Explanationism vs Modalism in Debunking (and theory choice)

Harjit Bhogal

Consider this line of thought: There seems to be a striking correlation between our moral beliefs and the moral truths. Given a robustly realist conception of morality (paradigmatically, Platonist non-naturalist realism) the correlation between our beliefs and the truth seems to be a coincidence. And this gives us reason to reject the combination of robust realism and the correlation between our moral beliefs and the moral truth. This line of thought is, of course, at the heart of many formulations of *debunking* arguments against robust realism.

There are two main strands of response to this thought in the literature. Firstly, there are *third-factor* responses, which, roughly speaking, claim that, even given robust realism, there is a common explainer of our moral beliefs and the moral truths, so the correlation between then isn't coincidental (e.g. Enoch, 2011; Skarsaune, 2011; Wielenberg, 2010, among many others). Let's put this response aside to focus on a different one in this paper.

Perhaps the most influential type of response in the recent literature, and what I'll focus on, are *Modalist* responses. Such responses accept that there is no explanation of the correlation — there is no explanatory connection between belief and truth on the robust realist view. But, they argue, the correlation between moral belief and moral truth has certain important modal features — notably that the correlation could not have easily been false — and these make the correlation, in a sense, unproblematic. It's common, in particular, to emphasize that our moral beliefs are *safe* and *sensitive* — and what more could we want from our moral beliefs than that? (Justin Clarke-Doane has been

most influential in developing this position in a series of papers. See, for example, Clarke-Doane (2016a,b, 2020) and Clarke-Doane and Baras (2019). Though, as Schafer (2017) and Faraci (2018) note, ideas in this modalist spirit run through a lot of the literature on debunking.)[1]

(Modalism, as I'm understanding it, accepts, at least for the sake of argument, that there is no explanation of the correlation, but claims that this does not debunk our moral beliefs. But you could also have a view, somewhat related to modalism, which claims that there is an explanation of the correlation, just an extremely minimal one (see Pust (2005) and Baras (2017)). For example, one could argue that merely giving a causal explanation of why we have the moral beliefs that we have, and noting that those beliefs are necessarily true, counts as an explanation of the correlation. This type of view won't be the focus of this paper, but as we will see in section 3.1, the core argument of this paper does give reasons to be doubtful of it.)

Explanationists, on the other hand, claim that we could, and do, want more than safety and sensitivity – we want an explanatory connection between belief and truth. Even if the correlation between belief and truth could not easily be false and even if our moral beliefs are safe and sensitive then the lack of such an explanation undermines our beliefs. (See, for example, Korman and Locke (2020a), Lutz (2020), Faraci (2019) among others.)

There's something of a standoff here. Explanationists claim that given robust realism our moral beliefs are explanatorily disconnected from the truth — there is no appropriate explanation of the correlation between belief and truth. And, they say, that's a problem – the lack of explanatory connection is a reason to be suspicious of our moral beliefs or to reject robust realism. Modalists reply that, sure, there is no explanatory connection, but this isn't a problem since our beliefs have the appropriate modal characteristics with respect to the truth.

This standoff is central to the modern debate about debunking. Let me put my cards on the table:

[1] Another type of response to debunking arguments, that I'm putting aside here, is what Korman (2019, section 5.1) calls *unstable minimalist responses*. The core idea is to accept that the correlation between our beliefs and truth is, given robust realism, a lucky coincidence, but there is good reason to believe that we were, in fact, lucky in that way. (See, for example, Dworkin (1996) and Setiya (2012, chapter 2).) This type of response doesn't address the form of debunking argument that I discuss in section 2 and onwards but I won't discuss this further here.

I don't think third-factor responses to debunking arguments work (though, as I noted, that won't be a focus of this paper). So, I think, if the explanationist is right there is a good debunking argument against robustly realist views of morality. And if the modalist is right there is not.

I'm going to argue that the explanationist is right. Or, more carefully, there is a version of the debunking argument that the explanationist is right about, and consequently the explanationist can give a powerful argument against robust realism. Notably, though, this version of the debunking argument differs in some important ways from versions that have received the most attention in the recent literature.

In particular, I'm going to argue that if we see debunking arguments as an instance of a broader class of arguments about how we choose between theories and in the process, deemphasize issues of whether our moral beliefs are justified, then this allows us to break the standoff between explanationism and modalism, whilst still providing an argument against robust realism.

Understanding the debunking argument in this way allows us to look at cases from science which are analogous to the moral case, seeing whether it is explanatory or modal considerations that are the driving factors behind theory choice in such cases. This, I'm going to claim, favors the explanationist.

1 Modalism and Explanationism

The place to start, though, is by looking in more detail at modalist and explanationist approaches to debunking. Exactly what an explanationist or a modalist will say will depend on the precise formulation of the debunking argument that is under consideration, and we certainly don't have time to survey all the different formulations. The goal in this section, then, will be to understand the spirit of the positions and then, in the next section, to see how they apply to the type of debunking argument that I'm focused on.

Most debunking arguments in the literature aim to argue that, given a robustly realist conception of morality our moral beliefs are not justified. The paradigm robustly realist position is a type of

non-naturalist realism where moral facts are mind-independent, non-natural, and causally inert. (Following much of the literature I will focus on this paradigm robustly realist position in what's to come. How debunking arguments apply to other versions of realism, or even to positions like constructivism, is a complicated question that I'm not going to deal with here. (See, for example, Bogardus (2016), Joyce (2016), Tropman (2014).))

