

## THE EPISTEMOLOGY OF DEBUNKING ARGUMENTATION

BY JONATHAN EGELAND

*There is an ever-growing literature on what exactly the condition or criterion is that enables some (but not all) debunking arguments to undermine our beliefs. In this paper, I develop a novel schema for debunking argumentation, arguing that debunking arguments generally have a simple and valid form, but that whether or not they are sound depends on the particular aetiological explanation which the debunker provides in order to motivate acceptance of the individual premises. The schema has three unique features: (1) it satisfies important desiderata for what any acceptable account of debunking would have to look like; (2) it is consistent with the inductively supported claim that there is no special debunking principle; and (3) it coheres with the plausible claim that what makes debunking arguments unique is that they rely on so-called genealogies for the justification of their premises.*

**Keywords:** debunking arguments, justification, rationality, safety, sensitivity, genealogy.

### I. INTRODUCTION

In Plato's *Republic* (1997: 2.358e2-359b5), Glaucon argues against Socrates' claim that justice is valuable in and of itself by inquiring into its origins. He says that there isn't any justice in the state of nature and, hence, that it was better to commit an 'injustice' than to suffer one. However, since the badness of suffering an 'injustice' generally outstrips the goodness of inflicting an 'injustice', less powerful individuals had to find another way to get the upper hand against their more powerful adversaries. They did this, Glaucon tells Socrates, by instituting 'laws and covenants' which they called 'lawful and just'. Thus, Glaucon attempts to undermine Socrates' claim about the nature and value of 'justice' by revealing its unflattering origins in a desire by weak and impotent individuals to get revenge on more dominant individuals who are powerful enough to commit 'injustice' with impunity.<sup>1</sup>

<sup>1</sup> 'The best is to do injustice without paying the penalty; the worst is to suffer it without being able to take revenge. Justice is a mean between these two extremes. People value it not as good

Glaucon's reasoning is an example of what now is called *debunking arguments*; i.e. arguments that purport to undermine a belief's rational standing by providing a plausible account of aetiological factors that have led to the formation or maintenance of the belief that casts the belief in an unflattering light. As evidenced by the paragraph above, debunking arguments have a long history, and they have been used by a diverse array of thinkers in order to (try to) undermine a multitude of beliefs and practices. Freud (1927/1961), for example, argued that belief in the existence of God has its origins in wishful thinking, and that this undermines the rationality of such belief (Dennett 2006; Griffiths and Wilkins 2013; cf. Hume 1757/1992; Nietzsche 1888/1982). Similarly, since the middle of the twentieth century, many ethicists have argued that the evolutionary origins of our moral intuitions and beliefs pressure us to revise those beliefs, to give them up, or to be anti-realists about their contents (Braddock 2017; Cofnas 2020; Gibbard 1990; Griffiths and Wilkins 2015; Harman 1977; Huemer 2008; Joyce 2006; Kitcher 2005; Rosenberg 2011; Ruse and Wilson 1986; Singer 2005; Street 2006). Other beliefs or practices whose rational standing also have been called into question by different debunking arguments are concerned with logic (Cooper 2003), probability (Handfield 2016), math (Benacerraf 1973; Field 1989), metaphysics (Ladyman and Ross 2007), epistemology (Williams 2004), science (Laudan 1981; Poincaré 1905/1952), naturalism (Plantinga 1993), social structures (Foucault 1975/1995), and aesthetics (Carroll 2001).<sup>2</sup>

However, despite their widespread use in the philosophical discourse (and outside it too), there isn't really much clarity as to how debunking arguments actually work, or indeed whether they work at all.<sup>3</sup> Consider, for example, a scenario where you grow up in a society of devout Mormons, and where you suddenly sometime during your teenage years realize that there are a great many religions in the world besides Mormonism and that you would have been just as confident in the truth of another set of religious dogmas had you grown up elsewhere.<sup>4</sup> How would you react in such a situation? How *should* you react?

