

MORAL CONFIRMATION VS. MORAL EXPLANATION: A TALE OF TWO CHALLENGES

Sarah McGrath

Abstract: In the first chapter of *The Nature of Morality* (1977), Gilbert Harman sets out what he takes to be the “basic issue” confronting moral philosophy: whether moral principles can be “tested and confirmed in the way that scientific principles can . . . out in the world” (3–4). Harman argues that they can’t be. In this paper I argue that if we reject the Harmanian view that confirmation is the converse of explanation, then we can agree with the naturalist realist on the basic epistemological issue of whether moral principles can be tested and confirmed in the way that scientific principles can. But I argue that there nevertheless is an important metaphysical way in which moral explanations differ from certain kinds of non-moral explanations. An upshot is that even realists who think that moral facts are necessary, causally inefficacious, and knowable a priori can agree that moral claims are subject to empirical confirmation in the way that scientific claims are.

1 Introduction

In the first chapter of *The Nature of Morality* (1977), Gilbert Harman sets out what he takes to be the “basic issue” confronting moral philosophy: whether moral principles can be “tested and confirmed in the way that scientific principles can . . . out in the world” (3–4). Harman argues that they can’t be.

There are two questions we might ask about this: First, is Harman right? Second, if Harman is right that there is this fundamental contrast between moral and scientific theories, then what is the upshot for moral knowledge?

Moral realists think that there are moral facts, and that these facts are objective in that they obtain independently of our opinions about whether they obtain. Moreover, moral realists are not skeptics; they think that we have a substantial body of moral knowledge. They have taken Harman’s challenge to be a challenge to realism, and their responses to Harman can be divided into two main strategies. The first strategy is to

argue that *pace* Harman, moral principles *can* be tested and confirmed in the way that scientific principles can be. This is the line that has been taken by naturalist realists, such as Boyd (1988), Railton (2003/1986), and Sturgeon (2001/1985). More recently, non-naturalist realists, such as Dworkin (2011), Enoch (2011), and Scanlon (2014) have argued that although Harman was right that moral principles can't be tested "out in the world" in the way that scientific principles can, this is not bad news for moral knowledge at all; it is a mistake to expect normative facts to play the same kind of explanatory role that scientific facts play. According to Dworkin, the idea that there are "moral particles" that "interact in some way with human nervous systems so as to make people aware of . . . morality or immorality or of . . . virtue or vice . . . is not even a remotely plausible thesis to attribute to" a person who thinks that actions really are right, wrong, virtuous or vicious, "quite apart from its insanity as a piece of physics" (Dworkin 1996, 104–105).

It has always seemed to me that there is something very sensible in both lines of response. On the one hand, it seems sensible for Sturgeon to insist that "there is a whole range of extremely common cases" (2001/1985, 221) in which moral beliefs are confirmed by our experiences. On the other hand, it seems perfectly sensible for Dworkin to say that the epistemology of morality is unlike that of particle physics.

I will argue for a view on which both of these claims are vindicated. My argument involves distinguishing two distinct challenges for realism that have not been clearly separated in the many realist responses to Harman. I will argue that if we reject the Harmanian view that confirmation is the converse of explanation, then we can agree with the naturalist realist on the basic *epistemological* issue of whether moral principles can be tested and confirmed in the way that scientific principles can. (They can be!) But I will argue that there nevertheless is an important *metaphysical* way in which moral explanations differ from certain kinds of non-moral explanations that naturalist realists seem to overlook. I will side with non-naturalist realists on the metaphysical issue of whether explanatory efficacy is a test for whether there are any moral facts (it isn't!), but I will disagree with them about *why* explanatory indispensability is too strong as an ontological requirement. An important upshot is that even realists who think that moral facts are necessary, causally inefficacious, and knowable *a priori* can agree that moral claims are subject to empirical confirmation in the way that scientific claims are.

2 Two Challenges

In this section, I will explain how we can distinguish two separate challenges concerning moral explanation and moral confirmation that go under the heading of "Harman's Challenge" or "Harman's Explanatory Requirement." The first challenge is a kind of Quinean metaphysical challenge.

According to this challenge, moral facts should be included in our ontology only if they play the right kind of explanatory role. The second challenge is purely epistemological; it concerns whether moral beliefs are ever confirmed by non-moral observation in the way that non-moral beliefs are, assuming the most widely held view of confirmation: Bayesianism. My hope is that separating the two challenges can help us to see what is right in both of the opposing realist positions described in the last section.

2.1 Harman Redux

First, let us review Harman's argument in "Ethics and Observation," the first chapter of *The Nature of Morality*. According to Harman,

Observational evidence plays a part in science it does not appear to play in ethics, because scientific principles can be justified ultimately by their role in explaining observations . . .—by their explanatory role. Apparently, moral principles cannot be justified in the same way. . . . Conceived as an explanatory theory, morality, unlike science, seems to be cut off from observation. (1977, 9)

Harman illustrates the contrast that he sees between science and ethics with the following pair of examples (1977, 4, 8):

PROTON: A physicist sees a certain track in a bubble chamber and makes the spontaneous judgment "There goes a proton." The fact that a proton passed through the chamber, we may suppose, helps to explain the physicist's judgment. The proton produces the track, which in turn causes a certain pattern of light to enter the physicist's eye, which in turn causes the physicist to make the judgment.

CAT: Jim rounds a corner and sees a group of young hoodlums pour gasoline on a cat and ignite it. Jim makes the spontaneous judgment "What the children are doing is wrong."

Harman thinks that the important contrast between CAT and PROTON is the following:

[I]t would be reasonable to assume, perhaps, that the children really are pouring gasoline on a cat and you are seeing them do it. But [there is no] obvious reason to assume anything about "moral facts," such as that it really is wrong to set the cat on fire. . . . It would seem that all we need assume is that you have certain more or less well articulated moral principles that are reflected in the judgments you make, based on your moral sensibility. It seems to

be completely irrelevant to our explanation whether your intuitive immediate judgment is true or false. (7)

Whereas the physicist needs to cite the fact that there was a proton in the cloud chamber in order to explain the vapor trail that in turn impacts his sensory apparatus, we don't need to suppose that any action really is wrong in order to explain Jim's judgment. In the moral case, nothing seems to play a role that is analogous to the role of the proton in the physics case.

