³¹ One such example is Williamson's (2011) case of the unmarked clock.

³² Why is Lutz's principle intuitively plausible to begin with? A natural diagnosis is that when a thinker learns (or acquires good reason to believe) that one of their beliefs fails to satisfy a constraint on knowledge, this usually comes along with knowledge (or good reason to believe) that the belief may not be true. And that would defeat the justification for the belief.

very worrisome, since it could still be maintained that one has justified true moral beliefs, which seems plenty good enough.

IV. Argument 2: Irrelevant Influences

Let me now turn to the second kind of epistemic argument against robust moral realism. This argument is based on the idea that, assuming moral realism, our moral beliefs were formed in a way that reflected the significant impact of irrelevant factors.³³ In more detail, the argument goes something like this:

- (i) If, under the assumption that some theory is true, a thinker has good reason to believe that one of her beliefs was formed in a way that reflects the significant impact of irrelevant influences, then there is rational pressure on the thinker against accepting the combination of the theory and the belief.
- (ii) Under the assumption that robust moral realism is true, we have good reason to believe that our moral beliefs were formed in a way that reflects the significant impact of irrelevant influences.
- (iii) So there is rational pressure against our accepting the combination of robust moral realism and our moral beliefs.

Here, an irrelevant influence is a factor in the causal history of the belief that is disconnected from the truth of the claim in question.

This kind of argument is one way of trying to make sense of “evolutionary debunking arguments” against robust moral realism, which try to show that robust moral realism is false on the grounds that our moral beliefs (or the underlying motivational tendencies that help to generate them) were shaped by evolution by natural selection, and assuming robust moral realism, evolutionary pressures are disconnected from the moral facts.³⁴ On such an argument, the irrelevant influences are evolutionary (e.g., survival and reproductive) pressures. But there can also be irrelevant influence arguments against robust moral realism that appeal to different kinds of irrelevant influences – for instance, psychological, sociological, or historical processes.

There is an immediate problem with this argument.³⁵ Namely, the central epistemic principle featuring in it, premise (i), is false. Many if not all of our beliefs were formed in a way that reflects the significant impact of irrelevant influences, and this does not count against their justification.

For an everyday example, consider the following case: It’s the first sunny day in a while, and I have a bit of free time, so I decide to go on a walk. I take a fairly random-ish path, and find myself on Ives St., where I happen to notice the newish vegan ice cream place. Out of curiosity, I look at its menu, and learn that in addition to vegan ice cream, it also serves knishes. I count as

³³ See Vavova (2018) for a general discussion of irrelevant influence arguments.

³⁴ See, for example, Street (2006).

³⁵ Some of the points in this section overlap with points made in Schechter (2017). This chapter is, in some ways, an update of that paper.

justified in my belief that the new vegan ice cream place serves knishes, despite the fact that my coming to have the belief reflects the significant impact of irrelevant factors – the weather, the fact I had some free time, the random-ish decisions to turn right and left along my walk, my curiosity about the menu, and so forth.

In response to cases like these, it is tempting to try to restrict or otherwise modify the epistemic principle. I can think of two main strategies for doing so. The first strategy is to say that irrelevant influences are not problematic when they lead to the discovery of new evidence. In the vegan ice cream case, the various irrelevant influences led me to read the menu, which provided me with evidence about what the vegan ice cream places serves. Such evidence “screens off” the impact of the irrelevant factors. According to this line of thought, learning about the significant impact of irrelevant influences on a belief, under the assumption that a theory is true, only generates rational pressure not to accept the combination of the theory and the belief if one is justified in believing that the irrelevant influences did not act via leading one to acquire new evidence that bears on the truth of the belief. Using this refined principle, we can argue that the irrelevant influences that generate our moral beliefs did not yield evidence for our moral beliefs, at least assuming robust moral realism.

This proposal faces several difficulties. One difficulty is that it is not clearly correct to claim that the irrelevant influences (evolutionary, historical, sociological, psychological, etc.) on our moral beliefs did not lead us to acquire evidence about moral matters. One could say that some of our evidence about moral matters comes from our intuitive judgments about specific (actual and hypothetical) cases, and that the irrelevant influences had an impact in generating these intuitions. So the influences did in fact play a role in leading us to acquire evidence about moral matters. A second difficulty with the proposal is that if learning about irrelevant influences can defeat the justification of a belief, it presumably can do so even if the influences led me to acquire new evidence – I might learn that the influences led me to acquire evidence in a biased way or to miscalculate my evidence. A third problem with the proposal is that learning of the significant impact of irrelevant influences can fail to defeat my justification even if the influence did not lead me to acquire new evidence – for instance, the influence might have led me to better evaluate the evidence that I already possessed. For instance, here is a case of that sort: I felt the need to relax and so took a shower, which enabled my mind to wander, which led me to figure out the crucial step in working out a practical or theoretical problem, which led me to a new belief. That I felt the need to relax had a significant impact on my belief but didn’t provide any new evidence bearing on its truth. Nevertheless, learning of this causal history would not defeat my justification for the belief.