The intuitive thought is that there is no 'connection' between our moral beliefs and the moral truths. The forces that lead us to our moral beliefs, in particular, the evolutionary forces, are not related to the actual moral truths — our moral beliefs don't explain the moral truths and the moral truths don't seem to explain our moral beliefs.[2] And the recognition that our beliefs are disconnected from the truth seems to undermine the justification in those beliefs.

The modalist response to such arguments doesn't dispute that there is a sense in which our moral beliefs are explanatorily disconnected from the truth, but argues that nevertheless, such beliefs have the modal characteristics required for justification. Or, perhaps more intuitively, even though our moral beliefs are explanatorily disconnected from the truth they are still reliable.

The debate over this type of response gets very technical very quickly — there are lots of different formulations of the relevant modal conditions and much discussion of the different ways in which justification can be defeated. But I'm just going to, in a very rough and ready way, describe the main thrust of the debate, leaving out a huge amount of nuance.

The modalist often argues that, given robust realism, our moral beliefs are *safe* and *sensitive*. Let's use the formulations from Clarke-Doane and Baras (2019):

Sensitivity: Our belief that P is sensitive iff had it been that ¬P, we would not still have believed that P, had we used the method that we actually used to determine whether P.

Safety: Our belief that P is safe iff we could not have easily had a false belief as to whether or not Q, where Q is any proposition similar enough to P, using the method that we actually used to determine

[2]Again, third-factor responders say that the connection between our beliefs and the truth consists in a common explainer, and we are putting this aside in this paper.

whether or not P.

Consider, for example, my belief that murder is wrong. Could I have easily had a false belief about whether murder is wrong (or about some similar proposition)? It seems not because moral truths are necessary, so murder is wrong in all possible worlds, and it seems that I could not easily have believed that murder is not wrong — I would have to have been an extremely different person to believe that murder is acceptable. So my belief appears to be safe. (See, for example, Clarke-Doane (2015, section 4.5). But as he notes, the idea that we could not easily have failed to hold at least some of our moral beliefs follows from a point commonly emphasized by defenders of debunking arguments (e.g. Street (2006)) – that evolution provides a robust explanation of our moral beliefs.)

Similarly, my moral belief seems to be sensitive. My belief is sensitive if, had it been the case that murder was not wrong, I would have not believed that murder is wrong (had I used the method that I actually used to determine whether murder is wrong). This counterfactual is normally taken to be trivially true, since there it could not have been the case that murder was not wrong — it is necessary that murder is wrong. (This kind of point goes back to Lewis (1986, section 2.4).) So, my belief is sensitive.[3]

And given that such moral beliefs are safe and sensitive, even given robust realism, then it's hard to see how debunking arguments could undermine our beliefs. When our beliefs are sensitive and safe it's natural to say that they are *reliable*. And once we have reliable beliefs what more could we want? In particular, if we know that our beliefs are safe and sensitive then it seems like those beliefs are not undermined – we are justified in having them. This type of of idea – that for our beliefs to be undermined we need reason to think that they are not safe or sensitive is called *Modal Security* in much of Clarke-Doane's work. (See Clarke-Doane and Baras (2019) for extensive discussion of Modal Security and the large recent literature discussing it.)

Clarke-Doane and Baras note, however, that many people find this line of argument extremely un-

[3]Whether such counterfactuals with metaphysically impossible antecedents are trivially true is controversial. See, for example, Berto and Jago (2018, esp. sections 2.5 and 6.5) for a survey. But I will assume, in the modalists favor, that they are.

intuitive. The worry is that it 'implies that one's beliefs can be secure from undermining, even upon learning that they bear no connection to the truth' (p.180).

In particular, explanationists, like Korman and Locke (2020b), Lutz (2020, 2018) and Faraci (2019), reject the modalist position. As Clarke-Doane and Baras put it, these explanationists claim that 'if one learns that there is no "explanatory" connection between the belief and the truth, then that in itself is undermining — regardless of the modal security of the belief' (p.180).

One way to put the disagreement between the explanationist and the modalist is to note that in most ordinary cases of belief we have both the relevant modal conditions and an explanatory connection between belief and truth. When I believe that there is a computer in front of me on the basis of my visual perception, my belief is safe and sensitive. And my belief is explained, in part, by the truth — the fact that there is a computer in front of me. In fact, we might think that in this case my beliefs are safe and sensitive because of the way that are explained by the facts. Perhaps the explanatory connection between my belief and the computer is what makes it modally responsive in the way safety and sensitivity require.

In the moral case, though, these modal and explanatory conditions seem to come apart — our beliefs are safe and sensitive but there is no explanatory connection. Modalists say that since the relevant modal conditions hold we are justified in our moral beliefs. Explanationists say that since the there is no explanatory connection between belief and truth we are not justified in our moral beliefs.

In fact, Clarke-Doane and Baras (section 10) say that the explanationist is misled by the fact that in ordinary cases the modal and the explanatory conditions go together. They say that an explanatory connection is predictive of safety and sensitivity in many cases, but it's really safety and sensitivity that is epistemically valuable. The explanationist, they say, confuses a proxy for what's valuable with the valuable thing itself. But of course, the explanationist could accuse the modalist of a similar mistake. Adjudicating between these intuitions seems rather hard.