2.2 The Metaphysical Challenge

Most commenters seem to agree that the important challenge that Harman raises is a kind of ontological challenge about whether there *are* any moral facts. Harman himself is not a nihilist, and does not present the argument as an argument for nihilism. Nevertheless, as [Sturgeon \(2001/1985\)](#) emphasizes, Harman's discussion strongly suggests an argument for nihilism. If moral facts don't explain any of our observations, the argument goes, then we shouldn't believe that they *exist* ([Sturgeon 2001/1985](#), 213).

Although naturalist realists and non-naturalist realists disagree about quite a lot, they nevertheless seem to agree on this ontological interpretation of Harman's challenge. In Sturgeon's classic reply to Harman, he is explicit that he is responding to an argument for "skepticism" or "nihilism." The issue is whether "moral facts are *needed*. . . . [whether] they will turn out to belong in our best overall explanatory picture of the world, even in the long run . . ." (220, emphasis mine). Scanlon similarly takes himself to be engaging in a discussion of ontology:

Harman's explanatory requirement makes good sense in this form: *we have reason to be committed to the existence of things of a certain sort in the natural world only if they play a role in explaining what happens in the natural world (including our experience of it)*. ([Scanlon 2014](#), 26, emphasis mine)

I take it that the argument for nihilism that commentators have extracted from "Ethics and Observation" can be made explicit roughly as follows:

- (P1) We have reason to believe that there are facts or entities of a certain kind only if they play the right kind of explanatory role.
- (P2) Moral facts don't play the right kind of explanatory role.
- (C1) Therefore, we have no reason to believe that there are any moral facts.

And it is a very short step from (C1) to a nihilist conclusion:

- (C2) There are no moral facts.

Reconstructing the argument in this way allows us to see the difference between the ways in which naturalist and non-naturalist realists standardly respond to the metaphysical challenge. The naturalist realist strategy has been to accept P1 and reject P2 (Boyd 1988; Railton 2003/1986; Sturgeon 2001/1985). The non-naturalist realist strategy has been to reject P1 (Dworkin 1996, 2011; Enoch 2011; and Scanlon 2014).

2.3 The Epistemological Challenge

Again, when Harman first explains what he takes the “the basic issue” of moral philosophy to be, he says that it is the issue of whether moral principles can be “tested and confirmed in the way that scientific principles can . . . out in the world” (3–4). Notice that this “basic issue” is not explicitly stated in terms of explanation, but rather, in terms of confirmation. So at first blush, the challenge seems to be an epistemological, rather than an ontological one.

Of course, in Harman’s discussion, the question of whether moral principles are ever confirmed by observation is closely connected to the question of whether moral principles ever explain anything. Harman seems to accept a broadly explanationist model of confirmation, according to which E’s confirming H involves H’s standing in an explanatory relation to E. (For this aspect of Harman’s epistemological views, see especially his early [1965; 1968a; 1968b] papers and Harman [1977]). For this reason, the primary focus of the Harman-Sturgeon debate is the relation of *explanation*, and in particular, the question of whether we should ever regard moral facts as relevant to the explanation of non-moral facts, especially psychological facts about why we make the moral judgments that we do.

However, it is contentious that the connection between confirmation and explanation is as close as Harman takes it to be. For example, on the most widely accepted view of confirmation, Bayesianism, there is no built-in connection between confirmation and explanation. If we distinguish sharply between, on the one hand, the question of whether moral beliefs or observations are ever best explained by moral facts and, on the other, the question of whether non-moral observations sometimes confirm moral beliefs, then the question of whether moral principles can be tested out in the world in the way that scientific principles takes on a different form.

According to the Bayesian, in typical cases of confirmation, the confirmation relation holds between an observation, a hypothesis, and the observer’s *prior probability distribution*.¹ Within this framework, a hypothesis H would be confirmed by observational evidence E just in case the prior probability that the observer assigns to H is lower than the prior probability that she assigns to H on the condition that E (i.e., $\Pr[H] < \Pr[H/E]$). The hypothesis H is confirmed for the observer when the relevant observation is actually made. On this account, the fact that a scientist can be in a

¹ Useful primers for Bayesian confirmation theory include Horwich 1982 and Easwaran 2011.

position to confirm a theoretical hypothesis about subatomic particles by observing a certain experimental outcome is grounded in the fact that she rationally assigns a higher probability to the hypothesis conditional on that experimental outcome than she does to the hypothesis unconditionally. Similarly, the fact that we are sometimes in a position to confirm a historical claim by observing that things are *currently* thus and so is grounded in our rationally assigning a higher probability to the historical claim conditional on things being currently arranged thus and so than we do to the historical claim unconditionally. Within this general framework, the question of whether moral beliefs can ever be confirmed or disconfirmed by observation in the way that scientific principles can amounts to this: Can a believer ever rationally assign a higher prior probability to the moral claim conditional on a certain non-moral observation than she does to the moral claim unconditionally, and then make the relevant observation?

Having divorced explanation from confirmation, we can reconstruct a version of Harman's challenge that doesn't involve explanation, or ontology:

- (P1) Moral beliefs are never confirmed/disconfirmed by empirical observation.
- (P2) Scientific beliefs are confirmed/disconfirmed by empirical observation.
- (C1) Therefore, moral beliefs can't be confirmed/disconfirmed in the way that scientific beliefs can be.

Without the assumptions that confirmation is the converse of explanation, and that explanatory indispensability is a test for reality, this argument doesn't have any ontological implications. One way to see this is to consider the fact that in the history of science there are many examples of claims that have enjoyed a high degree of confirmation even though the stuff that they were about turned out not to exist. Take, for example, phlogiston. There is no such substance as phlogiston; phlogistical nihilism is true. But although there is no such thing as phlogiston, there was in fact observational, confirming evidence in favor of the phlogiston theory of combustion.² The same is true of countless other now discarded theories. So even if moral nihilism is actually true, it doesn't follow that moral claims are not confirmed by empirical observation.

Now that we have seen that the question of whether moral beliefs can be confirmed "out in the world" in the way that scientific beliefs can be is a question that can be considered separately from the question of whether there is an explanatory argument for the existence of moral facts, let us turn to the question of how the realist should respond to these separate challenges.