The second strategy to restrict the epistemic principle is to distinguish between irrelevant influences and distorting or pernicious influences. On this line of thought, learning about the significant impact of some causal influences on a belief only defeats the justification of the belief if the thinker is justified in believing that the influences had a distorting effect on the belief. One difficulty facing this suggestion is that in many cases, it is a mistake to think of the casual influences as distorting some pre-existing belief. Rather, there would not have been a belief about the matter at all were it not for those influences. For instance, in the case of an

evolutionary debunking argument, it is not the case that were it not for the relevant evolutionary forces, we would have had undistorted moral beliefs. Rather, were it not for the relevant evolutionary forces, we would not have existed at all.

These concerns show that there are difficulties facing irrelevant influence-based arguments against robust moral realism. It is difficult to formulate a plausible version of the epistemic principle featuring in such arguments. But even if some restricted or modified version of the epistemic principle can be formulated that avoids counterexample, there is a still more significant difficulty with the entire line of thought. There is something wrong with the strategy of pointing to specific influences on our moral beliefs and arguing that such influences led us astray. There is a deeper intuitive problem facing robust moral realism – not that our moral beliefs reflect the impact of irrelevant influences, but that they fail to reflect the impact of relevant ones. Given robust moral realism, we don't need to learn about the causal history of our moral beliefs to locate a problem. We know that, since moral facts are acausal, they cannot play a role in generating our moral beliefs. And that is a deeper and intuitively more worrisome problem.

V. Argument 3: Harman's Challenge

There is a familiar argument against moral realism based on the thought that, assuming robust moral realism, the moral facts don't play a role in generating our moral beliefs. This argument is due to Harman and is sometimes called "Harman's challenge".³⁶ One way to formulate this argument is as follows:

- (i) If, under the assumption that some theory is true, the truths about a domain do not play a role in the best explanation of our beliefs about the domain, then there is rational pressure against accepting the combination of the theory and our beliefs about the domain.
- (ii) Assuming robust moral realism, the moral truths do not play a role in the best explanation of our moral beliefs.
- (iii) So, there is rational pressure against accepting the combination of robust moral realism and our moral beliefs.

(Harman's discussion focuses on whether the moral truths play a role in explaining what he calls our "moral observations", but the change to moral beliefs is harmless. Harman also targets a wider class of metaethical theories than robust realist ones, but the challenge seems strongest when presented against robust moral realism.)

The idea behind the second premise is that if robust moral realism is true, then the moral truths are acausal, and so can't play a role in explaining our moral beliefs. What is the idea behind the first premise?

There are three very different lines of thought that could support an epistemic principle like the first premise. The first is that a belief counts as knowledge only if the relevant truth plays a role

³⁶ See Harman (1977, pp. 3-10).

in the explanation of the belief. If there's also a principle linking necessary conditions on knowledge with justification, such as Lutz's principle, then we can derive something like the first premise. I've already discussed this kind of argument above and argued that it runs into serious problems.

The second line of thought motivating something like the first premise concerns reliability. If the truths about a domain do not play a role in explaining our beliefs about the domain, it is mysterious how it is that we believe the truths about the domain. In the absence of an explanation of our reliability, there is pressure against accepting the combination of our beliefs about the domain and the background beliefs that help to generate the mystery. (Those background beliefs include (i) Nonnaturalism, which rules out an explanation on which the moral facts explain our moral beliefs and (ii) Independence, which rules out an explanation on which our moral beliefs explain the moral facts.) I think that something like this line of thought is a challenging one for the robust moral realist, and I will discuss it in detail later in this chapter.

The third line of thought seems closest to what Harman himself seems to have had in mind. The idea is that for us to have reason to believe in the existence of some objects or the exemplification of some properties or relations, then commitment to those objects, properties, and relations had better do some explanatory work for us.³⁷ The objects, properties, or relations had better "earn their keep". This is a kind of Ockamite consideration – we shouldn't be committed to an ontology or ideology that is more expansive than necessary. Learning that (purported) objects, properties, or relations don't earn their keep defeats any justification we may have for being committed to them.

