



<http://www.diva-portal.org>

Preprint

This is the submitted version of a chapter published in *Ethics and Explanation*.

Citation for the original published chapter:

Tersman, F. (2016)

Explaining the Reliability of Moral Beliefs.

In: Neil Sinclair and Uri Leibowitz (ed.), *Ethics and Explanation* (pp. 37-57). Oxford University Press

N.B. When citing this work, cite the original published chapter.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-283616>

This is the pre-peer-reviewed version of the article. It has been published in final form in *Ethics and Explanation* (eds. Neil Sinclair and Uri Leibowitz), Oxford University Press, 2016, 37-57.

Explaining the Reliability of Moral Beliefs

Folke Tersman (Uppsala University)

1. Introduction

Moral realists believe that there are moral facts—facts about what is right and wrong, good and bad, etc.—and that these facts are radically mind-independent.¹ A common objection to that position is that it cannot adequately account for our having *knowledge* about what is right and wrong, good and bad, and so on. This worry is often supposed to be reinforced by an evolutionary account of the origins of our moral beliefs.² The idea is that a realist cannot reconcile such an account with the supposition that our moral beliefs are sensitive to their truth (or to the truth of their contents) in a way that is required in order for them to be reliable and justified and thus to count as knowledge. The evolutionary approach is therefore taken to commit a realist to an untenable form of skepticism.

There are a number of obvious ways of responding to this challenge. One is to question the evolutionary account, for example by stressing that we presently lack a sufficiently well developed version that is warranted by the evidence. Another is to reject the epistemic views that underlie the challenge, such as the view that a belief is justified only if it has the kind of reliability the evolutionary account of our moral beliefs is held to undermine. However, I shall focus on a third strategy; a strategy that neither questions the evolutionary approach nor the view that justification requires reliability. The aim of the third strategy is to argue that the evolutionary account does not after all exclude that our moral beliefs are reliable, even given

¹ See Enoch (2011) and Shafer-Landau (2003) for two prominent recent attempts to defend a realist view. Note that it is primarily non-naturalist versions of realism that are held to be vulnerable to the challenge discussed in this paper.

² For some widely discussed arguments along these lines, see Joyce (2006a), (2006b), and (2013) and Street (2006).

a realist view, by invoking substantive moral claims. This type of response has been pursued by, among others, David Enoch.³

The availability of this strategy might lead one to worry about the possibility of settling the debate about the meta-ethical implications of an evolutionary account of our moral beliefs on a neutral ground. On the one hand, anti-realists can complain that the strategy in question is illegitimate since it invokes beliefs of precisely the kind whose credentials the challenge is meant to undermine. On the other hand, a realist can insist that, by denying her the option of invoking moral beliefs, it is the anti-realist who is begging the question. For such a denial is plausible only under the assumption that the realist cannot legitimately regard her moral beliefs as being justified, which is exactly the conclusion the anti-realist wants to establish. Can we move forward in this situation? The aim of this paper is to argue that we can, and to make certain suggestions about how the discussion about the meta-ethical implications of the evolutionary account should proceed. I shall do so by illustrating that an anti-realist may offer alternative reasons for doubting that the target strategy will work; considerations whose significance and plausibility *can* be given a neutral defense. This allows an anti-realist to concede that the strategy is in principle legitimate and still argue that it ultimately will not help the realist.

The plan is as follows. The version of the challenge against realism that I shall focus on is the one developed by Sharon Street. I give a rough account of Street's argument in section 2. In section 3, I discuss an interesting recent attempt to respond to challenges such as Street's that differs from Enoch's and is due to Justin Clark-Doane.⁴ I am going to argue that, although Clark-Doane's proposal is highly thought provoking, it ultimately fails. However, seeing how it fails enables us to clarify the challenge and to consider in more detail what is required by a satisfactory response. In section 4, I reconstruct Enoch's proposal. In section 5, I start to discuss the types of criticism that are available to an anti-realist even if she concedes that the invocation of moral claims is in principle legitimate. I do so by stating two conditions that a satisfactory instance of the strategy arguably must satisfy. In section 6, I illustrate how

³ See Enoch (2010). For related suggestions, see Brosnan (2011), Schafer (2010) and Skarsaune (2011).

⁴ It is developed in Clark-Doane's contribution to this volume ("Genealogy and Reliability") and in the paper "What is the Benacerraf Problem?", to be published in the volume *New Perspectives on the Philosophy of Paul Benacerraf: Truth, Objects, and Infinity* (edited by Fabrice Pataut). A penultimate draft is available on Clark-Doane's web site.

one can use the conditions to evaluate particular instances of the strategy, by specifically focusing on Enoch's version. In section 7, finally, some concluding remarks are made.

2. *Street on evolution and ethics*

Moral realism is the view that there are moral truths that exist independently of our attitudes and beliefs in relevant ways. The target of Street's argument is not the thesis that there are moral truths but rather the assumption about mind-independence.⁵ Street's own view ("Humean constructivism") disagrees with realism on that point and she believes that it thereby avoids the skeptical worries she thinks realism is vulnerable to.⁶

To bring out how an evolutionary account of our moral beliefs is supposed to generate skeptical worries for a realist, Street begins by noting that the realist may have different views about the relation that holds between the forces that, according to such an account, have shaped the contents of our moral beliefs and the moral facts that the realist posits. The realist could either think that there *is* a relation or deny this and instead hold that the forces in question have "nothing whatsoever to do with evaluative truth".⁷ I shall call the latter claim "the no-relation assumption".

According to Street, if a realist chooses the first option and insists that there is a relation then she commits herself to a version of the evolutionary approach that Street calls "the tracking account". The evolutionary approach entails that we are disposed to have certain moral beliefs because the possession of that disposition gives us (or gave our ancestors) a selective advantage. According to the tracking account, they give us this advantage because the beliefs in question are often true. The problem with the tracking account, according to Street, is that there is a superior competitor, namely what she calls "the adaptive link account". According to the adaptive link account, the selective advantage provided by our moral beliefs is to be explained in terms of the behavior they give rise to rather than in terms of their truth. Why are we disposed to believe that it is wrong to be a free-rider? Roughly, on the adaptive link account, because the possession of that belief leads us to avoid free-riding,

⁵ Street (2006). Note that, although Street primarily focuses on non-naturalist forms of realism she also addresses naturalist versions.