In the scenario, there is without a doubt an aspect of arbitrariness to your belief in the tenets of Mormonism, insofar as they are caused by contingent features of your social environment that aren't directly relevant to the question of whether Mormonism is true. But does this undermine the rational standing

but because they are too weak to do injustice with impunity. Someone who has the power to do this, however, and is a true man wouldn't make an agreement with anyone not to do injustice in order not to suffer it. For him that would be madness' (359a5-359b4).

<sup>2</sup> This is not supposed to be an exhaustive list; in fact, it is far from it.

<sup>3</sup> Srinivasan (2015) has recently argued that debunking arguments invariably are self-defeating.

<sup>4</sup> For a similar, but more detailed example, see Knobe and Nichols (2008: 11). Cf. Vavova's (2018) interesting and instructive discussion of the case of G. A. Cohen, who reflectively claimed that he endorsed the analytic-synthetic distinction in part because he decided to go to Oxford rather than Harvard.

of your belief in the truth of Mormonism? Perhaps it does. After all, the fact that you happened to be raised in one among many different actual societies of religious believers—not to mention the even greater number of possible religious societies—and that this fact plays an important role in the formation and maintenance of your belief does seem to cast it into doubt. However, on the other hand, although you would have believed that the earth is flat had you been raised by a tribe of flat earthers, that does not undermine your actual belief that the earth is round. Indeed, the fact that a belief is the product of contingent causes that don't directly bear on the truth-value of its content is clearly not sufficient for undermining the rational standing of that belief.<sup>5,6</sup> Still, learning that you hold a certain belief for contingent causes which appear to contribute little, if any, support to it, and that you easily could have had some other, incompatible belief, should make you pause and question its veracity.

There is an ever-growing literature on what exactly the condition or criterion is that enables some (but not all) debunking arguments to undermine our beliefs. The goal of this article is to evaluate some prominent proposals in the literature and to provide a critical study of the epistemological functions of debunking arguments. However, doing so, I will not identify a single condition or criterion that is supposed to be involved in all valid debunking arguments and that is sufficient for undermining the rational standing of the target beliefs. Many epistemologists have in recent years set out to discover what this criterion might be, but with no more luck than explorers searching for Atlantis. It is my belief that both kinds of project have failed for the same reason: the thing searched for does not exist. So instead of trying to find a single feature that is common to all valid debunking arguments and that is responsible for undermining the beliefs they target, I will instead argue that debunking arguments generally have a simple and valid form (without relying on some special debunking principle), but that whether or not they are sound depends on the particular aetiological stories which the debunker provides in order to motivate acceptance of the individual premises.

The article is structured as follows. In Section II, I take a closer look at some prominent suggestions in the literature as to how debunking arguments should be formulated, and I explain why they fall short of satisfying a couple of desiderata for what any acceptable account of debunking would have to look like. In Section III, I provide a novel account of the epistemological functions of debunking arguments and illustrate the various ways in which they can be used in order to undermine our beliefs. In Section IV, I motivate the account

<sup>5</sup> The fact that the majority of factors that are causally responsible for our beliefs don't have any direct bearing on their truth-value does not by itself undermine their rationality.

<sup>6</sup> Cf. White (2010: 574) and Vavova (2018), although Vavova argues that evidence of irrelevant influence is *sometimes* evidence of error (135).

from the previous section and argue that debunking arguments generally differ from other kinds of argument, but that the reason why does not have anything to do with their form or with special debunking principles. In Section V, I summarize and conclude.