Given this methodological principle, the argument, then, goes like this: To justifiably retain a commitment to moral truths, properties, and relations, such truths, properties, and relations had better do some explanatory work for us. According to robust moral realism, moral truths, properties, and relations are not identical with, reducible to, or grounded in natural truths, properties, and relations. So our commitment to moral truths, properties, and relations cannot be defended by appealing to the explanatory roles that natural truths, properties, and relations play in our overall view of the world. The only explanatory work that moral truths, properties, and relations apparently can do is in helping to explain how it is we have our moral beliefs. But given robust moral realism, it seems that moral truths, properties, and relations cannot play a role in explaining our moral beliefs either. So there is pressure to give up our commitment to moral truths, properties, and relations, at least as the robust moral realist conceives of them.

This argument has some force. The general principle that we shouldn't be committed to an ontology or ideology that is more expansive than necessary is plausible. But Harman's challenge is less powerful than it is sometimes treated as being.

Harman's line of argument does, I think, show that we should not take robust moral realism to be the default view in metaethics. A line of argument that I've sometimes heard is that the pre-

³⁷ This claim presumably should be restricted to positive properties and relations. The property of being a non-unicorn, for example, presumably doesn't need to do work for us in order for us to have reason to think it is exemplified. I'll leave this restriction implicit in what follows.

theoretically most natural view is moral realism, so we should treat moral realism as the default view and only give it up if we have compelling reason to do so. What Harman's challenge, in effect, points out is that there is a cost in accepting moral truths, properties, and relations as the robust moral realist conceives of them. Once we notice that, it is hard to maintain that robust moral realism is the default position.³⁸ To justifiably endorse robust moral realism requires showing how moral truths, properties, and relations (as the robust moral realist conceives of them) earn their keep in our overall picture of the world.³⁹

However, Harman's line of argument seems much too quick in concluding that if moral truths (as conceived by the robust moral realist) do not play a role in explaining our moral beliefs, then they cannot do any work for us, and so we should give up our commitment to them. There are other ways that these truths can earn their keep beyond playing a role in explaining our beliefs. Versions of this point have been made by Sayre-McCord and by Enoch.⁴⁰ Sayre-McCord considers the view that our general explanatory practices rely upon value judgments to legitimate them, value judgments that may also legitimate our moral judgments. If that view is right, Sayre-McCord argues, then we may justifiably retain our commitment to moral truths as the robust moral realist conceives of them. Enoch argues that we are entitled to accept normative truths as the robust realist conceives of them, because a commitment to such truths is needed to make sense of our deliberative practices. Deliberation only makes sense if it aims at figuring out the independent normative truths. And once we allow normative truths as the robust realist conceives of them into our worldview, it is not costly to also allow specifically moral truths into our worldview.

Sayre-McCord and Enoch are right that we can respond to Harman's challenge by showing that moral truths, properties, and relations can earn their keep in ways other than by helping to explain our moral beliefs. But the point generalizes still further than they explicitly note. The basic idea behind Harman's challenge is that we shouldn't be committed to a more expansive ontology or ideology than is needed – or better, than we have good reason to accept. One way that a commitment to objects, properties, and relations can be defended is by showing that the objects, properties, and relations do explanatory work for us, for instance in helping to explain our beliefs. But the commitment to objects, properties, and relations can be defended in other ways, too. If we have any kind of good reason whatsoever to accept the objects, properties, and relations, that shows that the objects, properties, and relations can earn their keep. So all one needs to do to answer Harman's challenge is to show that we have some kind of positive reason

³⁸ Woods (2018) similarly argues that Harman's challenge should be seen as a "burden-shifting" argument.

³⁹ This understanding of Harman's challenge has several points of contact with Wright's discussion of Harman in Wright (1992, ch. 5.v). Wright argues that moral realism faces trouble if moral facts do not have a "wide cosmological role" in the sense that they cannot play a role in explaining anything other than our moral judgments and beliefs. There are two principal differences between this line of thought and the one I'm taking from Harman. First, Wright does not count explanations of our moral judgments and beliefs as a way that moral facts can earn their keep. Second, Wright identifies the distinction between domains in which the facts can play an explanatory role and domains in which they cannot with a distinction between realism and anti-realism. (On Wright's view, there are several realism/anti-realism distinctions.) By contrast, on my rendering of Harman's argument, that moral facts realistically construed cannot play an explanatory role provides reason to reject moral realism, which is understood as a distinct thesis.

⁴⁰ See Sayre-McCord (1988, pp. 278-80) and Enoch (2011, ch. 3).

to accept moral truths, properties, and relations as the robust moral realist conceives of them. Any good reason will do. That's why Harman's challenge is less powerful than it may at first appear.