⁶ This assessment can in fact be questioned. See Tropman (2013) for a discussion.

⁷ Street (2006, 121). Street refers to "evaluative truth" rather than "moral truth" (and to "normative realism" rather than "moral realism"). I trust, however, that this merely terminological difference will not cause any problems in what follows.

which in turn helps us to avoid many damaging collective dilemmas. Unlike the tracking account, this version of the evolutionary account is neutral towards moral issues and does not posit any moral facts.

Street thinks that the adaptive link account can be shown to be superior to the tracking account on purely scientific grounds, partly because it is simpler (since it doesn't posit any moral facts). A realist should therefore reject the tracking account. But, according to Street, if she rejects the tracking account then she is committed to the no-relation assumption. And this in turn means that she must concede that the Darwinian forces have pushed the contents of our moral beliefs in a certain direction independently of whether the truth lies in that direction. It *could* of course be the case that the truth does lie in that direction, but if that is so, and if some of our moral beliefs are indeed true, this is the result of a pure coincidence or fluke. In such a case, relying on them in trying to determine the truth about what is right and wrong would, according to Street, be like “setting out for Bermuda and letting the course of your boat be determined by the wind and tides”.⁸ Thus, either way—regardless of what view a realist takes on the relation between the Darwinian forces and the moral facts she posits—she ends up with a problematic conclusion. If she insists that there is a relation, then she is committed to the independently implausible tracking account. If she denies a relation, she is instead committed to a form of skepticism.

Street's challenge is sometimes classified as a “debunking argument”. Debunking arguments differ from other objections through being indirect. They try to undermine the target position by undermining the case that could be developed in *favor* of it rather than by offering considerations that are directly inconsistent with it. Street argues that, given realism and an evolutionary account of our moral beliefs, our moral beliefs are groundless. This is presumably why it is thought appropriate to label her challenge debunking.⁹

⁸ Street (2006, 117). Richard Joyce reasons similarly. He stresses that natural selection is “a process for which practical success rather than accuracy is the *summum bonum*” and that an evolutionary account of our moral beliefs therefore suggests that they “are products of a process that is entirely independent of their truth, which forces the recognition that we have no grounds one way or the other for maintaining these beliefs” (2006b, 135, see also 2006a).

⁹ In fact, Street suggests that, given a realist construal of the nature of the facts posited by our moral beliefs, the adaptive link account not only entails that they are unjustified but that they are likely to be false (Street (2006, 125). This is a stronger conclusion, as our belief in a proposition *p* could fail to be justified without *p* being likely to be false. For example, consider the belief that the number of living rabbits on Earth right now is odd. This is no more likely to be false than the opposite claim. Still, we hardly have sufficient reason for accepting

3. Safety and sensitivity

The challenge a realist has to face, according to Street, is to reconcile the evolutionary account with the supposition that our moral beliefs are reliably formed without thereby committing herself to a scientifically unviable version of that account. Before considering Enoch's answer to this challenge, I shall briefly consider another recent attempt to respond; an argument to the effect that it can quite trivially be met and that the challenge is not really a genuine challenge at all. In my view, this other attempt fails. However, by seeing how it fails, we get a better understanding of the nature of the challenge.

The argument in question has been developed by Justin Clark-Doane. He thinks that Street's argument is congenial with the so-called Benacerraf challenge against Platonism in mathematics and focuses especially on Hartry Field's formulation of it.¹⁰ Field assumes that Platonism is a viable position only if it appears in principle possible to explain the reliability of our mathematical beliefs on Platonist grounds (i.e., on the basis of Platonist assumptions about the nature of the facts and entities the beliefs posit).¹¹ The reason is that if their reliability cannot be explained on such grounds then a Platonist is committed to the view that our mathematical beliefs are unjustified.¹² And the point is that the task of providing such an explanation seems very hard, given that Platonists believe that the facts and entities postulated by our mathematical beliefs are mind- and language-independent, causally inert, and bear no spatio-temporal relations to us.

What is it, more specifically, that needs to be explained? Field suggests that it is a correlation between our mathematical beliefs and the relevant facts. What is to be explained, more specifically, is "that the following schema

it. I think the evolutionary considerations she invokes at best support the weaker conclusion (which also seems to be Joyce's view) but I shall leave that discussion aside in what follows.

¹⁰ See Field (1989, esp. 25-30, 230-9), Field (1996) and Field (2005) for his construal of the Benacerraf challenge. See Benacerraf (1973) for the initial formulation of the challenge.

¹¹ Field sometimes formulates the condition differently, by insisting that what matters is whether the Platonist can "make it plausible" that the reliability of our mathematical beliefs can be explained on Platonist grounds. See Field (1989, 232).

¹² That is, it is supposed to imply that the beliefs are not justified given that their contents are Platonistically construed. Field's underlying idea is that the apparent impossibility to explain the reliability of a set of beliefs tends to "undermine" them (Field (1989, 26)). I am, like Clark-Doane, assuming that Field thinks that it undermines them in the sense of showing that they are not justified. For textual support, see Field (1996, 377).

If mathematicians accept ‘p’ then p

(and a partial but hard to state converse of it) holds in nearly all instances, when ‘p’ is replaced by a mathematical sentence.” (Field (1989, 26).

The quote suggests that Field thinks that the correlation a Platonist must be able to explain is quite strong. But we can also imagine weaker versions of the requirement, by restricting the scope or strength of the correlation it must be possible to explain. For example, we may argue that we must only be able to explain the reliability of those of our mathematical beliefs that are relatively simple (those that I, following Clark-Doane, shall refer to as our “core” mathematical beliefs), and that the task of explaining the reliability of highly theoretical ones is less important. Weakening the requirement along these lines may be especially appropriate in the case of moral beliefs.¹³

Clark-Doane argues that, to meet any plausible requirement of the kind Field might have in mind, it is enough if the Platonist has resources that allow her to show, first, that if the target facts had been different, our (core) mathematical beliefs would have been likewise (in which case they are said to be “sensitive”), and, second, that the beliefs in question could not have easily been false (in which case they are said to be “safe”).¹⁴ And his point is that the Platonist *has* such resources, in spite of her peculiar views about the nature of mathematical entities. As for their sensitivity, Clark-Doane points out that mathematical facts are plausibly held to be metaphysically necessary, and that, given their necessity, the sensitivity of our mathematical beliefs follows trivially. After all, the necessity of the target facts excludes the existence of possible worlds in which those facts are different. So it also excludes the existence of worlds in which they are different while our beliefs remain the same.¹⁵ Moreover, given their sensitivity and their actual truth, all a Platonist has to do to establish the safety of the beliefs is to compellingly argue that they could not easily have been different. And this is something she could argue, in the case of our (core) mathematical beliefs, for

¹³ This is a point Enoch makes, see Enoch (2010, 427). I shall return to it further on.