## II. DEBUNKING AND THE EPISTEMIC PRINCIPLES ON WHICH IT ALLEGEDLY RELIES

A sound debunking argument does not demonstrate that the belief it targets is false; rather, it undermines the rational standing of the belief in some other way. This naturally raises a fundamental question: If not by demonstrating the falsehood of the target belief, how exactly do debunking arguments undermine its rationality? This is a very timely question, one that currently is discussed by many philosophers.<sup>7</sup> The challenge is to offer a general schema for how debunking arguments should be formulated, such that it satisfies the following desiderata:

*The Validity Desideratum:* insofar as the premises are true, debunking arguments that conform to the structure of the schema undermine the rational standing of the beliefs they target, while leaving the rational standing of other, unrelated beliefs unaffected.

*The Informativeness Desideratum:* the schema clarifies exactly *how* the debunking arguments that satisfy the validity desideratum are able to undermine the rational standing of the target beliefs.

Debunking arguments come in different varieties, but they generally conform to the following schema for a token belief B and a belief forming method M:

### *Debunking Schema*

1. B is formed or maintained on the basis of M.
2. M is epistemically defective.
3. Hence, B is irrational. (*Cf.* Nichols 2015: 101–2; Sauer 2018: 30)

The schema has two premises, the first of which is aetiological, in the sense that it specifies some causal factor which is responsible for a certain belief, and the second of which is epistemic, in the sense that the aforementioned causal factor is shown to be defective in a way that undermines the justificatory status of the belief in question.

As it stands, the schema is valid: any belief that can be shown to be formed or maintained on the basis of an epistemically defective method will be irrational. However, validity is not the only desideratum for a theoretically satisfying

<sup>7</sup> See e.g. the relevant discussions by White (2010), Bogardus (2016), and Sauer (2018: ch. 1).

schema for debunking argumentation. In addition, we also want it to be clear and informative. Unfortunately, however, the schema above is not very informative, since it is not at all clear exactly *how* debunking arguments are able to undermine the rational standing of the beliefs they target. And the reason is that the second premise is too vague.<sup>8</sup> We all agree that it is irrational to hold beliefs on the basis of some epistemically defective belief forming method. But we want to know the various ways in which a belief forming method can be epistemically defective, such that it undermines the rational standing of our beliefs. In other words, we want to know how the second premise should be spelled out, such that it becomes clear how successful debunking arguments are able to undermine the beliefs they target.<sup>9</sup>

So, as we have seen, although the schema proposed by (e.g.) Nichols satisfies the validity desideratum, it fails to satisfy the informativeness desideratum. However, when, in the literature, attempts have been made to provide a more informative account of how debunking arguments actually function, they seem to fall short of the validity desideratum. Indeed, whenever an account satisfies one of the desiderata, it appears unable to satisfy the other. Let's take a closer look at a few examples.

Richard Joyce, one of the most prominent proponents of debunking argumentation, especially when it comes to beliefs about morality, offers the following characterization of how debunking works:

I contend that on no epistemological theory worth its salt should the justificatory status of a belief remain unaffected by the discovery of an empirically supported theory that provides a complete explanation of why we have that belief while nowhere presupposing its truth. (Joyce 2006: 219)

And he tries to undermine the justificatory status of our beliefs about morality by applying this kind of argumentation to them.<sup>10</sup> He writes:

<sup>8</sup> Nichols, for example, doesn't say much about what he takes a belief-forming process to be or about the ways in which such a process may be epistemically defective. About the latter issue, he approvingly cites Alvin Goldman, who offers the following list of examples of epistemically defective belief-forming processes: 'confused reasoning, wishful thinking, reliance on emotional attachment, mere hunch or guesswork, and hasty generalization' (Goldman 1979: 9). However, he doesn't want to commit himself to any particular view about *why* these belief-forming processes are defective. It might be that Goldman is right and that they all are defective because they are unreliable, or it might be that Stewart Cohen (1984) is right and that they all are defective because the resulting beliefs aren't supported by the evidence (Nichols 2015: 103).

<sup>9</sup> I think a similar problem might be facing those whose conception of debunking is focused on the failure of the target belief to *track the truth* (see e.g. Cofnas 2020; Greene 2008; FitzPatrick 2015; Kahane 2011; Rini 2016). Without an analysis of the notion, one simply cannot be sure what is meant by 'failing to track the truth'.