VI. Argument 4: The Argument from Improbability

In discussing Harman's challenge, I put off discussion of the suggestion that the problem facing robust moral realism has to do with our reliability about moral matters. Let me turn to that suggestion now. The thought is that if robust moral realism is true, then it is mysterious how it is that our beliefs about morality are reliable, but since we are in fact reliable, we should give up moral realism. A version of this argument has recently been presented by Warren.⁴¹ Warren's argument is not aimed against robust moral realism in particular, but against what he calls "non-causal realism" about any domain. Somewhat simplifying and reordering Warren's presentation, and focusing on the specific case of morality, we can put the argument schematically as follows:

- (i) If robust moral realism is true, then our moral beliefs don't have feature F.
- (ii) If our moral beliefs don't have feature F, then they are unlikely to be reliable.
- (iii) Our moral beliefs are reliable.
- (iv) So robust moral realism is not true.

Warren uses the safety and sensitivity of our beliefs as feature F, but similar arguments can be made for other candidate features, such as standing in a causal or explanatory relation with the relevant facts or being non-accidentally true. This family of arguments thus resembles the first family of arguments I considered, involving features that are proposed as necessary conditions on knowledge, but without the tie to knowledge. Instead, the tie is between these features and reliability. (As before, beliefs about a domain count as reliable if they are true far more often than chance alone would predict.)

Notice that this argument, as I formulated it, is not deductively valid – the second premise concerns whether our moral beliefs are likely to be reliable and the third concerns whether they are in fact reliable. This is not, by itself, a problem with the argument. Plenty of good philosophical arguments are not deductively valid but nevertheless have significant epistemic force.

However, there is a difficulty with this family of arguments. There need not be much of a clash between the claim that such-and-such is unlikely to be true and such-and-such is in fact true. For example, suppose I flip a coin twenty times and the following sequence of heads and tails results: HHTTTHTTHTTHTTHTHHHTT. That this specific sequence resulted is highly unlikely, at least given my understanding about how coin flips work, etc. Indeed, the probability of that specific sequence is approximately one in a million. Nevertheless, there is no epistemic pressure on me whatsoever to give up my theory of how coin flips work based on the occurrence of this unlikely sequence of flips.

⁴¹ Warren (2017).

Warren responds to a version of this objection.⁴² He argues that in many cases of improbable events, such as the case of Lucky Larry winning the lottery, we have plenty of independent evidence for the occurrence of the improbable event. The difference is that in the case of our moral beliefs, we don't have independent evidence in support of their reliability. That's why the improbability of our reliability poses a problem but Luck Larry's lottery win does not.

I'm not convinced by this response. A first point is that we can think of cases in which we don't have much independent evidence of the occurrence of an improbable event. Consider a coin flip case in which I didn't flip the coin but my friend did, and he reports to me a specific sequence of heads and tails. Suppose my friend is not a very reliable testifier. The fact that what my friend reports is an unlikely occurrence isn't germane to whether or not I should believe my friend's report – I should believe it roughly as much or as little as I would believe his report of a not-unlikely occurrence, say that he had orange juice with breakfast.⁴³ So if there is a problem due to the lack of independent evidence, this has nothing to do with unlikelihood.

A second point is that the response misidentifies the target of the argument. Recall that Warren's argument is an argument against robust moral realism, not the reliability of our moral beliefs. So whether or not we have independent evidence of our moral reliability seems irrelevant to the argument. Indeed, consider a situation in which I have strong independent evidence for my moral reliability. Suppose, for instance, the epistemology oracle appears before me and tells me that my moral beliefs are highly reliable. Suppose the oracle is nearly always correct about what she pronounces upon (and I know this based on her excellent track record). I now have strong independent evidence of my moral reliability. But this doesn't seem to answer the intuitive problem for robust moral realism. The fact that I am reliable still seems mysterious, given robust moral realism, whether or not I have strong independent evidence in support of it.

What this suggests is that Warren's argument doesn't quite capture the problem for robust moral realism. The unlikelihood of an event isn't enough to generate an epistemic problem. The natural suggestion to make is that our reliability about morality isn't merely an unlikely fact (given robust moral realism) but is a striking fact that "calls out for explanation". The trouble for robust moral realism is that there doesn't seem to be such an explanation compatible with robust moral realism. And that is what generates the intuitive problem.

VII. Argument 5: The Reliability Challenge

The fifth and final kind of argument I'll discuss is based on the idea that robust moral realism runs into difficulty because it is not compatible with there being an explanation of how it is we

⁴² Warren (2017, pp. 1648-9).