¹⁴ I use the definitions of sensitivity and safety that Clark-Doane employs in “Genalogy and Reliability”.

¹⁵ The source of inspiration here is a well-known passage in Lewis (1986): “[I]f it is a necessary truth that so-and-so, then believing that so-and-so is an infallible method of being right. If what I believe is a necessary truth, then there is no possibility of being wrong. That is so whatever the subject matter...and no matter how it came to be believed.” (114-5).

example by invoking an evolutionary account of their origins. The idea seems to be that, given such an account, we would have had different beliefs only if our environment had been quite different, which means that it excludes the existence of near-by possible worlds in which our (core) mathematical beliefs are different.¹⁶ Clark-Doane furthermore notes that, as (some) *moral* facts are also commonly assumed to be necessary, and as our (core) moral beliefs are also, arguably, “evolutionarily inevitable”, the same type of reasoning is available also to someone who wants to meet Street’s challenge.¹⁷

The most obvious way to question this reasoning is perhaps to deny that the target facts are necessary. (That there is room for this response is perhaps most clear in the moral case.) Alternatively, we may deny that their necessity makes it vacuously true that if the facts had been different our beliefs had been correspondingly so.¹⁸ Or we could deny that the evolutionary account implies that they have the relevant kind of robustness. However, the response I shall pursue is to offer an interpretation of “explain reliability” such that that the considerations Clark-Doane invokes do *not* provide a sufficient basis for such an explanation. I shall primarily focus on mathematics, but the comments apply equally well to ethics.

Recall that the fact it must be possible for a Platonist to explain is, according to Field, a correlation between the target beliefs and the mathematical truths. The point of departure of my interpretation is the natural thought that something qualifies as an explanation of such a correlation only if it entails that it in fact obtains or at least that it is likely to obtain (i.e., that we *have* beliefs in the relevant area that are mostly true). And the problem with Clark-Doane’s argument from the inevitability of our (core) mathematical beliefs and the necessity of the mathematical facts is that it doesn’t. The conclusion we at best can derive from those considerations is the conditional thesis that *if* the target beliefs are true then they could not easily have been false, and that thesis does not suggest that any of them *is* in fact true. Consider an analogy. Suppose that two of our friends visited a certain city in the weekend and

¹⁶ Note that the idea here is *not* that an evolutionary account helps to explain the reliability of the target beliefs through suggesting that they are the results of mechanisms that would not have evolved if they had not generated mostly true beliefs or something to that effect. The evolutionary account is invoked *just* to argue that the target beliefs could not easily have been different.

¹⁷ Clark-Doane notes the ironic fact that, given his argumentation, an evolutionary account of our moral beliefs can be seen to *support* rather than undermine their reliability.

¹⁸ This is a response Field hints at in some passages. See Field (1996, 375). Note that Clark-Doane addresses and tries to respond to both this and the first objection.

that we wonder if they met. Suppose also that we know that each had a very determinate plan about her visit (about the places to go and so on). Then we may perhaps conclude that *if* their paths crossed it could not easily have been different. However, that conclusion does not in turn give us any clues as to whether they actually *did* meet.¹⁹

The requirement that an explanation of the reliability of a set of beliefs must entail or indicate that they are in fact reliable (i.e., that the pertinent correlation holds) helps to account for the significance Field ascribes to the possibility of developing such an explanation. Field presupposes that the principled inability to explain the reliability of the target beliefs undermines their justification. Now, note that there are two ways to undermine a belief's justification. One is to offer compelling reasons for thinking that it is false. Another is to argue that we lack sufficient grounds for believing in its truth. Given that an explanation of the reliability of a set of beliefs entails or indicates they are mostly true, we may argue that the principled inability to provide such an explanation undermines the target beliefs in the second way, as it rules out a possible way of defending them.

On the construal of the Benacerraf challenge that I propose, then, it is a debunking argument in the sense indicated earlier. This interpretation is congenial with Field's well-known analogy with beliefs about the goings-on in a remote village in Nepal:

Consider the Benacerraf worry about (certain kinds of) platonism in the philosophy of mathematics. The worry is that the same reasons that would lead us to advise against having beliefs about the happenings in a remote village in Nepal, when one has reason to think that there is no possible explanation of the reliability of those beliefs, should equally lead us to advise against having beliefs about mathematical entities platonistically construed, given that it appears that there is no possible explanation of the reliability of those beliefs. (Field (2009, 289))

Suppose for example that we are told by someone that an odd number of the villagers have been born on a Friday. Field suggests that if we find that there is no plausible way of explaining how our "informant" has acquired that knowledge (for example with reference to

¹⁹ If we were better informed about the details of their plans, then we would perhaps be able to figure out if they met. But the point is that this further information is not required to reach the conclusion that *if* they met it could not easily have been otherwise.

the assumption that she has made visits to the village), then we should remain skeptical, even if our beliefs about the village do not conflict with anything else we believe or have reason to believe. However, this is evidently not because the absence of such an explanation is a reason for thinking that it is *not* the case that an odd number of the villagers have been born on a Friday, but rather since it shows that our grounds for trusting that claim is too fragile.

In other words, on the construal I favor, it is the *positive* support an explanation of the reliability of the target beliefs provides that is the source of the significance Field attaches to the principled inability to construct such an explanation. And that requires that the concept of a reliability-explanation is be construed so that the availability of such an explanation *does* provide such support. That is why an explanation counts as being of the relevant type only if it entails or indicates that most of the target beliefs are in fact true. And, again, as the observation that our mathematical beliefs are inevitable and such that, if they are true, then they are necessarily true, does *not* suggest that any of them are true, that is also why Clark-Doane's argument does not qualify as a explanation of their reliability in the relevant sense.