² For a useful overview, see especially [Chang 2010](#).

3 How the Realist Should Reply to the Metaphysical Challenge

Recall our reconstruction of the metaphysical argument:

- (P1) We have reason to believe that there are facts or entities of a certain kind only if they play the right kind of explanatory role.
- (P2) Moral facts don't play the right kind of explanatory role.
- (C1) Therefore, we have no reason to think that there are any moral facts.

From there, the nihilist makes the move to:

- (C2) There are no moral facts.

How should the realist respond to this challenge? I will argue that she should reject the first premise. But first, I will explain why many naturalists have instead rejected P2. Since Sturgeon's response to Harman is particularly well developed, I will focus on Sturgeon as my representative naturalist. Again, Harman claims that the important difference between Proton and Cat is that, whereas the physicist needs to cite the fact that there was a proton in the cloud chamber in order to explain the vapor trail that in turn impacts his sensory apparatus, we don't need to suppose that any action really is wrong in order to explain anything we observe. Sturgeon denies this claim, but he agrees with Harman that there is a tight connection between confirmation and explanation. And he accepts the idea that if moral facts do not explain anything that we observe—if they do not explain “any non-moral facts we have reason to believe in” (Sturgeon 2001/1985, 219)—then moral facts are not facts that we have reason to believe in. Thus, the debate between Harman and Sturgeon focuses on the issue of whether moral facts ever explain anything.

In favor of a “yes” answer, Sturgeon emphasizes that moral facts very often *are* cited in perfectly reasonable explanations of non-moral phenomena that we have reason to believe in. For example, we can explain why Hitler did the things that he did by citing the fact that he was morally depraved. We can explain why we believe that Hitler was morally depraved by citing the fact that he really was morally depraved (Sturgeon 2001/1985, 217). We can explain why opposition to slavery arising in the eighteenth and nineteenth centuries in French and English speaking North America was relatively vigorous and widespread by citing the fact that the slavery then and there was morally worse than the kinds of slavery that existed at other times and places (Sturgeon 2001/1985, 222–223).

On the face of it, these look like perfectly respectable explanations. And I think that Sturgeon is right that they *are* perfectly good explanations. But I believe that this does not suffice to show that P2 is false. My suggestion is that while Sturgeon is right that moral explanations are perfectly good

explanations, his examples are not examples in which moral features play the kind of explanatory role that would suffice to show that realism wins out over nihilism.³

What would it take for a moral feature to play the right kind of explanatory role to favor realism over nihilism? The answer has to do with *inference to the best explanation*. Plausibly, the physicist knows that there is a proton in the cloud chamber by inference to the best explanation of the vapor trail: without the proton, the vapor trail is unexplained. She sees the vapor trail and knows that there is a proton, because the hypothesis that there is a proton explains the vapor trail. Supposing that Jim knows that what the kids are doing is wrong, how does he know that? I believe, contra Sturgeon, that the answer cannot be: by inference to the best explanation of his own judgment that what they are doing is wrong.⁴

Notice that in cases of inference to the best explanation reasoning, *first* there is something that needs to be explained, and *then* there is an inference to the hypothesis that best explains it. For example, I observe that the cheese has been nibbled. I *then* judge that there is a mouse. The nibbled cheese is evidence for my mouse hypothesis because my mouse hypothesis would explain the nibbled cheese. If my hypothesis that there is a mouse is confirmed, that is because it explains the nibbled cheese, which is what prompted me to make the inference to the hypothesis. But, in the case of Jim, it is hard to see what there is that needs to be explained, that can play a role that is analogous to the nibbled cheese. Let's call the claim that what the kids are doing is wrong *Jim's hypothesis*. What does Jim's hypothesis explain? Sturgeon's discussion suggests that Jim's hypothesis explains Jim's judgment that what the kids are doing is wrong. But this cannot be right. Jim's hypothesis—that *what the kids are doing is wrong*—just is the content of his judgment. The analogous claim in the mouse case would be: I infer that there is a mouse, not because I find the nibbled cheese, but because I find myself judging that there is a mouse. Even if we were to grant that it is reasonable to infer that p at t_2 from the fact that one found oneself judging that p at t_1 , still, when I *originally* find myself judging that p at t_1 , *that* judgment at t_1 cannot have been supported by inference to the best explanation of its own occurrence at t_1 . Thus, if my original judgment is justified, it is not justified by an inference to the best explanation. My original judgment could only be supported by inference to the best explanation of something that I observed *prior* to making my

³ For related discussion of whether Sturgeon's examples favor realism over nihilism see Sober 2015. Sober argues against the claim that Hitler's deeds make the hypothesis that *moral realism is true and Hitler was depraved* more probable than the hypothesis that *moral nihilism is true and Hitler was depraved*, on the grounds that the latter conditional probability is not defined.

⁴ I make this point (2004) in support of the claim that there is moral knowledge by perception.

judgment, such as some nibbled cheese or a vapor trail.⁵ So for this reason, I think Harman is clearly right that there is an important contrast between Proton and Cat. Jim's hypothesis is not supported by inference to the best explanation.

What about Sturgeon's other examples? Consider the claim that the explanation of why there was relatively vigorous opposition to the forms of slavery practiced in the eighteenth and nineteenth centuries is that those forms of slavery were morally worse than the kinds of slavery practiced previously. In this case, the thing to be explained is not a judgment with the candidate explanans as its content, but rather, a feature of an independently occurring event. Could the realist reply to the nihilist by arguing that we must posit a rise in moral atrociousness in order to explain the increased opposition to slavery?