⁴³ There are two caveats to make here. First, if the reason my friend is unreliable is that he has a dodgy memory, perhaps I should be less confident in his report on the coin toss than on the orange juice, simply because a sequence of twenty tosses is harder to remember. So we should understand the case to be one in which the reason for his unreliability is different. Second, if my friend tells me a highly implausible story (a fish tale), I should have less confidence in that report than that he had orange juice at breakfast. But this is not solely due to the low probability of the story. Rather, it is due to the fact that my evidence suggests that the best explanation of my friend's recounting the story is not that the story is true but that he made it up.

have reliable moral beliefs. This argument is often called “the reliability challenge” against moral realism. It is analogous to an argument made by Field against mathematical Platonism which is often called “the Benacerraf-Field problem”.⁴⁴

The argument depends on the idea that some facts are striking or “explanatorily urgent”.⁴⁵ Whether a fact is striking doesn’t depend only on whether it is unlikely. For instance, the fact that a coin was tossed a hundred times and resulted in a random-seeming pattern of heads and tails is not striking. The fact that a coin was tossed a hundred times and resulted in heads every time is striking. So is the fact that a coin was tossed a hundred times and resulted in the first few lines of *Hamlet* in Morse code. This is so despite the fact that every sequence of heads and tails is equally likely for a fair coin.

Given the notion of a striking fact, the reliability challenge can be put as follows:

- (i) It is a cost of a theory if it is incompatible with there being a plausible explanation of some striking fact.
- (ii) According to robust moral realism, our moral beliefs are reliable in the sense that they are true far more often than chance alone would predict.
- (iii) Given robust moral realism, the fact that our moral beliefs are reliable is a striking fact.
- (iv) Robust moral realism is incompatible with there being a plausible explanation of the fact that our moral beliefs are reliable.
- (v) This is a cost to accepting robust moral realism.

The first premise is a general principle governing theory choice. It is a cost of a theory if the theory requires us to treat some striking fact as inexplicable. In other words, there is at least some epistemic pressure to not accept the theory. (This is compatible with the overall balance of considerations epistemically supporting the acceptance of the theory.)

The second premise of the argument follows from the definition of robust moral realism provided above. (If one doesn’t want to include our reliability about morality as a component of moral realism, the argument could be refashioned to conclude that there is a cost to accepting the combination of moral realism with the claim that our moral beliefs are reliable.)

In support of the third premise, it would be nice to provide a general theory of when a fact is striking. Unfortunately, I do not have such a theory to provide.⁴⁶ Nevertheless, it is plausible that the reliability of our moral beliefs, if true, is a striking fact. Part of the reason the fact is striking is that, assuming robust moral realism, it is intelligible and genuinely possible that our moral beliefs are largely false, but our moral beliefs are in fact largely true, or at least are true

⁴⁴ See the Introduction and title essay in Field (1989) for the Benacerraf-Field problem for mathematical Platonism. See Enoch (2011, ch. 7) for the analogous problem for robust moral realism. See Schechter (2010) for the analogous problem for realism about logic, as well as a brief discussion of the cases of mathematics, modality, and morality.

⁴⁵ I borrow the phrase “explanatory urgency” from White (2005).

⁴⁶ There are plausible heuristics for when a fact is striking: (i) A general fact is striking if it describes a simple pattern – e.g., “all heads”; (ii) A particular fact is striking if it violates an otherwise well-confirmed generalization – e.g., “the coin has started to float in mid-air”; (iii) A fact is striking if there is a salient (but perhaps improbable) theory that would explain it – e.g., “someone is intentionally controlling the coin with their mind”.

considerably more reliable than chance alone would predict. How striking our reliability about morality is depends on just how reliable our beliefs are – the more reliable we are, the more striking our reliability is. It also depends on who the “we” is in the statement of reliability – the bigger and more diverse the group, the more striking our reliability is.

The fourth premise depends on a particular understanding of what exactly calls out for explanation. There are different explanatory tasks one could undertake in explaining our reliability about morality. One question is: How do our belief-forming methods work so as to generate reliable moral beliefs? A second question is: How is it that we have belief-forming methods that generate reliable moral beliefs? In earlier work, I called the first question “the operational question” and the second question “the etiological question”.⁴⁷

For the purpose of formulating the reliability challenge, the second question is the relevant one. There are plausible candidate answers to the first. For example, here is one possibility: We have a cognitive mechanism for making moral judgments that involves implicit commitment to general moral principles that are more-or-less correct.⁴⁸ (One such principle could be that causing gratuitous suffering is wrong). Our mechanisms of concept application make use of these principles as well as our beliefs about the world to yield moral verdicts on actual and hypothetical cases. This is what generates some of our moral judgments. We then use our general reasoning abilities to arrive at further moral beliefs. Since the general moral principles embedded in our cognitive mechanism are more-or-less correct, this explains how it is that our belief-forming methods work so as to yield reliable moral beliefs.