The thesis that a reliability-explanation must entail or indicate that most of the target beliefs are true just states a necessary condition. What else is needed? The Benacerraf challenge is sometimes said to be a request to explain why, assuming that most of our beliefs in the pertinent area are true, it is not a mere *coincidence* that they are true.²⁰ What Clark-Doane can be seen as offering along those lines is a composite answer. He offers one explanation of why it is not a coincidence that we have the mathematical beliefs that we do have (the explanation that appeals to evolution) and a separate explanation of why it is not a coincidence that the (contents of the) beliefs are actually true (the explanation that appeals to the necessity of the target facts). What is needed is instead, as Field puts it, something more "unified" (Field (1996, 370-71)). Here is my proposal: To explain the reliability of the target beliefs is to explain the psychological fact that we have the beliefs with reference to claims whose truth also suggests that most of them are correct.²¹ Thus, an explanation of the reliability of a set of beliefs is a kind of *argument* for the truth of the target beliefs; an

²⁰ See for example Field (2005).

²¹ Field himself rather tries to explicate "unified" in terms of "counterfactual persistence". Clark-Doane takes this to mean that the feature he is after is really sensitivity (as defined above). If so, then it is vulnerable to the type of worries Clark-Doane raises. The interpretation I am offering is supposed to avoid the worries.

argument whose premises essentially include the assumptions it invokes to account for the fact that we have them. This is the interpretation I will use in what follows.

I shall consider two objections to my criticism of Clark-Doane's proposal. On the first objection, requiring that a Platonist can explain the reliability of our (core) mathematical beliefs in the sense I have just indicated is to beg the question against her. This objection is related to Clark-Doane's view that the Benacerraf challenge is aimed to be "non-skeptical" and to his observation that Field, in stating the challenge, grants that our (core) mathematical beliefs are in fact true.

That the challenge is non-skeptical means, in Clark-Doane's use of the term, that it is not a request to justify our mathematical beliefs in a way that would convince a mathematical skeptic (someone "who doubts that there are any (non-vacuous) mathematical truths at all").²² To further clarify the term, he makes a comparison with perceptual beliefs. One way to challenge the epistemic credentials of our perceptual beliefs is to require a defense that does not rely on claims whose justification presupposes the very thesis we are asked to validate (i.e., that perceptual beliefs are worth trusting). A challenge of this kind excludes attempts to justify perceptual beliefs with reference to, say, the theory of evolution or to theories about how our sense organs work, as it is perceptual beliefs that ultimately provide our evidence for such theories. A "non-skeptical" challenge, by contrast, allows us to invoke such theories.

However, the way I have construed the Benacerraf challenge does not exclude that it is non-skeptical in this sense. According to the central idea of the challenge, the target beliefs are not justified unless it appears possible to develop a justification of them in the form of an explanation of their reliability. The moral of the comparison with perceptual beliefs is not that this isn't a reasonable requirement, not even given "my" view about what it is to provide such an explanation. The moral is rather that, when we try to develop such an explanation, it is legitimate to invoke claims whose justification presupposes the reliability of the target beliefs.²³ This is consistent with the interpretation I am employing. In fact, to offer an explanation of that kind is exactly the strategy that is instantiated by Enoch's attempt to meet Street's challenge.

²² See "What is the Benacerraf Problem?" for Clark-Doane's use of this concept.

²³ Similarly, it is not to beg the question against a theorist to require that she can do more in to justify her beliefs than to simply stress their truth. That is, we can (as Field does) grant that the Platonist's beliefs are true and still insist that she is justified in having them only if she has the ability or resources to construct a more complex argument for them.

According to the second objection, although the Platonist's capacity to show that, given the truth of our (core) mathematical beliefs, they could not easily have been false, does not exclude that there is a sense in which she can still be unable to explain their reliability, it *does* exclude that this inability *undermines* them. Clark-Doane offers in this context an epistemic principle that is stated as follows:

Modal Security: Information, E, cannot undermine our beliefs of a kind, D, without threatening our ability to show that our D-beliefs are both safe and sensitive.²⁴

Clark-Doane suggests that this principle, which he takes to be “extremely weak”, rules out the existence of an interpretation of “explain the reliability” which makes both of the following claims come out true:

- A. If it appears impossible to explain the reliability of our mathematical beliefs then the justification of those beliefs is undermined.
- B. This *does* appear impossible.

The argument Clark-Doane seems to have in mind for thinking that *Modal Security* conflicts with A and B goes as follows: Given *Modal Security*, A and B imply that there is a consideration that threatens our ability to show that the target beliefs are safe and sensitive. However, since we *can* arguably establish their sensitivity and safety, by appealing to their evolutionary inevitability and to the necessity of the facts they represent, we know this implication to be false. Thus, *Modal Security* implies that no interpretation of “explain reliability” is such that, given this interpretation, it is both the case that it appears impossible to explain the reliability of our mathematical belief beliefs and that this undermines them.

However, I find the objection that appeals to *Modal Security* unpersuasive. Note in this context that the concept of a belief being safe can be given two different readings. On the first reading, a belief cannot be safe unless it is actually true. On the second, safety doesn't imply truth and the claim that a belief is safe merely means that *if* it is true then it could not easily have been false. Now, as we saw, the argument to the effect that *Modal Security* excludes the truth of the conjunction of A and B presupposes the soundness of the argument from

²⁴ “Genealogy and Reliability”

inevitability and necessity. However, that argument is sound only under the second interpretation of “safe”, since it at best establishes the conditional claim that, given the actual truth of the beliefs, they could not easily have been false. And the point is that, under the second interpretation, *Modal Security* is not a plausible principle. For, on the second interpretation, *Modal Security* implies that we cannot undermine a belief by showing that we lack sufficient grounds for it, as the fact that a belief is safe in the second sense does not guarantee that there are such grounds. This is clearly an implausible implication.