The moral hypothesis that slavery practiced in the eighteenth and nineteenth centuries was morally worse is at least chronologically well-placed to be confirmed by the thing it is supposed to explain—namely, the relatively vigorous opposition to slavery. But that is not likely to impress the nihilist. Let us assume that, if there are any moral facts, then these facts supervene on underlying non-moral facts. Given that assumption, it is far from clear that this kind of example can vindicate the existence of moral facts against skeptics. The worry about this kind of example is that the skeptic already has an explanation that does not mention any moral facts. The underlying non-moral features that ground the moral fact are not in dispute between the realist and the nihilist, and they are sufficient, by themselves, to bring about the rise in opposition. If the forms of slavery practiced in the eighteenth and nineteenth centuries were morally worse than previous forms or slavery, then presumably that moral fact is grounded in underlying non-moral facts about how the slaves were treated. If the underlying non-moral facts are sufficient to explain the rise in opposition, then the additional moral fact is not an explanatory essential. Returning to Harman's example: *that what the kids are doing is wrong* is such that if it is true, then its truth is fixed by the natural, non-moral facts on which it supervenes—that the kids are pouring gasoline on the cat, that they are causing it pain, etc. Those underlying, non-moral facts are not in dispute between Sturgeon and his nihilist opponent: the nihilist who thinks torturing animals is *not* wrong agrees that the kids are pouring gasoline on the cat, etc. Moreover, those facts (together with some facts about Jim's psychology) are sufficient to bring about the effect: that Jim judges that what the kids are doing is

⁵ The problem is not that Jim's judgment, unlike the vapor trail and the cheese, is something *psychological*. Indeed, a similar inference to the best explanation of Jim's judgment made from the third person point of view might be perfectly kosher. Suppose that I know that Jim has highly reliable moral judgment. Then I might correctly infer from the fact that Jim thinks that the kids did something wrong that they really did. The point is rather that if Jim's judgment itself is justified when he originally makes it, it can't be, at that moment, justified by inference to the best explanation of itself.

wrong.⁶ He doesn't have to posit an instantiation of wrongness in order to explain anything he observes. By contrast, that there is a proton is not a higher-level fact that supervenes on lower-level facts that both sides agree obtain, and that are sufficient to bring about the vapor trail. Instead, if there is a proton, then it *causes* the vapor trail all by itself. In the moral case, there simply isn't anything analogous.

That is why the proton, and not the wrongness, plays what I am calling "the right kind of explanatory role": belief in its existence is supported by an inference to the best explanation. When we have a causal structure of the first kind—where a "higher-level" fact, the existence of which is in dispute, supervenes on underlying facts which are sufficient to bring about the effect irrespective of whether the higher-level fact obtains—it is hard to see how we could offer our nihilist opponent a straightforward inference to the best explanation argument that the higher-level facts must exist. Given the sufficiency of the underlying facts to bring about the effect, the moral fact's existence *wouldn't make a difference*: someone who doubts that there are any moral facts already has an explanation of the non-moral observation.⁷ Thus, while the existence of certain kinds of non-moral facts or features can be vindicated, even against skeptics, by inference to the best explanation, it seems to me that this is a very restricted kind of case. Assuming that the facts of fundamental physics do not supervene on more basic facts, the proton-skeptic can't point to more basic facts that would explain the vapor trail. In this way, PROTON is different from CAT.

This contrast between the moral hypothesis, on the one hand, and hypothesis of particle physics, on the other, has to do with whether an inference to the best explanation can vindicate a certain kind of fact or feature against a skeptic. While I have argued that inference to the best explanation arguments don't suffice to vindicate the existence of moral facts or features against a skeptic, I have not argued that moral facts do not, in fact, explain anything. Whether we have to posit the existence of a certain feature or entity in order to explain an observation, and whether that feature or entity actually plays a role in explaining that observation, are two separate issues. So, for example, even if Jim's hypothesis is *not* supported by an inference to the best explanation, it might nevertheless be true that the fact that what the kids are doing is wrong explains his judging that what they are doing is wrong. Nothing that I have said so far—and, I take it, nothing that Harman has said so far—casts doubt on the claim that Jim's judgment is in fact explained by the moral fact it represents.

To see this, notice that there are plenty of non-moral claims p such that p explains q , and I know p , but I do not know p by inference to the best

⁶ See [Schechter 2017](#), 449; [Sturgeon 2001/1985](#), 227–229; and [Thomson 1996](#), 88–91, for discussion of this point.

⁷ I do not mean to suggest that none of our moral knowledge is by inference to the best explanation reasoning. Indeed, I believe that some of it is. My point is just that our justification for believing that there are moral facts and features is not by inference to the best explanation.

explanation of q . Consider the case of color. When I look at a red apple, the fact that the apple is red explains why I judge that it is red, and why I experience it as red. But it is not the case that I believe that the apple is red *because* that hypothesis would explain why I judge that the apple is red, or because that hypothesis would explain why it seems to me that this apple is red. In other words, that the apple is red explains my color experience and my color judgment, but I don't know that the apple is red by inference to the best explanation of my judgment or anything else. On the most plausible story about how I know this, I know it by perception: my visual experience represents the apple as red.⁸ Whether there is some inference to the best explanation of our color judgments seems irrelevant to the question of what justifies my judgment that the apple is red in the first place.⁹ More generally, the most obvious answer to the question of why we have reason to believe that things are colored is that we see that they are. It is not as though the ordinary person is justified in believing that things are colored only after she has shown that she can justify her belief in colors by an inference to the best explanation argument.

Moreover, it is plausible that, in order for a belief to amount to knowledge, there must be some kind of explanatory connection between its truth and its being believed. This requirement seems to hold with respect to both moral and empirical beliefs. If the fact that it is raining plays no role in the explanation for why I believe that it is raining, then this casts doubt on whether I know that it is raining, even if my belief is true. Similarly, if the oracle were to inform us that, whenever we believe that a moral proposition is true, the explanation for why we believe this is completely neutral with respect to whether the proposition *is* true, then we should be worried about whether we have any moral knowledge at all.¹⁰ In this respect, moral and empirical beliefs are on par.

⁸ Of course, there is also a tradition in epistemology that sees perceptual knowledge in general as in effect a special case of knowledge by inference to the best explanation. This tradition goes back at least to Locke's *Essay Concerning Human Understanding* (at least, on some interpretations of that text), and includes among others Russell in some of his writings on skepticism (see, e.g., Russell 1912), John L. (1976), Frank Jackson (1978), Laurence Bonjour (Bonjour and Sosa 2003) and Vogel (2012).

⁹ See McGrath 2004 for further discussion of this point. Harman himself discusses the case of color (1977, 13–14): his view is that even if color facts were explanatorily dispensable, “this would not prove that there are no facts about colors; it would only show that facts about colors are not additional facts, over and above physical and psychological facts.” In that case, we would conclude that color facts must be reducible to facts that “can help explain our observations” (14). Thus, ultimately, we are justified in believing that there are facts about colors by a reduction of color facts to facts belief in which can be justified by inference to the best explanation reasoning. While this might be a dialectically appropriate response to a color nihilist, my point is that our ordinary beliefs to the effect that things are colored are not justified by inference to the best explanation reasoning.