My claim here is not that this explanation is correct. Rather, it is that this explanation is a plausible explanation compatible with robust moral realism. More importantly, it is a proof of concept that there are plausible explanations to be had that are compatible with robust moral realism.

Given robust moral realism, the second explanatory task is, at least *prima facie*, much more challenging. Suppose the above answer to the operational question is correct. How is it that we have a cognitive mechanism for making moral judgements that involves the implicit commitment to more-or-less correct moral principles? Given robust moral realism, the answer cannot be that the moral truths caused us to have these commitments – moral truths are acausal and so cannot cause anything. Given robust moral realism, the answer cannot be that our moral commitments explain the moral truths – the fundamental moral truths are mind-independent and so cannot be explained by our commitments. So it appears much more difficult to provide a plausible explanation of how it is we have a cognitive mechanism that yields reliable moral beliefs. It is this explanatory task that the reliability challenge should be understood to focus on.

Notice that the conclusion of the argument is not that robust moral realism is mistaken. Rather, it is that there is a cost to accepting robust moral realism. Even if the conclusion is correct, it could

⁴⁷ Schechter (2010).

⁴⁸ This is what Peacocke (1998) calls an “implicit conception”.

be that this cost is outweighed by the considerations supporting the theory. It could be that robust moral realism is superior to its alternatives.

What should we make of this argument? In my view, it is a serious concern facing robust moral realism. The premises are all plausible. And the argument seems to get to the heart of what is disquieting about robust moral realism, at least from an epistemic point of view. So this argument is what I take to be the central epistemic argument against robust moral realism.

There are responses in the literature to the reliability challenge that aim to show that it is not as pressing as it appears. One such response is to argue that the epistemic principle (i.e., that it is a cost of a theory if it is incompatible with there being a plausible explanation of a striking fact) is unmotivated. Relatedly, it could be argued that it is unmotivated to think there is a distinction between striking and non-striking facts at all. Baras argues for this latter conclusion by arguing, among other things, that most of the examples that have been taken to motivate the idea that some facts are striking can be accounted for using purely Bayesian reasoning, and so the notion of strikingness does no explanatory work for us.⁴⁹

A different kind of response is to argue that the reliability challenge can be answered with a trivial kind of explanation: We first explain how we came to have the moral beliefs that we do, appealing to some combination of evolution, history, sociology, and psychology. We then explain why the moral truths are the way they are, appealing to the fundamental moral principles as well as contingent claims about the way the world actually is. (The fundamental moral principles themselves either have no explanation, are explained by the fact that they are necessary truths, or are explained by the fact that they follow from the nature of moral properties and relations.) Finally, we put both explanations together and argue that there is a causal story of why we believe such-and-such claims, and such-and-such claims are in fact largely true, so this explains how it is we believe moral truths.⁵⁰

These are interesting responses. I don't have the space to adjudicate them in detail here.⁵¹ But I do want to claim that intuitively, some facts seem to require explanation more than others. For example, flipping a coin many times and seeing a sequence of all heads requires explanation in a way that a random-seeming sequence of heads and tails does not, despite the fact that the two sequences are equiprobable (for a fair coin). And, intuitively, an "explanation" of how we have reliable moral beliefs that merely points to an explanation of how we came to believe the claims that we do and then says that these claims are largely true does not seem like a genuine explanation of our reliability. Of course, it would be nice to have a worked-out account of what it

⁴⁹ Baras (2022).

⁵⁰ This is close to what Linnebo (2006) calls "the boring explanation" in the context of explaining our reliability about mathematics. Also see Lewis (1986, pp. 114-5), who considers (and rejects) the view that explanations of our reliability are needed only for contingent subject matters. Also see Clarke-Doane (2017), who argues that the only legitimate explanatory demands concerning our reliability about a domain are to show that our beliefs about the domain are safe and to show that our beliefs about the domain are sensitive. Since the fundamental moral principles are necessary truths, our beliefs in such truths are automatically sensitive. An evolutionary, historical, sociological, or psychological explanation suffices to show that our beliefs are safe.

⁵¹ See Schechter (2010; 2018) for discussion of the latter kind of response, in the context of the reliability challenge for logic. I hope to address the first kind of response in future work.

would take to explain our reliability. But this follows from a more general point: It would be nice to have a general account of explanation. Formulating such an account has proved difficult, even though there is generally broad agreement about when some purported explanation does and when it does not count as a genuine explanation.