The point could be put differently. Under the interpretation of “safe” according to which safety implies truth (the first sense), *Modal Security* may be defensible, since we, on that interpretation, cannot show that a belief is safe without showing that it is true. However, on that interpretation, there is no conflict between *Modal Security* and the conjunction of A and B. The conjunction of A, B and *Modal Security* does not imply a contradiction, it merely implies that a necessary condition for our having the ability to establish the sensitivity and safety of the target beliefs is that it appears in principle possible to explain their reliability. Clark-Doane denies that this is a necessary condition, on the ground that the safety of the beliefs can be established with reference to their inevitability and to the necessity of the relevant facts. However, as I have insisted, those considerations only establish their safety under the second interpretation of “safe”. And, as we have seen, on that reading of “safe”, *Modal Security* is not plausible. Thus, either way, regardless of if we take safety to imply truth or not, the argument that appeals to *Modal Security* to the effect that there is no interpretation of “explain the reliability” which makes both A and B come out true fails.

The discussion about Modal Security might seem complex, but the upshot of this section is simple: The fact that a theorist has the capacity to show that, given the truth of some of her beliefs, they could not easily have been false, does not necessarily mean that she is justified in having them, *not even if we assume that they are in fact true*. This is why the observation that a Platonist and a moral realist has that capacity as regards our (core) mathematical and our (core) moral beliefs is not, as far as I can see, really relevant to the possibility of meeting challenges of the kind Field and Street have proposed.

4. Enoch and the value of survival

Let us turn to Enoch’s attempt to respond to Street’s argument. Enoch is a “robust” moral realist. This position is congenial with Platonism in that it holds that the facts postulated by

our moral beliefs are mind-independent, causally inert and irreducible to natural truths,²⁵ and may therefore seem vulnerable to similar epistemological worries. Enoch acknowledges the need for a robust realist to say something about how to explain the reliability of our moral beliefs. However, unlike Street, he does not think that the adaptive link account excludes the possibility of providing such an explanation. Instead, like Clark-Doane, he thinks that it helps the realist to respond to the challenge.

The salient feature of Enoch's explanation is that it invokes a substantive moral claims. Like Field, Enoch assumes that what needs to be explained is a correlation; a correlation between our moral beliefs and the facts in virtue of which they are supposed to be true. He notes that there are different ways of explaining such a correlation. An obvious example is to argue that the target beliefs are caused by the facts. This is the possibility Street thinks is undermined by her objections to the tracking account. However, the type Enoch is pursuing is different. He is going for a "third-factor explanation", according to which the correlation is explained by an external factor that is taken to be responsible both for the facts and for the beliefs. Let us focus on "rightness beliefs"; i.e., beliefs to the effect that certain types of behavior are morally right. According to the adaptive link account, roughly, we are disposed to have the rightness beliefs that we do have because the behavior they prescribe increases the agent's fitness. It often does so, moreover, through helping the agent or her kin to survive. The fact that the behavior helps the agent or her kin to survive is the "third factor" to which Enoch wants to appeal. And it is to argue that this fact may also be responsible for the *truth* of the beliefs that he invokes a substantive moral assumption.

The assumption in question is the claim that "survival or reproductive success [...] is at least somewhat good" (Enoch (2010, 430)). The claim is clarified as follows:

Not, of course, that it is always good, or that its positive value is never outweighed by other considerations, or even that it is of ultimate or of intrinsic value, or anything of the sort. Furthermore, I am not asking you to assume that the evolutionary "aim" is of value *because* it is the evolutionary aim. All I will be relying on is the assumption that survival (or whatever) is actually by-and-large better than the alternative (430).

²⁵ A potentially important difference is that it does not, unlike Platonism about mathematics, imply that the target beliefs postulate any special moral entities. I will ignore this difference, however.

Now, if we combine the claim that survival is somewhat good with the idea that the fact that an action promotes something good suggests that it is right then we may reach the sought-for conclusion. For, on the evolutionary account, roughly, the factor that explains our rightness beliefs is the fact that the actions they ascribe rightness to tend to promote survival. And, given the view that survival is good, this factor also tends to make the actions right. Thus, it suggests that many of them are at least “somewhat in line with the normative truths” (430).

In other words, the adaptive link account is supposed to generate skeptical conclusions (given a realist construal of them) because it suggests that the existence of our beliefs is not appropriately sensitive to the truth of their contents. However, if we are disposed to believe that an action is right to the extent that it promotes our chances of survival, and if the fact that it does so indicates that it is right, then our rightness beliefs *are*, to some extent, sensitive to their truth. What he thereby argues (to illustrate how his suggestion fits with the dialectic of Street’s paper) is that a realist can deny the no-relation assumption without committing herself to the tracking account.

A central concept in this account is that of a fact being “responsible” for some other fact. Enoch takes this term to cover both a causal relation and a metaphysical relation of constitution. It is the latter type of relation that is supposed to hold between the fact that an action promotes survival and the truth of the belief that it is right. He writes:

Now, unlike the paradigmatic pre-established harmony kind of explanation—of the mind-body harmony, for instance—my explanation of the correlation between the normative truths and our normative beliefs is not causal in both directions. Perhaps the evolutionary “aim” —whatever exactly it is—causally shapes our normative beliefs, but the fact that the evolutionary “aim” is of value does not causally shape the normative truths. It seems more appropriate to say that it is related to them in some constitutive way, a way the details of which depend on the details of your favorite first-order, normative theory. For instance, the fact that survival is good is plausibly related in coherence relations to many (though perhaps not all) other normative truths, like that pain is pro-tanto bad, that some close relationships are good to have, etc. Relations of this nature between that survival is good and many other normative truths—though not causal—still allow me, I think, to explain the correlation between the normative truths and our normative beliefs in a way

that resembles pre-established accounts sufficiently to merit the name.²⁶

Enoch concedes that the type of reliability we can squeeze out from his account is not overwhelming, as he thinks that it merely suggests that the moral beliefs we have been endowed with through evolution are “reasonably good starting-points” that are not “too far off”. However, that conclusion is arguably sufficient to avoid the type of radical skepticism Street thinks a realist is committed to, given the adaptive link account. Remember that Street thinks that realists are committed to the view that, if our moral beliefs are true, this is just a matter of fluke or chance. So if the likelihood can be shown to be slightly greater than chance, on the basis of the adaptive link account, that is enough to refute her claim. Moreover, Enoch thinks that we have reason for an even stronger optimism when we consider the impact that ordinary reasoning processes may have on our beliefs. He writes:

Given a starting point of normative beliefs that are not too far-off, presumably some reasoning mechanisms (and perhaps some other mechanisms as well) can get us increasingly closer to the truth by eliminating inconsistencies, increasing overall coherence, eliminating arbitrary distinctions, drawing analogies, ruling out initially justified beliefs whose justificatory status has been defeated later on, etc.