¹⁰ In saying this I am disagreeing with Dworkin's claim that moral facts differ from empirical facts in that the latter, but not the former, can play a role in explaining our beliefs, see McGrath 2014.

Thus, supposing that there are moral facts, we should agree with Sturgeon that these facts figure in perfectly good explanations, and that Harman hasn't given us any reason to doubt this. But insofar as Sturgeon thinks that Harman hasn't pointed to any important difference between moral facts and subatomic particle facts, we should disagree. If there are any moral features, they don't make a causal contribution all by themselves in the way that protons do. Therefore, inference to the best explanation gets a kind of traction against the proton skeptic that it doesn't get against the moral skeptic, but this is not bad news for moral realism. For it is not true, in general, that the only reason we have to countenance facts of a certain kind in virtue of their playing this very *special* kind of explanatory role. In other words, the first premise of the metaphysical challenge is false.

Now, in saying that the moral realist should respond to the metaphysical argument by rejecting the first premise, I am agreeing with the standard non-naturalist realist line. But I am offering a different rationale. I think that the first premise states a requirement that is too strong across the board: there are plenty of empirical facts that differ from proton facts in the same way that moral facts differ from proton facts. By contrast, non-naturalists think that while this requirement is appropriately applied with respect to some domains or some kinds of facts, it is not a requirement that can be deployed in argument for skepticism about morality.

There are at least two lines of argument that non-naturalist realists have offered for thinking P1 is too strong in this way. *Relaxed* non-naturalist realists think that the reason why we should reject the first premise is that there are different standards for knowledge in different domains. They think that in the natural world, we have reason to be committed to the existence of things of a certain sort only if they play the right kind of explanatory role. They think that moral facts are not the kind of facts that can explain anything: the requirement doesn't make sense in the moral domain. Thus, Scanlon argues:

In the domain of natural science . . . Harman's explanatory requirement makes good sense in this form: we have reason to be committed to the existence of things of a certain sort only if they play a role in explaining what happens in the natural world (including our experience of it). But this maxim is specific to the domain of natural science. It does not apply, as Harman's explanatory requirement is often held to apply, to every domain, for example to the normative domain, or to mathematics. (Scanlon 2014, 26)

Similarly, Dworkin (1996, 119) concedes that Harman's test, or something in the near vicinity, is appropriate for "beliefs about the physical world." But he denies that the test is appropriate for moral beliefs.

By contrast, the *robust* non-naturalist realist thinks that we should reject the first premise because we can have a different kind of reason to believe

that there are facts or entities of a certain kind—namely, that they play the right kind of *deliberative role*. Thus, David Enoch argues that in order to deliberate about what to do, we have to assume that there is an answer to the question of what we should do. That in turn requires assuming that there are normative facts. Thus, “deliberative indispensability . . . justifies belief in normative facts, just like the explanatory indispensability of theoretical entities like electrons justifies belief in electrons” (Enoch 2011, 50).

I have argued that we should reject the first premise for reasons that have nothing to do with the reasons that these non-naturalist realists have offered for rejecting it. To get on board with the companions-in-guilt case against P1 on offer here, you need not sign on to the view that there is an explanatory requirement that makes good sense as applied to the empirical domain, but not to morality, or that moral facts are deliberatively indispensable. You only need to agree that in general, a fact or feature can be explanatory even if it is not something that we know about by inference to the best explanation. Even if we can’t satisfy the moral nihilist by deploying an inference to the best explanation argument, moral facts are nevertheless in good company with other, non-moral, higher-level facts.¹¹ Thus, we neither have to posit distinct standards for knowing in the empirical versus moral domains, nor to insist that moral facts are indispensable to something *other* than explanations.

To summarize, Harman is right that there really is an important difference between Proton and Cat. The scientist knows that there is a proton by inference to the best explanation reasoning, but that’s not how Jim knows that what the kids are doing is wrong. Moreover, proton facts can be vindicated against skeptics by inference to the best explanation. So on that particular issue, I think that we should side with Harman. But we can take this position while rejecting the view that there is an “explanatory requirement” that makes sense in the domain of the natural sciences but not in ethics. Insofar as there is a true explanatory requirement, the requirement makes sense across domains: if the fact that *p* plays no role in explaining the belief that *p*, that belief won’t amount to knowledge. Thus, we can hold on to the view that moral facts explain our moral beliefs while rejecting the idea that the only way we get reasons to believe in things is by inference to the best explanation. This requirement is too strong not just in ethics but in the natural sciences as well. For there are many properties that (1) properly belong to the domain of the natural sciences but (2) are such that our reasons to believe in them don’t involve inference to the best explanation.

¹¹ Again, see Harman’s discussion of the case of color (1977, 14).

4 How the Realist Should Reply to the Epistemological Challenge

Let's now turn to the epistemological challenge.¹² If we do not follow Harman and Sturgeon in assuming that there is an especially close connection between confirmation and explanation, can moral beliefs be confirmed by empirical observation, in the way that empirical beliefs can be? If we assume a Bayesian picture, then the answer seems to be "yes." To a first approximation, for a person to make an observation that confirms one of her moral beliefs, she would satisfy the following conditions:

- (i) she makes an observation,
- (ii) she treats the content of that observation as a reason to increase her confidence in the moral view, and
- (iii) it is rational for her to treat the content of the observation in that way.

As a purely psychological matter, it seems clearly *possible* that a believer could distribute her prior probabilities in the relevant way. Presumably, the rationality proviso is essential: even if my psychology is such that I am hard-wired to increase my confidence that *the death penalty is intrinsically immoral* in response to observing that *there is a mouse in my kitchen*, actually observing a mouse in my kitchen would not confirm my belief about the death penalty. Thus, a theorist who wishes to deny that moral principles are ever confirmed by non-moral observations should argue that it would never be *rational* for a believer to distribute her credences in this way. That is, the theorist should deny that it would ever be rational for a believer to assign a higher probability to a moral view conditional on a non-moral observation that is higher than the probability that the believer assigns to the same moral view unconditionally.¹³ But it *can* be perfectly

¹² The discussion in this section overlaps with discussion in McGrath 2019, 120–127.