<sup>10</sup> Joyce's argument builds on Gilbert Harman's (1977) explanatory challenge to the belief in moral facts. Harman's challenge focuses on the fact that there is a disanalogy between scientific and moral beliefs: whereas the best explanation of the scientist's beliefs typically refers to the empirical objects that the beliefs are about (such as electrons or benzene rings), the best

Our moral beliefs are products of a process that is entirely independent of their truth, which forces the recognition that we have no grounds one way or the other for maintaining these beliefs. (Joyce 2006: 211)

Thus, according to Joyce, all we need in order to undermine the justificatory status of someone's beliefs about morality (or, indeed, any other subject matter) is to provide a 'complete explanation' of why the person has those beliefs, where the explanation doesn't imply that they are true. However, Nichols (2015: 98–101) has recently pointed out a couple of problems with Joyce's conception of debunking. First, we can explain our beliefs about the future in a way which doesn't presuppose their truth, but without becoming any less justified in holding those beliefs. For example, let's say that I believe that I will go to work at 8 a.m. tomorrow. Even though the fact that I have this belief can be completely explained in terms of my (access to my) intentions and expectations (or perhaps in terms of my current neural state, if you prefer a more low-level analysis), it doesn't follow that my belief is unjustified. In fact, it may even be the case that my belief is false—perhaps there is a strike among the employees at the train company, thus preventing me from getting to work on time—but this, together with the explanation above, intuitively doesn't do anything to the justificatory status of my belief; I still seem to be just as justified in believing that I will get to work on time tomorrow.

Secondly, not only does Joyce's conception of debunking appear to be problematic in the case of future-oriented beliefs, but it also appears to be in trouble when it comes to any justified belief that is false. The reason is simply that when you have a justified false belief, the best explanation for why you have that belief *cannot* presuppose its truth—in which case Joyce has to say that the belief (or, more precisely, its justificatory status) is debunked. But this clearly seems wrong. To illustrate the problem, Nichols (2015: 99) has us consider Lavoisier's belief in the caloric theory of heat. Since the caloric theory is false, the best explanation for Lavoisier's belief in that theory cannot presuppose its truth—which means that, according to Joyce, it should be debunked. However, just because the best explanation for Lavoisier's belief doesn't mention its truth, it doesn't follow that it is unjustified. As long as he bases his belief on good supporting evidence, his belief will be justified despite being false and despite being explained in a manner which is open to the possibility that it is false. The general worry which Nichols seems to have identified with Joyce's conception of debunking is that a subject can hold a justified belief for reasons which don't imply its truth. Indeed, it is possible to have justifying, *non-factive* reasons for belief.<sup>11</sup>

explanation of our moral beliefs does not refer to moral facts in the robust realist sense of the term, but rather to facts about our human psychology.

<sup>11</sup> However, no philosophical view, no matter how plausible, is immune to dissent. At least some of the proponents of what has been called 'the factive turn in epistemology' argue that

Another idea as to how debunking argumentation functions in order to undermine the justificatory status of some target belief concerns the issue of luck. Some commentators, such as Matthew Bedke (2009) and Sharon Street (2008), argue that a belief is vulnerable to debunking if its truth-value depends on the believer being lucky. Consider the following passage by Street, who offers a debunking argument against moral realism:

The realist must hold that an astonishing coincidence took place—claiming that as a matter of sheer luck, evolutionary pressures affected our evaluative attitudes in such a way that they just happened to land on or near the true normative views among all the conceptually possible ones. (Street 2008: 208–9)

However, a problem with the claim that the central ingredient in any successful debunking argument is luck is that it appears possible for justified beliefs to be true as a function of luck. Consider, for example, a world in which every person at infancy is given spectrum inverting colour lenses that they have to wear for the rest their lives. As a result, when people see green things, such as grass or the flag of Saudi Arabia, they experience them as being another colour. However, due to a manufacturing mistake, the lenses of a particular person are double-inverting in a way which cancels out the inverting effect. Now although the person's true colour beliefs will be epistemically lucky, that does not prevent them from being justified. Indeed, not only would they be justified, they would plausibly constitute knowledge—or at least so I contend.<sup>12</sup> This shows that luck is compatible with justification, and that the view which claims that luck opens up for undermining via debunking cannot be true.