It is the combined effect of these factors—the fact that our initial beliefs about what is right and wrong depend on the extent to which they promote survival and the benign effects of subjecting to a reflective equilibrium-style reasoning process—that are thought to be sufficient for fending off the type of skeptical worries Street and others have raised.

5. Coherence and explaining error

Enoch’s attempt to respond to Street’s reliability challenge raises a number of questions. One objection goes as follows. Enoch acknowledges that a realist must avoid having to conclude that it is a mere coincidence that we have ended up with true moral beliefs. That such a coincidence exists can hardly, he thinks, be treated as a brute fact that we may legitimately posit in the absence of any reasonable explanation. As we have seen, Enoch tries to provide

²⁶ Enoch (2010, 431). One might wonder how the assumption that a relation of constitution of this kind can hold between the non-moral fact that some action promotes survival and a moral fact can be reconciled with the type of non-naturalism robust realism is associated with. However, I shall set that complication aside.

the requested explanation by assuming that survival is good. However, it might be thought that Enoch thereby just pushes the buck, in that he manages to avoid postulating one surprising fact simply by positing another. What Enoch in effect is assuming is that the moral facts happen to be such that the adaptive link account supports rather than undermines the reliability of our moral beliefs. Couldn't that also be seen as a mysterious coincidence that requires an explanation?

Enoch responds to this objection in part by trying to play down the alleged mysteriousness of the truth of the moral assumption his explanation invokes (see Enoch (2010, 433-4)). However, I shall not enter into that discussion. Instead, I shall focus on other ways of challenging examples of the strategy his proposal instantiates.

In the introductory section, I mentioned the possibility of complaining that the strategy in question is illegitimate on the ground that it tries to justify the target beliefs by invoking beliefs of the very kind whose epistemic credentials that are challenged. One might respond by insisting that the view that one must be able to explain the reliability of the target beliefs on purely independent grounds is implausible because it leads to a too indiscriminate form of skepticism.²⁷ Take for example perceptual beliefs. The only available account of their reliability that shows any promise involves the theory of evolution and various claims about how our sense organs interact with our environments.²⁸ This account is circular in the same way as Enoch's, as the claims it invokes are empirical and as the evidence we have for them is exhausted by our perceptual beliefs. So if we want to avoid skepticism also about perception we should allow for a similar circularity in the case of moral beliefs.²⁹

In other words, it appears too strict to require that a non-skeptic about our moral beliefs must be able to show that they can be justified on purely independent grounds (i.e., without invoking any of their contents). The point I presently want to make, however, is different, namely that there are ways of criticizing explanations of the reliability of our moral beliefs

²⁷ This presupposes that the requirement to explain the reliability of the target beliefs is supposed to apply generally to all domains, which seems to be Field's view. See Field (1996, 377).

²⁸ For example, we could argue that our perceptual beliefs are reliable (when formed in suitable conditions) since they are the results of mechanisms that have evolved, as Joshua Schechter puts it, because it "conferred a heritable survival or reproductive advantage on our ancestors to correctly represent their environment" (2010, 444).

²⁹ Note also that Field acknowledges that the account that is available in the perceptual case is satisfactory, in spite of the circularity. See Field (1989, 25-30).

that invoke substantive moral claims even if one concedes that this strategy is in principle legitimate. I shall indicate the available types of complaints by stating two conditions that a satisfactory instance of the strategy arguably must satisfy.

The reliability-explanation that is available in the perceptual case appeals to claims such that, to determine their truth, we have to consult beliefs of the very kind whose reliability we want to explain. An important feature of that account is that our perceptual beliefs actually *do* confirm the empirical claims it invokes (in the sense that the claims arguably contributes to the best explanation of the truth of the contents of some of those beliefs). It is easy to overlook the importance of this fact, and to think that the type of fit that thus has been achieved is somehow trivial. However, that is not so. Having established the fit is a real achievement, as the initial presumption in favor of the reliability of the target beliefs could conceivably have led to a less positive verdict about their reliability.³⁰

A similar fit must be obtained in the case of our moral beliefs. In other words, if our explanation of the reliability of our moral beliefs invokes certain moral assumptions, there should be no tension between those assumptions and (the contents of) the beliefs. I am going to call this condition COHERENCE. Here is an attempt to formulate it a bit more carefully: The moral claims invoked by the explanation must not be disconfirmed by the beliefs whose reliability it is supposed to explain, where they are “disconfirmed” to the extent that they conflict with the conclusions we are led to accept, through competent reasoning, by taking the contents of the beliefs as starting points in a reflective equilibrium-style reasoning process. The obvious reason for imposing COHERENCE is that the violation of it would make the defense provided by the explanation self-defeating. For the assumption that it *does* justify the target beliefs would, in such a case, also show the defense to be flawed (as (the contents of) those beliefs would undermine some elements of it).

The second condition has to do with the strength of the correlation (between our moral beliefs and the facts a moral realist posits) that it must be possible for her to explain. Given the extensive disagreement that exists regarding moral issues, the correlation the realist may hope to be able to explain is not going to be perfect. This is not in itself a problem, but it does give rise to a certain constraint that her explanation arguably should satisfy. Thus, she can try to explain the reliability of the target beliefs by arguing that they are the results of a process that generally, but not always, leads to true beliefs. But then we should require that the

³⁰ See Shogenji (2000) for a discussion of this point and its significance.

considerations that are supposed to support thinking that the process is thus reliable also help explain why it sometimes leads to error. For example, she may try to blame the errors on some type of distorting factor—a factor that is supposed to interfere with a process that in other cases are held to work fine. But then her account should provide resources for an argument to the effect that the factor in question *deserves* being regarded as distorting. I am going to call this admittedly rather vague condition EXPLAINING ERROR. It is clearly satisfied in the perceptual case, as the reliability-explanation that is available in that context does provide a ground for thinking that the factors to which we typically attribute perceptual errors (e.g., bad lighting) are distorting in the relevant sense.