The most natural attempt to adjudicate is to look at other cases where the modal and explanatory conditions come apart to see whether our beliefs are intuitively justified in such cases. Faraci (2019,

pp. 12-13) for example, gives such a case that he claims favors the explanationist. The problem is that it is extremely hard to formulate cases which meet the conditions of sensitivity and safety without an explanatory connection. The resulting cases are very strange, so that it's really not clear whether the agents in this case are justified, and further, it's not even clear that the cases are coherent. Here is Faraci's case:

> Eula is defeasibly justified in forming beliefs about which numbers are prime by consulting the Source, and there is no available evidence that the Source is untrustworthy. In fact, the Source is outputting numbers at random. Eula consults the Source to form beliefs about which numbers are prime. The numbers the Source outputs at random are all prime numbers. Eula's resultant beliefs are therefore true; and there is no good evidence to the contrary. At every possible world, Eula's counterpart forms beliefs about which numbers are prime by consulting the Source's counterpart, and at every possible world, the Source's counterpart delivers the same answers as at the actual world. (p.13)

Such a case would be one where Eula's beliefs are safe and sensitive without an explanatory connection between belief and truth. And it seems clear in this case that learning about the lack of explanatory connection – that the Source is outputting numbers at random – undermines Eula's beliefs. But it's not clear that this case is coherent. In particular, why should we believe that 'At every possible world, Eula's counterpart forms beliefs about which numbers are prime by consulting the Source's counterpart, and at every possible world, the Source's counterpart delivers the same answers as at the actual world'? As Clarke-Doane and Baras (p.175) note, such claims about metaphysical necessity can't just be stipulated as part of a case.

There's a lot more we could say about such cases, but the point I want to make now is just that it's hard to use them to clearly adjudicate between explanationism and modalism.

2 Debunking and Theory Choice

The way to make progress on this question of modalism vs explanationism, I think, is to move away from the type of debunking argument we have just been considering. So far we have been considering debunking arguments where the conclusion is supposed to be that we lack justification in our moral beliefs, given robust realism. Perhaps there is a stalemate between the modalist and the explanationist with respect to this argument. But if we consider a slightly different form of debunking argument then a range of new considerations open up for us.

In particular, it's fruitful, I think, to construe debunking arguments as an instance of more general considerations about theory choice, rather than being an argument that is narrowly about the nature of justification or the reliability of our beliefs. Consider, for example, this argument:

(1) On robustly realist conceptions of morality the correlation between our actual moral beliefs and the moral truths is a big coincidence.
(2) There is reason to reject theories that lead to big coincidences.
So,
(3) There is reason to reject either the correlation between our moral beliefs and the truth, or the robustly realist conception of morality.

This argument is very much in the spirit of Field's (1989) argument against mathematical realism.

This type of debunking argument is an instance of a very general consideration — when we have a striking correlation and some theory which implies that the striking correlation is a coincidence, then that's reason to reject the theory or to reject the correlation. In the moral case under investigation the striking correlation is between our beliefs and the truth. But, on this way of understanding the issue, that's rather incidental. The correlation could be between all sorts of things and an argument of this form would still apply.

Imagine a friend told me that she tossed a coin 50 times and it landed heads every time. The theory that it's a fair coin that she tossed normally implies that it's just a big coincidence that the coin landed

heads every time. This should lead us either to reject the theory — to think that it's not a fair coin — or to reject the correlation — to think that my friend was lying and the coin didn't land heads every time.

This form of argument plays an important role in scientific practice too. To take a very simple example, objects tend to fall towards earth. If a theory implied that this correlation in the motion of objects was just a big coincidence then that theory should be rejected. To take a less trivial example, Bhogal (2020, pp. 677-8) discusses the 'cosmic coincidence problem': roughly speaking, the observation that the amount of energy in the universe that comes from dark energy is almost the same as the amount of energy that comes from dark matter. Our current best cosmological theories seem to say that this is a coincidence though — at other times in the universe this equality does not hold. And he notes that some physicists react to this by rejecting the theories that imply that this equality of energy is a coincidence, while other physicists are inclined to reject the data which suggest that there is equality between the two types of energy. Either way, there is pressure to avoid the coincidence.

We will talk about other scientific cases soon. But to be clear, the point is not that there can never be any coincidences, or that when a theory implies that a correlation is a coincidence that's a conclusive reason to reject the theory. It's just that coincidences that are sufficiently big can give us reasons to reject a theory, and sometimes those reasons can be rather powerful.

In discussing this argument I've been appealing to an intuitive notion of coincidence, and I will continue to do that. But it will be useful to say just a little more about the notion of coincidence at work here. For our purposes a coincidence is, roughly, a striking correlation or matching between facts that does not have an explanation (see, for example, Hart and Honoré (1985, p. 74), Lando (2017), Bhogal (2020) and the wide range of views that Lando (section 2) labels as 'traditional views').

This is only rough. To be more precise we would say more about what a striking correlation is. Some judgments are clear — if a coin lands heads 50 times in a row that's striking, but if the coin lands HTHHHTTHHTTHTTT…that is not striking — but giving a full account of strikingness is not

at all easy (for discussion see, for example, Horwich (1982, chapter 7), Schlesinger (1987), White (2005) Wong and Yudell (2015), Baras (Baras)). But luckily it won't matter much for our purposes, since everyone agrees that the matching between our moral belief and the moral truths is striking. And further, to be more precise about the notion of coincidence we would have to say more about exactly what it is for the matching between facts not to have an explanation. It turns out that there are a lot of complicated issues here (Lando, 2017; Bhogal, 2020). These complicated issues would be very important if we were discussing third-factor responses, but given our focus is on explanationism vs modalism the details don't matter here. (Though related issues will be discussed in section 3.1.)