¹³ According to *Subjective* Bayesians, *any* prior probability distribution is rational, so long as it is probabilistically coherent (i.e., obeys the laws of the probability calculus). Given the truth of Subjective Bayesianism, the thesis that non-moral observations can confirm moral claims follows immediately, for it is clearly possible for a believer to have a coherent prior probability on which the conditional probability of some moral claim conditional on a certain non-moral observation is higher than the unconditional probability of the moral claim itself. However, many hold that Subjective Bayesianism is too liberal in allowing for confirmation, and that not every prior probability is rational. As White (2015) notes, the subjective Bayesian will hold that it can be perfectly rational for me, in advance of any observational evidence, to be extremely confident that when I look to the heavens I will see that the stars are aligned in such a way as to constitute a perfect representation of my face. Similarly, applied to the moral case, Subjective Bayesianism entails that the observation that there is a mouse in my kitchen *does* confirm or provide evidence for me *the death penalty is immoral*, provided that (i) my psychology is eccentric in the right way, and (ii) my degrees of belief are probabilistic coherent. I agree with the common thought that this approach makes empirical confirmation and rationality too easy to come by, in both the moral and non-moral domains. According to *Objective* Bayesians, not every coherent prior probability distribution is rational; rather, there

reasonable for a believer to assign a higher probability to a moral view conditional on a non-moral observation that is higher than the probability that the believer assigns to the same moral view unconditionally.

First, consider a testimonial case. Suppose that you initially find yourself uncertain about whether some moral claim is true; you thus assign a probability of n to the hypotheses, where n is some value significantly less than 1 but significantly more than 0.¹⁴ Suppose that you decide to ask a friend, whom you reasonably regard as having good moral judgment. Because you regard your friend as having good moral judgment, you assign probability $n + m$ to the moral claim conditional on your friend's judging that it is true (where $m > 0$). Your friend then tells you that she thinks the moral claim is true. In response, you increase your confidence in the claim. In doing so, you rationally treat a certain non-moral observation as evidence in favor of a moral claim.¹⁵

Thus, given modest assumptions about the way in which it can be rational to respond to empirical discoveries about the moral opinions of others, we should agree that moral beliefs can be confirmed or disconfirmed

are substantive constraints on prior probability distribution beyond the demand for purely formal coherence. Having noted that *if* Subjective Bayesianism is true, then the conclusion that this section aims to establish follows straightaway, I will assume that Subjective Bayesianism is false. The burden of the paragraphs that follow is to show that, even on a more objective picture of confirmation, we still have compelling reasons to accept that conclusion.

¹⁴ Here I assume, in the standard way, a scale on which probability 1 represents perfect certainty that a proposition is true and a probability of 0 represents perfect certainty that it is false.

I also assume that there are circumstances in which it would be reasonable to invest some level of credence in a moral view that is between 0 and 1, even if we restrict our attention to fundamental moral claims and assume that any such claim is metaphysically necessary if it is true at all. (Compare: even if the claim that *Water is H₂O* is a metaphysically necessary truth, there are many possible circumstances in which it is reasonable to be far from certain that it is true.)

¹⁵ This description of the case is not completely uncontroversial. On some views of the epistemology of testimony (see especially McDowell 1980), in paradigmatic cases one does not move (either consciously or unconsciously) from a belief of the form 'S said that p ' to a belief with the content p ; rather one directly takes up the content which is presented as true by the testifier. (For example, in a conversation in which you report that you were born in England, in the usual case I will simply add the proposition that you were born in England to my stock of beliefs, without inferring this from the belief that *you asserted that you were born in England*, or something similar.) If this model is applied to the moral case, then increasing one's confidence in a moral view on the basis of someone else's testimony would *not* involve a non-moral observation that has as its content the fact that someone engaged in a certain speech act; rather, one would simply adjust one's attitude toward the moral content directly. However, even if McDowell is right that this is the usual case of belief revision via testimony, it is obviously not the only case. For example, you might learn from the testimony of a third person that your trusted friend endorses the moral view of which you are uncertain. In that case, even on McDowell's view of testimony, the empirical evidence that you gain via testimony consists in a non-moral proposition. If McDowell's account is correct, then the case described in the main text can be amended in the relevant way, so that we once again have a (now more complicated) case in which a non-moral observation is rationally treated as evidence in favor of a moral belief.

by non-moral observations. Still, that leaves open the following question, which might be thought to be the question of greatest philosophical interest in the neighborhood: Are moral beliefs ever confirmed by non-moral observations, where the non-moral observations are not evidence about the moral opinions of others? Here is a hypothetical, but realistic case in which it's plausible to think that *disconfirmation* occurs:

The Story of Ted. At time t_0 , Ted believes that same-sex marriage is morally objectionable. His opposition to same-sex marriage involves both a belief that people of the same sex should not enter into such marriages given the opportunity to do so, as well as the belief that it is wrong for society to recognize such marriages, or to treat them as having equal standing to marriages between a man and a woman. Ted also believes that the recognition of same-sex marriages, and an increase in their number, would have numerous negative consequences. Some of these consequences he would characterize in explicitly moral terms, while others he would characterize in non-moral terms (e.g., “a significant increase in the divorce rate”). However, although Ted believes both that same-sex marriage is morally objectionable and that it would lead to consequences that he regards as bad, he does not believe that it is morally objectionable *because* it would lead to bad consequences, or that its (alleged) wrongness is in any way reducible to its having those consequences. Because of this, if he were asked at time t_0 whether he thinks that same-sex marriage would still be morally objectionable even if it did not have the bad consequences that he fully expects it to have, he would truthfully report his opinion as follows: “Yes, the wrongness of same-sex marriage does not consist in its having such-and-such consequences; rather, it is intrinsically immoral.” On the other hand, Ted does believe that there is the following connection between the (alleged) wrongness of same-sex marriage and the (alleged) fact that it would lead to negative outcomes for society as a whole; its moral wrongness is part of why it can be expected to lead to negative outcomes.