My view is that the best response to the reliability challenge for moral realism is not to reject the challenge or to provide a trivial kind of explanation, but to respond to the argument by providing a substantive explanation of our reliability. In carrying out this task, a question arises: Is it legitimate for the contents of our moral beliefs to play a role in the explanation? Can we, for instance, appeal to the claim that causing gratuitous harm is wrong in trying to formulate an explanation of why it is we correctly believe that claim (or other moral claims)? Or would that be problematically question-begging?⁵²

Given the structure of the reliability challenge, there is no problem with appealing to our substantive moral views in responding to the challenge. The challenge is an explanatory one – what is the explanation of our reliability? In general, in providing an explanation of some phenomenon, it is perfectly acceptable to make use of our views about the world.⁵³ The challenge tries to point to an internal tension in the moral realist’s package of views. In responding to the claim that one’s view has an internal tension, it is perfectly acceptable to show how the tension can be resolved using materials provided by the view.

What would be objectionably question-begging is if someone argued as follows: Our moral beliefs are reliable. This is a striking fact. So there must be an explanation of our reliability. So there is no problem facing the realist. To answer the challenge, it is not enough to conclude that there is an explanation. Rather, we need to exhibit such an explanation, a sketch of such an explanation, or some kind of proof of concept that such an explanation can ultimately be provided.

What could an explanation of our reliability about morality look like, if it is to be compatible with moral realism? The most promising suggestion is to provide a third-factor explanation. The idea is that there is some factor that both causally explains our moral beliefs and metaphysically grounds the moral truths, thus ensuring some kind of match between them. There are several candidate third-factor explanations that appear in the literature.⁵⁴ While it is beyond the scope of this chapter to evaluate them, I will say that I am not satisfied with the third-factor explanations that have so far been developed. If the robust moral realist is going to directly respond to the reliability challenge, this is where they should focus their attention. If no such account is workable, then the best they can do is to argue that the reliability challenge poses a cost for robust moral realism, but it is a cost which they think they should bear.

⁵² See Street (2008, pp. 216-7) for a version of this concern, aimed at responses to the “Darwinian dilemma” for moral realism.

⁵³ What is prohibited is to explain a phenomenon in terms of itself or in terms of anything that is itself (in part) explained by the phenomenon.

⁵⁴ See, for instance, Wielenberg (2010), Enoch (2011, pp. 168-74), and Sarksaune (2011).

VIII. Conclusion

Let's take stock. There are many different families of argument with an epistemic flavor that have been presented against robust moral realism. What I've claimed here is that many of these arguments are not very pressing. Arguments appealing to (purported) necessary conditions on knowledge face several difficulties, most notably that their conclusion is not very threatening. Arguments appealing to irrelevant influences face difficulty in formulating a plausible epistemic principle and also seem not to capture the intuitive epistemic problem facing robust moral realism. Harman's challenge has some force, but only ends up showing that robust moral realism is not the default metaethical position. The argument from improbability is not very threatening, since the occurrence of an improbable event, by itself, does not pose an epistemic problem for a view. What does seem to be the most challenging problem is the reliability challenge. It makes use of plausible premises. It also captures the intuitive epistemic problem with robust moral realism. (And the same is true for our beliefs about other domains, such as logic, mathematics, modality, and so forth.) So this is the problem on which robust moral realists should focus their attention.⁵⁵

References

- Balaguer, Mark. (1998) *Platonism and Anti-Platonism in Mathematics*. Oxford: Oxford University Press.
- Baras, Dan. (2022) *Calling for Explanation*. Oxford: Oxford University Press.
- Barnett, Zach. (2021) "Rational Moral Ignorance," *Philosophy and Phenomenological Research* 102: 645-64.
- Benacerraf, Paul. (1973) "Mathematical Truth," *The Journal of Philosophy* 70: 661-79.
- Bengson, John. (2015) "Grasping the Third Realm," *Oxford Studies in Epistemology* 5: 1-38.
- Blackburn, Simon. (1984) *Spreading the Word*. Oxford: Oxford University Press.
- Burgess, John and Rosen, Gideon. (1997) *A Subject with No Object*. Oxford: Oxford University Press.
- Chisholm, Roderick. (1966) *Theory of Knowledge*. Englewood Cliffs, NJ: Prentice-Hall.
- Clarke-Doane, Justin. (2017) "What is the Benacerraf Problem?," in Fabrice Pataut (ed.), *New Perspectives on the Philosophy of Paul Benacerraf: Truth, Objects, Infinity*. Cham: Springer, pp. 17-43.
- Cuneo, Terence and Shafer-Landau, Russ. (2014) "The Moral Fixed Points: New Directions for Moral Nonnaturalism," *Philosophical Studies* 171: 399-443.

⁵⁵ Thanks to David Christensen for helpful discussion. Thanks also to the two editors of this volume for numerous helpful suggestions.