2.1 The Modalist response to (1)-(3)

Again, the argument (1)-(3) has a different form from many of the debunking arguments that have received the most attention in the recent literature[4] — it's not distinctively about belief or justification or knowledge. But this shift doesn't, of course, mean that the modalist has to give up and accept the debunking argument. A modalist can respond to (1)-(3) in a way that is closely analogous to their response to other version of the argument. In this section we will consider how that response will go.

Notice that, the argument (1)-(3) embodies a kind of explanationist position. It says that we have reason to reject the combination of the correlation between our moral beliefs and the truth and the robustly realist conception of morality because that would make the correlation a coincidence. And the correlation being a coincidence is an *explanatory* deficiency — it's that the striking correlation is unexplained.

The modalist move we discussed in the last section is to argue that even though moral belief and truth seem to be explanatorily disconnected — that is, the correlation between them seems unexplained — this isn't epistemically important. Rather, what is important is that our moral beliefs are safe and sensitive.

[4]Though certainly this type of argument has received attention, e.g. Field (1989), Enoch (2010), Schechter (2010).

How can the modalist make this kind of move with respect to (1)-(3)? Well, with respect to (1)-(3) the explanationist claims that the fact that the correlation between belief and truth is not explained given robust realism makes the combination of the correlation and robust realism problematic — we have reason to reject one or the other. The modalist move will be to accept that the correlation is not explained, but then to claim that this doesn't provide a reason to reject the correlation or robust realism, because given robust realism, the correlation belief and truth has certain appropriate modal characteristics.

This move is, in effect, a rejection of premise (2). The notion of coincidence we are using is an explanatory one — again, a coincidence, for our purposes, is a striking correlation that is unexplained. The modalist response to (1)-(3) — analogous to their responses to other debunking arguments — is that we should be happy to accept theories that lead to big unexplained correlations, as long as the theory implies that the correlations have appropriate modal features.

2.2 Appropriate Modal Features

There is, though, an obvious question: Just what are these appropriate modal features?

As we discussed, the standard modalist move is that it is the safety and sensitivity of our beliefs that make them unproblematic. But as we noted in the last section our version of the debunking argument is not distinctively about belief.[5] Rather, it's an instance of a more general argument that when we have a striking correlation and some theory which implies that the striking correlation is a coincidence, then that's reason to reject the theory or to reject the correlation. This argument applies to correlations that have nothing to do with belief.

The issue is that safety and sensitivity *are* distinctively about belief. So what the modalist needs are generalizations or analogues of safety and sensitivity which apply not just to the correlation between our beliefs and the truth, but to correlations or matchings between facts more generally. Roughly speaking, the idea of safety is that our beliefs could not easily have been false — that is, the correlation

[5] Though, of course, the argument does have implications for what we should believe.

between our beliefs and the truth could not have easily failed to hold. And the idea of sensitivity is that if the facts had been different our beliefs would have been different.

It's easy enough to see how to extend these ideas to correlations more generally. Take, for example, a correlation or matching between fact A and fact B.

Safety*: The correlation is safe* if and only if it could not easily have failed to hold, that is, if A could not easily have failed to match B.

Sensitivity*: The correlation is sensitive* if and only if had it been the case that one side of the correlation did not hold the other would not have held. That if, if A did not hold then B would not have held and vice versa.

Again, the modalist avoids the argument (1)-(3) by claiming that we should be happy to accept theories that lead to big unexplained correlations, as long as the theory implies that the correlations have appropriate modal features. And the natural candidates for these appropriate modal features are safety* and sensitivity*.

3 Protons and Electrons

Perhaps it may seem like this shift to debunking arguments like (1)-(3) hasn't helped us at all. There is still, it seems, a standoff between the explanationist and the modalist. The explanationist says that the correlation between moral belief and truth is unexplained given robust realism, and that's a problem with either the correlation or robust realism. The modalist replies that it's not a problem because, given robust realism, the correlation has the appropriate modal features — in particular safety* and sensitivity*.

But moving to arguments like (1)-(3) in fact opens up a range of new cases that we can use to evaluate explanationism and modalism. In particular, since (1)-(3) is an instance of general considerations about theory choice and which are not distinctively about belief then the cases we can appeal to to

evaluate (1)-(3) don't have to be correlations between belief and truth. In this section I'm going to appeal to a scientific case that, I will argue, ends up favoring the explanationist.

The case is one where there is a striking correlation, and a theory, given which the correlation is safe* and sensitive*. So, the modalist should say that the combination of the theory and the correlation is unproblematic. But, nevertheless, it seems clear that we should reject either the correlation or the theory, because the correlation would be unexplained given the theory.

I'll focus on a simple case, but later I'll point toward how it's illustrative of more complicated cases which are currently under dispute in scientific practice. Here is the case:

Protons and Electrons Protons are positively charged. Electrons are negatively charged. However, the absolute value of their charge is the same — call this fact the *charge correlation*. Specifically, protons have a charge of $1.602176634{\times}10^{-19}$ coulombs, while electrons have a charge of $-1.602176634{\times}10^{-19}$ coulombs.

The charge correlation is very striking. Consider a theory, T, that implies that it is just a complete coincidence that the charges are the same – that there is no explanation of the matching. This seems like a big problem for T. Perhaps it wouldn't be a conclusive reason to reject the theory T – if T is successful enough it might be justified to still accept it – but it's a bad-making feature of the theory.

Now let's imagine what such a theory, T, might look like. Let T state that it is a basic law of nature that protons have a charge of $1.602176634{\times}10^{-19}$ coulombs, and also it is a basic law of nature that electrons have a charge of $-1.602176634{\times}10^{-19}$ coulombs. And T has nothing more to say about why the particles have these charges.Importantly, given T the charge correlation is nomically necessary.