The reason for imposing EXPLAINING ERROR is also straightforward. If we must concede that a belief-forming process leads to inexplicable errors, our ground for treating it as reliable and for relying on *any* of the beliefs it generates on that ground, would be diminished. This is especially so when we (such as in the moral case) lack an independent, direct, way of confirming that many of the beliefs the process leads to are true.³¹

6. Applying the conditions

Neither COHERENCE nor EXPLAINING ERROR is trivially satisfied, and supposing that they must be met by a satisfactory explanation of the reliability of our moral beliefs does not beg the question against the realist. Their significance can be established on neutral grounds, and whether they are satisfied or not can be independently determined. The purpose of this section is to illustrate how one may argue that a particular candidate explanation that invokes substantive moral claims violates the conditions and thus fails to provide a compelling response to the reliability challenge. I shall in particular focus on Enoch's version. However, note that my aim is primarily to criticize Enoch. In fact, I think his proposal might ultimately turn out to be successful. The aim is rather to illustrate how the discussion about the meta-ethical implications of an evolutionary account of the origins of our moral beliefs, in my view, should proceed.

The critical strategy associated with COHERENCE is to argue that the moral assumptions we need to make to generate the conclusion that our moral beliefs are reliably

³¹ For a view that is congenial with the assumption that EXPLAINING ERROR is necessary in order for an explanation of the reliability of a set of target beliefs to be satisfactory, see Weinberg (2007). For further discussion of these and other possible conditions, see Tersman (2012).

formed cannot plausibly be reconciled with those very beliefs. More specifically, it is to argue that the assumptions in question are inconsistent with the conclusions we would be led to accept if we were to use the contents of the target beliefs as our starting points in a reflective equilibrium-style reasoning process.

In the case of Enoch's proposal, the relevant assumption is the claim that survival is ("somewhat") good. Determining whether this proposal satisfies COHERENCE requires that we know something about where we *would* end up when trying to achieve a reflective equilibrium. This issue is, in a way, precisely what the current discussion in normative ethics is all about, and it is interesting to note that the discussion of the reliability of our moral beliefs thus depends on what is the main project in normative ethics. It might seem premature to say anything confident about the outcome of that discussion. However, we can still make progress. A limited number of general accounts about moral rightness have emerged in the discussion and are commonly seen as the main competitors, including Kantianism, utilitarianism and virtue theory. One way to investigate if Enoch's proposal satisfies COHERENCE is to explore the implications of each of those competitors.

For example, consider classical utilitarianism, according to which an action is morally right if and only if it maximizes the total net sum of happiness or welfare, and suppose that this is the position that the moral beliefs we have been endowed with through evolution ultimately would lead us to accept after using the method of reflective equilibrium. Given that assumption, in order for COHERENCE to be met by Enoch's proposal, it is not enough that there is an interpretation of the claim that survival is good that is consistent with classical utilitarianism (which there obviously is). The interpretation must also be such that, given utilitarianism, the claim that survival is good generates, together with the evolutionary account, the conclusion that the target beliefs are (sufficiently) reliable. And that is doubtful.

The evolutionary account implies that we judge actions to be right to the extent that they promote the agent's survival or the survival of her kin, and classical utilitarianism implies that an action is right to the extent that it maximizes welfare. So, in order for the claim that survival is good to support the conclusion that the target beliefs are (sufficiently) reliable, it must ensure that there is a significant correlation between actions that promote the survival of the agent or her kin and actions that maximize welfare. Now, to be consistent with classical utilitarianism, the claim that survival is good must be taken to mean that it is merely instrumentally good (as classical utilitarianism conceives of happiness or welfare as the sole (positive) intrinsic value). And there are at least two reasons to doubt that the view that it is instrumentally good is sufficient for ensuring that a correlation of the relevant type obtains

(even granted Enoch's point that it might be relatively modest). First, the fact that an act promotes the agent's survival, or the survival of her kin, does not exclude that it often is detrimental to the survival of others. Second, survival is evidently not the only thing that has instrumental value, and the fact that an action promotes the survival of some individuals might be outweighed by other consequences. On a utilitarian view, of course, *all* the effects of an action that affect the total amount of welfare that exists in the world are relevant to its rightness, regardless of whether these effects are brought about by survival or something else, and the fact that it has *some* consequences that can be expected to be favorable does not automatically give us any ground for verdicts about its *total* outcome.

This provides, in my view, some ground for thinking that, if classical utilitarianism would come out the winner in the reflective equilibrium-style process, Enoch's proposal violates COHERENCE. Of course, establishing that conclusion is not sufficient for concluding that it violates COHERENCE full stop. To reach such a verdict, we would have to systematically examine also the other main competitors. That this is a worthwhile project is one of the suggestions I want to make about how the discussion about the meta-ethical implications of an evolutionary account of our moral beliefs should proceed.

In the case of EXPLAINING ERROR we can be briefer. Enoch's explanation of the reliability of the target beliefs involves two parts. The first is the combination of the evolutionary account of the origins of the beliefs and the claim that survival is good. The second is the assumption about the benign effects of reasoning and a competent search for coherence. Given the existence of disagreement, the realist must concede that some of the target beliefs are false. Since the dispositions that have evolved through natural selection presumably have had at least some role in all those cases, the realist must attribute those errors to the influence of other factors, such as, perhaps, historic contingencies of different kinds. For example, some people may have idiosyncratic rightness beliefs because they have grown up in an idiosyncratic cultural environment. Are those additional influences distorting or truth-conducive? In order for Enoch's proposal to satisfy EXPLAINING ERROR it must contain resources for giving an answer to that question, and the problem is that it doesn't appear to do that. The reason is partly that the claim that survival is good does not exclude that there are other factors that determine an action's moral status, besides its capacity to promote the agent's survival or the survival of her kin. An exception is provided by factors that manifest the subject's incompetence in the search for coherence (such as fallacious reasoning). But it is doubtful if these resources are sufficient, as disagreements obtain also between people who do not appear subject to any such shortcoming. Again, however,

reaching a confident verdict about this issue requires investigating the nature of the disagreements that actually occur and to see how the proposed account can be extended and developed in more detail to account for error.³²

7. Conclusion

Sharon Street, Richard Joyce and others stress that the emergence of evolutionary accounts of the origins of our moral beliefs faces realists with a certain skeptical challenge. The challenge is to explain how, given such an account, it can be more than a fluke that we have arrived at true moral beliefs. Some have tried to respond to this challenge by providing explanations that invoke substantive moral assumptions and thus the contents of (some of) the very beliefs whose reliability we seek to explain. Anti-realists might be unwilling to grant that such a strategy is legitimate. I have argued, however, that, even if they *do* concede its legitimacy, there is still room for criticism.