- Enoch, David. (2011) *Taking Morality Seriously*. Oxford: Oxford University Press.
- Feldman, Richard. (2003) *Epistemology*. Upper Saddle River: Prentice Hall.
- Field, Hartry. (1989) *Realism, Mathematics, and Modality*. New York: Blackwell Publishers.
- Gettier, Edmund. (1963) "Is Justified True Belief Knowledge?," *Analysis* 23: 121-23.
- Goldman, Alvin. (1967) "A Causal Theory of Knowing", *Journal of Philosophy* 64: 357-72.
- Goldman, Alvin. (1976) "Discrimination and Perceptual Knowledge," *The Journal of Philosophy* 73: 771-791.
- Grundmann, Thomas. (2018) "Saving Safety from Counterexamples," *Synthese* 197: 5161-85.
- Harman, Gilbert. (1977) *The Nature of Morality*, Oxford: Oxford University Press.
- Lewis, David. (1986) *On the Plurality of Worlds*. Oxford: Blackwell Publishers.
- Lewis, David. (1989) "Dispositional Theories of Value," *Proceedings of the Aristotelian Society, Supplementary Volume*, 63: 113-137.
- Linnebo, Øystein. (2006) "Epistemological Challenges to Mathematical Platonism," *Philosophical Studies* 129: 545-574.
- Lutz, Matt. (2020) "The Reliability Challenge in Moral Epistemology," *Oxford Studies in Metaethics* 15: 284-308.
- Mackie, John. (1977) *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin.
- Nozick, Robert. (1981) *Philosophical Explanations*. Cambridge, Mass.: Harvard University Press.
- Peacocke, Christopher. (1998) "Implicit Conceptions, Understanding and Rationality," *Philosophical Issues*, 9, *Concepts*, 43-88.
- Plunkett, David. (2020) "Conceptual Truths, Evolution, and Reliability about Authoritative Normativity," *Jurisprudence* 11: 169-212.
- Pritchard, Duncan. (2005) *Epistemic Luck*. Oxford: Oxford University Press.
- Sarksaune, Knut. (2011) "Darwin and Moral Realism: Survival of the Fittest," *Philosophical Studies* 152: 229-43.
- Sayre-McCord, Geoffrey. (1988) "Moral Theory and Explanatory Impotence" in his *Essays on Moral Realism*, Ithaca: Cornell University Press, pp. 256-81.
- Schechter, Joshua. (2010) "The Reliability Challenge and the Epistemology of Logic," *Philosophical Perspectives*, 24, *Epistemology*, 437-64.

- Schechter, Joshua. (2017) "Explanatory Challenges in Metaethics," in Tristram McPherson and David Plunkett (eds.), *The Routledge Handbook of Metaethics*. New York: Routledge, pp 443-58.
- Schechter, Joshua (2018) "Is There a Reliability Challenge for Logic?," *Philosophical Issues* 28, *Philosophy of Logic and Inferential Reasoning*, 325-47.
- Shafer-Landau, Russ. (2003) *Moral Realism: A Defense*. Oxford: Oxford University Press.
- Sosa, Ernest. (1999) "How to Defeat Opposition to Moore," *Philosophical Perspectives*, 13, *Epistemology*, 141-53.
- Stevenson, Charles. (1937) "The Emotive Meaning of Ethical Terms," *Mind* 46: 14-31.
- Street, Sharon. (2006) "A Darwinian Dilemma for Realist Theories of Value," *Philosophical Studies* 127:109-66.
- Street, Sharon. (2008) "Reply to Copp: Naturalism, Normativity, and the Varieties of Realism Worth Worrying About," *Philosophical Issues*, 18, *Interdisciplinary Core Philosophy*, 207-28.
- Vavova, Katia. (2018) "Irrelevant Influences," *Philosophy and Phenomenological Research* 96: 134-52.
- Warren, Jared. (2017) "Epistemology versus Non-Causal Realism," *Synthese* 194: 1643-62.
- Wedgwood, Ralph. (2001) "Conceptual Role Semantics for Moral Terms," *The Philosophical Review* 110: 1-30.
- White, Roger. (2005) "Explanation as a Guide to Induction," *Philosophers' Imprint* 5: 1-29.
- Wielenberg, Erik. (2010) "On the Evolutionary Debunking of Morality," *Ethics* 120: 441-64.
- Williamson, Timothy. (2011) "Improbable Knowing," in Trent Dougherty (ed.), *Evidentialism and its Discontents*. Oxford: Oxford University Press.
- Woods, Jack. (2018) "Mathematics, Morality, and Self-Effacement," *Noûs* 52: 47-68.
- Wright, Crispin. (1992) *Truth and Objectivity*, Cambridge, Mass.: Harvard University Press.