Such a case, I claim, provides an argument against the modalist. Given T there is no explanation of the charge correlation but the correlation seems to be safe* and sensitive*. We still have reason, however, to reject either the charge correlation or T. So, it is the explanatory factors — the way in which, given T, the charge of the proton and electron are explanatorily disconnected — that are relevant for theory choice, and not the modal features of the correlation. It is the explanationist who

is right about arguments of the form (1)-(3), and not the modalist.

The best way to explain and defend this argument is, I think, by considering some possible modalist responses. I'll do that in the rest of the section. So how might a modalist respond? Firstly, they could claim that there is an explanation of the correlation. Secondly, they could claim that the correlation is not safe* or sensitive*. Thirdly, they could claim that we do not, in fact, have reason to reject the combination of the correlation and T. Let's consider these possible responses in turn.

3.1 No Explanation

Let's start with the claim that given theory T there is no explanation of the correlation between the charges. This, I take it, is fairly intuitive — after all, T doesn't seem to provide any connection between the charge of the proton and the electron. But there is a natural way that someone might try to deny it. They might claim that charge correlation is explained as follows: The fact that it's a basic law that electrons have a charge of $-1.602176634{\times}10^{-19}$ coulombs explains why electrons have a charge of $-1.602176634{\times}10^{-19}$ coulombs. The fact the it's a basic law that protons have a charge of $1.602176634{\times}10^{-19}$ coulombs explains why protons have a charge of $1.602176634{\times}10^{-19}$ coulombs. $-1.602176634{\times}10^{-19}$ and $1.602176634{\times}10^{-19}$ have the same absolute value. All this together explains why the absolute value of the charges of protons and electrons match.

I'll make three points about this strategy.

Firstly, claiming that there is an explanation of the charge correlation isn't really a way for the modalist to defend their position. The result that we want is that we should reject theory T. But, of course, claiming that theory T does in fact explain the charge correlation doesn't help us get this result. What this is, rather, is a criticism of the explanationist, saying that they face problems with the protons and electrons case too.

(An aside: In the introduction I mentioned the view that, contra the modalist – there is an explanation of the correlation between our moral beliefs and the moral truths, given robust realism — a

very minimal explanation (Pust, 2005; Baras, 2017). And, consequently, this is a reason to be happy accepting both the correlation and robust realism. We can now see a major problem with that strategy — the existence of a similar minimal explanation of the charge correlation, analogous to those suggested by Pust and Baras, should not make us happy to accept both the charge correlation and theory T, so it's hard to see why it should in the moral case. There is a lot more to say about this strategy, but that will have to be for another time.)

Secondly, not withstanding the previous parenthetical, the focus of this paper is on the modalist strategy of accepting that there is no explanatory connection between our moral beliefs and the moral truths but claiming that this doesn't matter, since our beliefs have the relevant modal features with respect to truth. Consequently, it's fair for us to assume that the modalist will similarly take there to be no explanation with respect to the charge correlation.

Thirdly, it's intuitively clear that the proposed explanation of the charge correlation, given T, is deeply unsatisfying and explanatorily deficient. This argument doesn't really explain the matching between the charges, rather it merely separately explains the precise value of the charge of the proton, and the precise value of the charge of the neutron. Similarly, separately explaining why each coin tossed landed heads, by giving the microphysical details of exactly how each coin was tossed, is not a satisfying explanation of why the coin landed heads every time my friend tossed it. Consequently, the explanationist should be happy to accept that this proposed 'explanation' doesn't count as an explanation of the charge correlation. (Or, alternatively, they could accept that it does, strictly speaking count as an explanation, but even in light of this type of explanation the charge correlation is still a huge coincidence.) Making precise the exact sense in which the proposed explanation is deficient is an interesting and complicated task.[6] Going into it further would take us too deep into the literature on scientific explanation. But still, it's easy to recognize that the proposed explanation is deficient.

For these reasons, claiming that the charge correlation has an explanation, given T, is not a promising

[6] See Field (1996, section V), Lange (2010), Tersman (2016, section 3), Baras (2017), Faraci (2019), Bhogal (2020) and Baras (Baras, chapter 4) for discussion.

way for the modalist to defend themselves.

3.2 Modal Characteristics

Let's consider the next possible response. Given T, is the correlation between the charge of the proton and electron safe* and sensitive*?

3.2.1 Safety*

Let's start with safety*. It's pretty clear, I think, that the correlation is safe*. T implies that it is nomically necessary that the absolute value of the charge of the electron matches that of the proton — the correlation is guaranteed by the laws of nature. In this clear sense, then, it could not easily have failed to hold, because that would require the laws to have been different.

The view that the laws of nature are modally robust and so could not easily have been false is, I take it, the standard, mainstream position. For example, it's standardly taken to be the case that a possible world that contains widespread violations of the actual laws of nature is, in virtue of that, very distant and dissimilar from the actual world (Lewis, 1979). And this distance from the actual world makes it the case that it could not easily have held.

Furthermore, we have reason to think that the modalist, in particular, would accept that the laws could not easily have been false.

As we discussed in section 1 the modalist typically argues that my beliefs about murder being wrong could not easily have been false. The reasoning is that the closest world where I have a false belief about murder being wrong is one where I believe murder is acceptable. But I could not easily have believed murder is acceptable — I would have to have been an extremely different person to do so.