At this point, let me offer two remarks about The Story of Ted. First, there is nothing formally incoherent about Ted's views. Ted's views toward same-sex marriage are much like the views that many people hold toward lying. Many people hold the following beliefs about lying: (1) lying is wrong; (2) the wrongness of lying does not consist in, nor is it any way reducible to, facts about its consequences; (3) nevertheless, lying *does* tend to lead to bad consequences; and (4) the fact that lying tends to lead to bad

consequences is not independent of the fact that it is wrong. (Perhaps the last belief is in some cases due to a more general belief to the effect that actions that are morally wrong tend to give rise to bad consequences.)

Second, this combination of views toward same-sex marriage does not seem in any way unrealistic or even uncommon among people who oppose same-sex marriage. Indeed, most people who oppose same-sex marriage are not consequentialists, and thus would deny that its (alleged) wrongness in any way consists in or is reducible to its consequences. Nevertheless, they expect that same-sex marriage will lead to bad consequences, and that its having such consequences is intimately bound up with what they take to be its wrongness. (That is, it seems plausible that a significant number of actual people would accept the analogues of (1)-(4) in the case of same-sex marriage.)

Now consider:

The Story of Ted, continued. The society in which Ted lives grants full legal recognition to same-sex marriage and begins to treat those marriages as having equal standing with marriages between a man and a woman. The number of such marriages steadily increases. Ted retains his expectation that over time these developments will lead to (among other things) a significant increase in the overall rate of divorce within society. However, no increase in the divorce rate occurs, an observed fact that Ted acknowledges to be true. In response to the falsification of his prediction, Ted decreases his confidence in the views that led him to make that prediction, including the view that same-sex marriage is morally wrong.

It seems like it would be reasonable for Ted to revise his view about same-sex marriage in this way just as it is often reasonable to decrease your confidence in a non-moral view that leads you (typically in conjunction with other views) to make a prediction that is ultimately falsified. Indeed, given the description of the case, Ted's responding in this way seems at least as reasonable, if not more reasonable, than the alternatives (e.g., adopting a new sociological hypothesis that "explains away" the fact that the expected increase in the rate of divorce failed to materialize).

If this is the right verdict about the case of Ted, then it follows that moral beliefs can sometimes be *disconfirmed* by non-moral observations. From there, it is a short step to the conclusion that moral beliefs can sometimes be confirmed by non-moral observations as well. For, in a standard Bayesian framework, it is generally true that any belief that can in principle be disconfirmed by an observation can also in principle be confirmed by some (other) observation.¹⁶ In general, the possibility that reality could turn out

¹⁶ This is because a possible observation E disconfirms a hypothesis H just in case one's probability for H conditional on E is lower than one's unconditional probability for H (i.e.,

in such a way as to provide observational evidence in favor of a view that one holds is simply the flipside of the possibility that reality could turn out to provide observational evidence against that view.

Thus, the argument that there are possible circumstances in which a moral belief is confirmed by observational evidence runs as follows. If there are possible cases in which people have observational evidence that disconfirms their moral beliefs, then there will also be possible circumstances in which people have observational evidence that confirms their moral beliefs. But there are possible cases in which people receive evidence that disconfirms their moral beliefs (as witnessed by *The Story of Ted*). Therefore, there are possible cases in which people receive observational evidence that confirms their moral views.

Even a relatively fundamental moral conviction (of a kind whose truth or falsity seems maximally independent of any contingent matters of fact) can give rise to expectations about how the world will turn out to be when it is conjoined with other views. When those expectations are disappointed, it can make sense to treat that outcome as evidence against the views that gave rise to them. When those expectations are fulfilled, it can make sense to treat that outcome as evidence in favor of those views.¹⁷

I have offered the *Story of Ted* as a case in which a moral belief is confirmed by non-testimonial, non-moral observation. But I want to emphasize that Sturgeon's own examples can be used to make the same point. Someone in the eighteenth or nineteenth century might have thought that the forms of slavery currently practiced were morally worse than the kinds of slavery practiced before, and have predicted that morally worse forms of slavery would lead to relatively vigorous opposition. When such opposition arose, fulfilling her expectations, she might rationally increase her confidence that the form of slavery currently being practiced was morally worse. *The Story of Ted* has the same basic structure. But my point here is that, within a Bayesian framework, moral claims can be confirmed and disconfirmed by non-moral observations, irrespective of whether we can

$\Pr(H/E) < \Pr(H)$); given that one's unconditional probability for H is the weighted average of one's probability for H on E ($\Pr(H/E)$) and one's probability for H on not- E ($\Pr(H/\text{not-}E)$), it follows that $\Pr(H/E) < \Pr(H)$ only if $\Pr(H) < \Pr(H/\text{not-}E)$ —that is, only if the probability of H conditional on not- E is higher than the unconditional probability of H . But then, it follows that observing not- E would *confirm* H , since in a Bayesian framework, confirmation consists in probability-raising. Although it is generally true that any belief that can in principle be disconfirmed by an observation can also in principle be confirmed by an observation, there are deviant cases where this fails to hold. For example, the hypothesis “Nothing exists” is falsified by any possible observation, and is thus a claim that can be disconfirmed even though it cannot be confirmed. Thanks to Yoav Isaacs, via Tom Kelly, for this example.

¹⁷ Cf. the holistic picture of confirmation famously sketched in Quine 1951, sec. 6—although it seems clear that Quine himself had no sympathy for extending that picture to the moral domain. (On this last point, see especially Quine 1979). I want to emphasize, however, that one can accept the kind of holism suggested here without endorsing some of Quine's more extreme claims, for example, that particular observations test “the totality of our so-called knowledge or beliefs” (Quine 1961, 42).

infer the existence of moral features, based on their explanatory role. Even non-naturalist realists who think that moral facts are necessary, causally inefficacious, and knowable *a priori* can agree that, assuming that there *are* moral truths, they are subject to empirical confirmation in the way that scientific claims are.

To be clear, even if moral claims can be confirmed by empirical observation in the same way that non-moral claims can be, this doesn't amount to any kind of response to the nihilist that we met in [Section 3](#). The Bayesian has little interest in addressing any kind of skeptic, moral or otherwise. For the Bayesian, some of us give relatively high credence to claims that require moral properties—and others don't. The Bayesian does not take sides on the issue. Inference to the best explanation, I have argued, gives us little traction against the moral nihilist. But Bayesianism is not even in the game of offering traction against skeptics in the first place.