Maybe it needs to be stressed at this point that I am not supposing that the evolutionary challenge presupposes that, unless a reliability-explanation has actually been *articulated* and been *shown* to satisfy the relevant conditions, skepticism about the status of the target beliefs follows. That is clearly not a plausible view. After all, as is often pointed out, the reliability-explanation that is available in the case of perception is of a rather recent vintage, and people were surely justified in trusting their perceptual beliefs even before it emerged. Recall also that Field thinks that what it takes for a Platonist to respond to the Benacerraf challenge is “merely” that she can make plausible that it is “in principle possible” to explain the reliability of our mathematical beliefs.

Still, the best way to make plausible that it is possible to do something is presumably to do it, which in itself motivates conducting the investigations I propose. Moreover, even if the requirement advocated by Field may seem weak, there is a risk of underestimating its strength. Recall that one of his sources of inspiration is the story about the village in Nepal. The idea is that, unless it is possible to explain the reliability of a person’s reports about what happens in this village, we should have less confidence in them. A potential explanation of their reliability is that she has frequently visited the village. However, the mere fact that that scenario is *possible* and, say, such that it is not excluded by anything we know, is surely not enough. We should also require some positive evidence for supposing that such visits have

³² See Tersman (2006) for a more extensive discussion along these lines about how moral disagreement can have skeptical implications.

actually taken place. It might similarly be held that, unless we can offer some positive evidence to the effect that an account could eventually be developed that adequately explains the reliability of our moral beliefs, skepticism looms. And gathering such evidence involves engaging in the type of projects that I have proposed.

References

- Benacerraf, P. (1973). "Mathematical Truth", *Journal of Philosophy* 70, 661-79.
- Brink, D. *Moral Realism and the Foundations of Ethics*, Cambridge: Cambridge University Press, 1989.
- Brosnan, K. (2011). "Do the evolutionary origins of our moral beliefs undermine moral knowledge?", *Biology and Philosophy* 26, 51-64.
- Clark-Doane, J. "What is the Benacerraf Problem?", in Pataut, Fabrice (ed.), *New Perspectives on the Philosophy of Paul Benacerraf: Truth, Objects, Infinity*, forthcoming.
- Clark-Doane, J. "Genealogy and Reliability", this volume.
- Enoch, D. (2010). "The epistemological challenge to metanormative realism: how best to understand it, and how to cope with it", *Philosophical Studies* 148, 413–38.
- Enoch, D. (2011). *Taking Morality Seriously: A Defense of Robust Realism*, Oxford: Oxford University Press.
- Field, H. (1989). *Realism, Mathematics, and Modality*. Oxford: Blackwell.
- Field, H. (1996). "The A Prioricity of Logic", *Proceedings of the Aristotelian Society* 96, 359-79.
- Field, H. (2005). "Recent Debates About the A Priori", in Gendler and Hawthorne, eds., *Oxford Studies in Epistemology* (Oxford University Press 2005), pp. 69–88.
- Field, H. (2009). "Epistemology without metaphysics", *Philosophical Studies* 143, 249-90.
- Harman, G. (1977). *The Nature of Morality*, New York: Oxford University Press.
- Harman, G. (1986). "Moral Explanations of Natural Facts — Can Moral Claims Be Tested Against Moral Reality?", *Southern Journal of Philosophy* 24 (suppl.), 57-68.
- Joyce, R. (2006a). *The Evolution of Morality*, Cambridge, Mass.: MIT Press.
- (2006b). "Metaethics and the empirical sciences", *Philosophical Explorations* 9, 133-48.
- (2008). "Precis of *Evolution of Morality* and Reply to Critics", *Philosophy and Phenomenological Research* 77, 213-67.

- (2013). “The Evolutionary Debunking of Morality”, in J. Feinberg & R. Shafer-Landau (eds.), *Reason and Responsibility*, 15th edition, Cengage, in press.
- Kahane G. (2011). “Evolutionary Debunking Arguments”, *Nous* 45, 103-25.
- Lewis, D. (1986). *On the Plurality of Worlds*. Oxford: Blackwell.
- Mackie, J. (1977). *Ethics. Inventing Right and Wrong*, New York: Penguin.
- Schafer, K. (2010). “Evolution and Normative Scepticism”, *Australasian Journal of Philosophy* 88, 471-88.
- Schechter, J. (2010). “The reliability challenge and the epistemology of logic”, *Philosophical Perspectives* 24, 437-64.
- Shafer-Landau, R. (2003). *Moral Realism: A Defense*, Oxford: Oxford University Press.
- Shogenji, T. (2000). “Self-Dependent Justification Without Circularity”, *British Journal for the Philosophy of Science* 51, 287-98.
- Skarsaune, K. (2011). “Darwin and Moral Realism: Survival of the Fittest”, *Philosophical Studies* 152, 2011, 229-43.
- Street, S. (2006). “A Darwinian Dilemma for Realist Theories of Value”, *Philosophical Studies* 127, 109-66.
- Tersman, F. (2006). *Moral Disagreement*, New York: Cambridge University Press.
- Tersman, F. (2012). “Intuitional Disagreement”, *Southern Journal of Philosophy* 50, 639-59.
- Tropman, E. (2013). “Evolutionary debunking arguments : moral realism, constructivism, and explaining moral knowledge”, *Philosophical Explorations*, 10.1080/13869795.2013.855807 (DOI).