But if I could not have easily believed that murder is acceptable — something that would have occurred if my upbringing had been very different, or if I was hypnotized, or if I was in the grip of a deeply mistaken philosophical theory — then it seems clear that it could not easily have been the

case that the basic physical laws are different. And so, given T, it could not easily have been the case that the correlation between the protons and electrons failed to hold.

3.2.2 Sensitivity*

So much for safety*. What about sensitivity*? Can the modalist respond by denying that the charge correlation is sensitive*?

Whether the charge correlation is sensitive* is a harder question than whether it is safe*. It's particularly hard because the counterfactuals we have to evaluate are *counternomics* — that is, they have nomically impossible antecedents. They ask what would happen if electrons or protons had a different charge. Often it's extremely hard for us to evaluate what would be the case if the laws were different. What would be the case, for example, if general relativity were false? Would special relativity still be true? Would the world be Newtonian? It's hard to see what considerations we could bring to bear upon this.[7]

It's hard to know, in particular, in a situation where the charge of electrons is nomically necessary, what would be the case if the charge was different. So, does that mean that the modalist can reasonably claim that the correlation fails to be sensitive* and, further, that this this the reason that we are inclined to reject the combination of the correlation and T?

I don't think this is an attractive option for the modalist. The reason is that there are background assumptions we can add which guarantee that sensitivity* holds, but these background assumptions don't make a difference to how we should react to the combination of the theory T and the charge correlation — we should still reject the combination of T and the correlation.

In particular, it's very plausible that electron and protons have their charge essentially. And, if they have their charge essentially, then they have their charge as a matter of metaphysical necessity.

Assuming this, for a moment, consider the counterfactuals relevant for the sensitivity of the charge

[7] This is not to say that every case is like this. There are some counternomics which are easier to evaluate and which seem clearly true (Tan, 2019).

correlation, for example: If the charge of the proton hadn't been $1.602176634{\times}10^{-19}$ coulombs then the charge of the electron would not have been $-1.602176634{\times}10^{-19}$ coulombs.

As we discussed in section 1 the modalist takes the counterfactual *if it had been the case that murder was not wrong then I would have not believed that murder is wrong* to be trivially true, since it could not have been the case that murder is not wrong. Similarly, if the charge of protons is necessary then the counterfactual *if the charge of the proton hadn't been* $1.602176634{\times}10^{-19}$ *coulombs then the charge of the electron would not have been* $-1.602176634{\times}10^{-19}$ *coulombs* would be trivially true.

The same is true for the counterfactual *if the charge of the electron hadn't been* $-1.602176634{\times}10^{-19}$ *coulombs then the charge of the proton would not have been* $1.602176634{\times}10^{-19}$ *coulombs*. The modalist's prior reasoning implies that it is trivially true.

So, if the charges of protons and electrons are essential, and therefore metaphysically necessary, then the charge correlation is sensitive*, given T.

But why think that protons and electrons have their charge essentially? Well, one reason is if we accept one of the variety of nomic essentialist positions which imply that the laws of nature are metaphysically necessary. There are a variety of different approaches to laws of nature that would get that result. The most common version of the view appeals to the nature of the dispositions or powers that properties have essentially (e.g. Shoemaker (1980); Swoyer (1982); Ellis (2007); Bird (2007); other versions appeal to claims about how possible worlds are generated (e.g. Wilson (2020)).

But even if we don't hold any of these positions, it's still natural to think that electrons and protons have their charge necessarily for broadly Kripkean reasons. If there was some possible world where there was something electron-like but it was, in fact, positively charged rather than negatively charged then we would take it to be a different particle, and not an electron.

Getting into the detail of these positions that imply that protons and electrons have their charge essentially would take us too far afield. But importantly, if any of these positions are true, then electrons and protons have their charge necessarily and the charge correlation is metaphysically necessary and consequently is sensitive*.

Can't the modalist simply deny these positions though? Can't they deny these claims about scientific essentialism and the Kripkean considerations, saying that the charge correlation isn't sensitive* given T and that's the reason we should reject the theory T?

They could take this approach, but the resulting view is not attractive. The problem with this approach is that it makes the acceptability of T, given the charge correlation, depend upon these background philosophical views about the nature of scientific laws, or about what properties are essential. But this isn't right. Whether we should accept the combination of the charge correlation and theory T is a first order scientific question, that should not be affected by the truth of the philosophical theories about laws and modality. And in particular we should continue to reject T, given the charge correlation, regardless of whether the charges are had essentially.

So the strategy of denying sensitivity is not, I think, an attractive way for the modalist to resist the argument.

3.3 Don't Reject T

Another possible way for the modalist to respond is by claiming that we should be happy to accept the combination of the charge correlation and T. This, they could try to argue, is not a bad result.

This, I think, is implausible. That the charge of the proton and the charge of the electron seem unrelated given T is a powerful reason to think that T is unlikely. But we can go further in making this case by noting that the structure of **Protons and Electrons** is the same as other cases that are under serious scientific investigation. And those cases suggest that we should reject T.

In particular, the 'hierarchy problem' of the mass of the Higgs Boson is a case of this kind. Going into detail about this problem would require a long discussion of particle physics and, of course, now is not the time for that. But the problem is, again, one which involves a strikingly coincidental match between two values, both of which hold with nomic necessity. Very roughly speaking, the matching is between the different factors that contribute to the mass of the Higgs Boson. The bare mass of the Higgs Boson and the radiative corrections appear to match, and therefore to cancel, in

a seemingly miraculous way leaving the total mass extremely small (see, e.g. Arkani-Hamed et al. (2016), Torrente-Lujan (2014), Craig (2020)).