5 Conclusion

I have argued that Harman's "Ethics and Observation" raises two distinct challenges: one has a metaphysical conclusion, the other has an epistemological conclusion. In closing, I will summarize what I take to be the lessons that emerge from separating the challenges.

By separating the two challenges, we can see that the question of how to respond to the nihilist is a different question from the question of how we know which moral claims are true. The question of whether Hitler's deeds confirm his cruelty can be separated from the question of whether we can know by inference to the best explanation of Hitler's deeds that there are moral facts. The answer to the question, "can moral principles be confirmed by observation?" is a resounding "yes"; the answer to the question, "do we know that there are moral facts by inference to the best explanation?" is "no."

Thus, separating the two challenges can also help us to resist a certain kind of motivation for the view that we need a special story about where moral knowledge comes from. Realists can agree that there is an important contrast between Proton and Cat without signing on to the view that when it comes to moral knowledge, ordinary ways of acquiring empirical knowledge won't do. Non-naturalist realists have sometimes motivated the idea that moral knowledge would have to emerge by piecemeal application of "the method of reflective equilibrium" by something like the following line of reasoning:

Harman's Challenge mistakenly applies an epistemological standard to the moral domain that is in fact only appropriately applied to the domain of the natural sciences. We believe in subatomic particles because they make a difference to what we observe, but our moral beliefs are not

shaped by the impact of “moral particles.” There are no “morons” that make a difference to what we observe. That is not bad news for moral knowledge, because the moral domain has its own epistemological methods and standards for justification.

In fact, the contrast between PROTON and CAT does not offer any motivation for thinking that the moral domain requires its own distinct methods of investigation. We don’t infer the existence of moral features in the same way that we infer the existence of subatomic particles, but in this way, moral facts and features are in good company with many other kinds of facts and features, including color facts and features. But from the fact that we don’t have to posit “moral particles” to explain anything, it doesn’t follow that there isn’t empirical confirmation of moral claims, or that we need a special story about where moral knowledge comes from.

Sarah McGrath
Princeton University
E-mail: smcgrath@princeton.edu

References:

- BonJour, Laurence and Ernest Sosa. 2003. *Epistemic Justification: Internalism vs. Externalism, Foundations Vs Virtues*. Malden, MA: Wiley-Blackwell.
- Boyd, Richard. 1988. “How to be a Moral Realist.” In *Essays in Moral Realism*, edited by Geoffrey Sayre-McCord. Ithaca, NY: Cornell University Press.
- Chang, Hasok. 2010. “The Hidden History of Phlogiston.” *International Journal for Philosophy of Chemistry* 16 (2): 47–79.
- Dworkin, Ronald. 1996. “Objectivity and Truth: You’d Better Believe It.” *Philosophy and Public Affairs* 25 (2): 87–139.
- Dworkin, Ronald. 2011. *Justice for Hedgehogs*. Cambridge, MA: Harvard University Press.
- Easwaran, Kenny. 2011. “Bayesianism I: Introduction and Arguments in Favor.” *Philosophy Compass* 6 (5): 312–320.
- Enoch, David. 2011. *Taking Morality Seriously*. Oxford: Oxford University Press.
- Harman, Gilbert. 1965. “The Inference to the Best Explanation.” *Philosophical Review* 74 (1): 88–95.
- Harman, Gilbert. 1968a. “Enumerative Induction as Inference to the Best Explanation.” *The Journal of Philosophy* 65 (18): 529–533.
- Harman, Gilbert. 1968b. “Knowledge, Inference, and Explanation.” *American Philosophical Quarterly* 5 (3): 164–173.
- Harman, Gilbert. 1977. *The Nature of Morality: An Introduction to Ethics*. Oxford: Oxford University Press.
- Harman, Gilbert and Judith Jarvis Thomson. 1996. *Moral Relativism and Moral Objectivity*. Cambridge, MA: Blackwell.
- Horwich, Paul. 1982. *Probability and Evidence*. Cambridge: Cambridge University Press.
- Jackson, Frank. 1978. *Perception: A Representative Theory*. Cambridge: Cambridge University Press.
- Mackie, John L. 1976. *Problems from Locke*. Oxford: Oxford University Press.
- McDowell, John. 1980. “Meaning, Communication, and Knowledge.” Oxford: Clarendon Press.
- McGrath, Sarah. 2004. “Moral Knowledge by Perception.” *Philosophical Perspectives* 18 (1): 209–228.

- McGrath, Sarah. 2014. "Relax? Don't Do It! Why Moral Realism Won't Come Cheap." In *Oxford Studies in Metaethics*, Vol. 9, 186–214.
- McGrath, Sarah. 2019. *Moral Knowledge*. Oxford: Oxford University Press.
- Quine, Willard Van Orman. 1951. "Two Dogmas of Empiricism." *The Philosophical Review* 60 (1): 20–43.
- Quine, Willard Van Orman. 1961. *From a Logical Point of View*. Cambridge, MA: Harvard University Press.
- Quine, Willard Van Orman. 1979. "On the Nature of Moral Values." *Critical Inquiry* 5 (3): 471–480.
- Railton, Peter. 2003/1986. *Facts, Values, and Norms: Essays Toward a Morality of Consequence*. New York: Cambridge University Press.
- Russell, Bertrand. 1912. *The Problems of Philosophy*. Oxford: Oxford University Press.
- Scanlon, Thomas M. 2014. *Being Realistic About Reasons*. Oxford University Press.
- Schechter, Joshua. 2017. "Explanatory Challenges in Metaethics." In *The Routledge Handbook of Metaethics*, edited by Tristram McPherson and David Plunkett. New York: Routledge.
- Sturgeon, Nicholas. 2001/1985. "Moral Explanations." *Moral Philosophy* 212–232.
- Vogel, Jonathan. 2012. "The Refutation of Skepticism." In *Contemporary Debates in Epistemology*, edited by Matthias Steup, John Turri, and Ernst Sosa, Vol. 2. Malden, MA: Blackwell.
- White, Roger. 2015. "The Problem of Induction." *Episteme* 12 (2): 275–290.