Importantly, the hierarchy problem is taken to be a *problem*. The matching between these values is taken to be a problem for our current theories and a reason to look for new physics. The nomic necessity of the matching, given our current theories, isn't a reason to be happy with the combination of our theory and the matching.

The response that we shouldn't reject T goes against how similar cases are dealt with in scientific practice. Consequently, we shouldn't accept this response.

3.3.1 What this case tells us

The takeaway from cases like **Protons and Electrons** then, is that explanatory factors are most relevant to theory choice, not modal ones. And, in particular, when a striking correlation is unexplained that's a problem. Just making that correlation modally robust doesn't make the unexplained correlation unproblematic.

Though, if we are being careful, there is no need for us to say that such modal considerations are *never* relevant to theory choice. But the point is that even when a correlation could not easily have been false, and when the correlation satisfies the analogues of safety and sensitivity, it is still the case that we have reason to reject a theory that implies that the correlation has no explanation.

The modalist response to (1)-(3), then, is unsuccessful. When we consider the correlation between our moral beliefs and the moral truths the explanatory disconnection is a reason to reject the combination of the correlation and robust realism — this is not undermined by pointing out that the correlation has certain modal features, given robust realism. It is the explanationist that wins out over the modalist.

Of course, philosophical arguments like this are rarely conclusive. So I'll just point to one place where someone might try to resist my argument (of course, this is in addition to the variety of responses that

I considered in section 3). The modalist move is to say that we should be happy to accept theories that lead to big unexplained correlations, as long as the theory implies that the correlations have appropriate modal features. And the appropriate modal features I considered were the analogues of safety and sensitivity. This was the natural place to look because of the way that modalists in the literature have put a lot of weight on safety and sensitivity.

But perhaps the modalist could claim that it is, in fact, other modal features that are relevant. They could claim, that is, that there are some modal features that the correlation between our moral beliefs and the moral truths have but are not had by the correlation in **Protons and Electrons** and analogous scientific cases. And, further, that those modal features are importantly connected to theory choice — they are the reason that we reject the theory T in **Protons and Electrons** but should be happy to accept robust realism about morality. This is certainly a pathway for the modalist, but it seems to be a very difficult task to find such modal features. It's a challenge to the modalist for them to be specific about what such features are.

4 Conclusion

The modalist response to (1)-(3), I have argued, fails. We have good reason to think that (1)-(3) is a powerful argument against robust realism.

But, there is a sense in which a modalist like Clarke-Doane might look at everything I've said in this paper, but complain that I haven't really said why his argument is wrong. As we discussed in section 1 modalists in the literature so far have argued that, given robust realism, our moral beliefs are safe and sensitive, and so we are justified in our moral beliefs. How, the modalist might ask, have I shown this to be false?

To be clear, I haven't. My aim has not been to claim that robust realism leads to moral skepticism. Rather, as I noted in section 2 my aim is to defend a type of debunking argument — argument (1)-(3) — which works differently and doesn't focus on the justification of our beliefs.

In fact, it might be possible for me to simply accept that the we are justified in our moral beliefs, given robust realism. The argument (1)-(3) is designed to show that given robust realism we have reason to think that there isn't a correlation between our moral beliefs and the truth. Whether this undermines the justification we have in specific moral beliefs is a difficult and controversial question. If we accept (1)-(3) we come to believe that there isn't a correlation between our moral beliefs and the truth and, therefore, many of our moral beliefs are false. Does this undermine the justification for specific moral beliefs that we have, for example, our belief that murder is wrong?

This type of question, about the epistemic force of higher-order evidence, has been hotly contested in the recent epistemology literature. If the answer turns out to be no, then accepting the argument (1)-(3) is perfectly consistent with modalists arguing that we have justification in our moral beliefs because they are safe and sensitive.

But again, the key dialectical point is that I don't need to take a stand on such questions about justification. These questions about justification have been at the center of the recent literature on debunking. But, if we understand the issues as an instance of more general considerations about theory choice, as I have been, then issues of justification are secondary. We can simply conclude that we have reason to reject the combination of robust realism and the correlation between our belief and truth.

References

Arkani-Hamed, N., T. Cohen, R. T. D'Agnolo, A. Hook, H. D. Kim, and D. Pinner (2016, December). Solving the hierarchy problem at reheating with a large number of degrees of freedom. *Physical review letters 117*(25), 251801.

Baras, D. *Calling for Explanation*. Oxford University Press.

Baras, D. (2017, June). Our reliability is in principle explainable. *Episteme 14*(2), 197–211.

Berto, F. and M. Jago (2018). Impossible Worlds. https://plato.stanford.edu/archives/fall2018/entries/impossible-worlds/.

Bhogal, H. (2020). Coincidences and the grain of explanation. *Philosophy and phenomenological research 100*(3), 677–694.

Bird, A. (2007, January). Nature's metaphysics. *OUP*.

Bogardus, T. (2016, April). Only all naturalists should worry about only one evolutionary debunking argument. *Ethics 126*(3), 636–661.

Clarke-Doane, J. (2015). Justification and explanation in mathematics and morality. *Oxford studies in metaethics 10*, 80–103.

Clarke-Doane, J. (2016a). Debunking and dispensability. In U. D. Leibowitz and N. Sinclair (Eds.), *Explanation in Ethics and Mathematics: Debunking and Dispensability*. Oxford University Press.

Clarke-Doane, J. (2016b). What is the benacerraf problem? In *Logic, Epistemology, and the Unity of Science*, pp. 17–43. Cham: Springer International Publishing.

Clarke-Doane, J. (2020, March). *Morality and Mathematics*. Oxford University Press.

Clarke-Doane, J. and D. Baras (2019, September). Modal security. *Philosophy and phenomenological research 19*, 1.

Craig, N. (2020, November). Naturalness hits a snag with higgs. *Physics 13*.

Dworkin, R. (1996, April). Objectivity and truth: You'd better believe it. *Philosophy & public affairs 25*(2), 87–139.

Ellis, B. (2007, January). *Scientific Essentialism*. Cambridge University Press.

Enoch, D. (2010, April). The epistemological challenge to metanormative realism: how best to understand it, and how to cope with it. *Philosophical studies 148*(3), 413–438.

Enoch, D. (2011, July). *Taking Morality Seriously: A Defense of Robust Realism*. OUP Oxford.

Faraci, D. (2018). Explanation in ethics and mathematics: Debunking and dispensability by uri d. leibowitz and neil sinclair. *Analysis 78*(2).

Faraci, D. (2019). Groundwork for an explanationist account of epistemic coincidence. *Philosophers Imprint*.

Field, H. (1989, January). *Realism, Mathematics & Modality*. Realism, Mathematics & Modality. Oxford: Basil Blackwell.

Field, H. (1996). The a prioricity of logic. In *Proceedings of the Aristotelian Society*, Volume 96, pp. 359–379.

Hart, H. L. A. and T. Honoré (1985, May). *Causation in the Law*. OUP Oxford.

Horwich, P. (1982, January). *Probability and Evidence*. Cambridge University Press.

Joyce, R. (2016). Reply: Confessions of a modest debunker. *Explanation in ethics and mathematics. Oxford University Press, Oxford*, 124–148.

Korman, D. Z. (2019, December). Debunking arguments. *Philosophy Compass 14*(12), e12638.

Korman, D. Z. and D. Locke (2020a). Against minimalist responses to moral debunking arguments. *Oxford Studies in Metaethics 15*.

Korman, D. Z. and D. Locke (2020b). Evolutionary debunking and moral relativism. In M. Kusch (Ed.), *The Routledge Handbook of Philosophy of Relativism*, pp. 190–199. New York: Routledge.

Lando, T. (2017, March). Coincidence and common cause. *Noûs 51*(1), 132–151.

Lange, M. (2010, July). What are mathematical coincidences (and why does it matter)? *Mind; a quarterly review of psychology and philosophy 119*(474), 307–340.

Lewis, D. (1979). Counterfactual dependence and time's arrow. *Nous 13*(4), 455–476.

Lewis, D. (1986, January). *On the Plurality of Worlds*. London: Basil Blackwell.

Lutz, M. (2018). What makes evolution a defeater? *Erkenntnis. An International Journal of Analytic Philosophy 83*(6), 1105–1126.

Lutz, M. (2020). The reliability challenge in moral epistemology. *Oxford Studies in Metaethics 15*.

Pust, J. (2005, June). On explaining knowledge of necessity. *Dialectica 58*(1), 71–87.

Schafer, K. (2017). Review of leibowtiz and sinclair (eds.) explanation in ethics and mathematics: Debunking and dispensability. *Notre Dame Philosophical Reviews 2017*.

Schechter, J. (2010). The reliability challenge and the epistemology of logic. *Philosophical Perspectives. A Supplement to Nous 24*(1), 437–464.

Schlesinger, G. N. (1987). Miracles and probabilities. *Noûs 21*(2), 219–232.

Setiya, K. (2012, November). *Knowing Right From Wrong*. OUP Oxford.

Shoemaker, S. (1980). Causality and properties. In P. Van Inwagen (Ed.), *Time and Cause: Essays presented to Richard Taylor*, pp. 109–135. Dordrecht: Springer Netherlands.

Skarsaune, K. O. (2011). Darwin and moral realism: survival of the iffiest. *Philosophical studies 152*(2), 229–243.

Street, S. (2006). A darwinian dilemma for realist theories of value. *Philosophical studies 127*(1), 109–166.

Swoyer, C. (1982, January). The nature of natural laws. *Australasian journal of philosophy 60*, 203–223.

Tan, P. (2019, March). Counterpossible non-vacuity in scientific practice. *The journal of philosophy 116*(1), 32–60.

Tersman, F. (2016). Explaining the reliability of moral beliefs. In N. S. A. Leibowitz (Ed.), *Explanation in Ethics and Mathematics: Debunking and Dispensability*. Oxford University Press.

Torrente-Lujan, E. (2014, February). The higgs mass coincidence problem: why is the higgs mass $m_h^2 = m_z m_t mh2 = mzmt$? *The European Physical Journal C 74*(2), 2744.

Tropman, E. (2014, May). Evolutionary debunking arguments: moral realism, constructivism, and explaining moral knowledge. *Philosophical explorations: an international journal for the philosophy of mind and action 17*(2), 126–140.

White, R. (2005, January). Explanation as a guide to induction. *Philosophers' Imprint 5*, 1–29.

Wielenberg, E. J. (2010, April). On the evolutionary debunking of morality. *Ethics 120*(3), 441–464.

Wilson, A. (2020, January). *The Nature of Contingency: Quantum Physics As Modal Realism*. Oxford University Press.

Wong, W.-H. and Z. Yudell (2015, October). A normative account of the need for explanation. *Synthese 192*(9), 2863–2885.