Another suggestion, related to one above, is that debunking arguments undermine the beliefs they target by showing them to be insensitive, in the sense that one would still hold the beliefs even if they were false (at least if one were to rely on the same belief-forming process). This view is endorsed by Ruse and Wilson (1986), Justin Clarke-Doane (2012), and perhaps the commentators who focus on the property of 'truth-tracking' too (*cf. fn. 9*).<sup>13</sup> These commentators sometimes claim that the insensitivity of our moral beliefs shows them to lack any rational standing and, moreover, that morality in some sense is 'redundant' since our moral beliefs can be explained without referencing their content as true.<sup>14</sup> However, a problem with this suggestion is that sensitivity is not a

all (epistemic) reasons for belief are factive. Mitova (2018: 1), for example, says that 'when you believe something for a *good reason*, your belief is in a position to enjoy all the cardinal epistemic blessings: it can be rational, justified, warranted, responsible, constitute knowledge, you name it.'

<sup>12</sup> For a similar argument that draws upon Hawthorne's (2002) swampwatch scenario, see Bogardus (2016: 655).

<sup>13</sup> Commentators like Kahane (2011) and FitzPatrick (2015) clearly appear to have sensitivity in mind when they talk about 'off-track' belief-forming processes.

<sup>14</sup> See e.g. Ruse and Wilson (1986: 183ff) who explicitly reject objective morality as redundant and illustrate their point by explaining how belief in the immorality of sibling incest has been shaped by certain epigenetic rules.

necessary condition for either knowledge or justification. To see why that is so, consider the following objection from Roger White:

[T]he problem arises more generally from cases of empirical underdetermination. As jury members we have overwhelming evidence that the defendant is guilty (finger prints, eyewitness reports, DNA matching blood on the murder weapon. . .). Although we are more than justified in believing that he is guilty, our evidence doesn't strictly entail it. There is always logical room for some wildly ad hoc conspiracy theory according to which the defendant was framed and all the evidence was planted. We will have to be justified in denying this conspiracy theory if we're justified in believing the defendant to be guilty (since they can't both be true). But if the conspiracy theory were true we would have all the same evidence and hence quite reasonably (but mistakenly) believe the defendant to be guilty and hence the conspiracy theory to be false. So again, by Truth Sensitivity we derive the absurd conclusion that we aren't justified in denying the conspiracy theory or believing that the defendant is guilty. (White 2010: 581)

If a belief can be justified or constitute knowledge without having the property of sensitivity, as White plausibly argues, then it cannot be debunked by an argument demonstrating that it does not have said property.

The last suggestion that I will discuss as to how debunking arguments should be formulated focuses on the property of safety. On this view, (successful) debunking arguments undermine the justificatory status of the beliefs they target by showing them to be unsafe, in the sense that they easily could be false (at least if one were to rely on the same belief-forming process).<sup>15</sup> The view is held, or at least hinted at by (e.g.) Charles Darwin (1879/2004) and Alex Rosenberg (2011), and it is motivated by the idea that relatively small changes in our evolutionary history—a history which has been determined to a large extent by chancy factors outside of our control, such as random (beneficial/neutral/deleterious) genetic mutations—would have caused the belief-forming process we have used in order to arrive at our current moral beliefs to produce other moral beliefs that clearly appear to be false, which naturally calls into question the veracity of our current beliefs about morality.

As an example of this sort of reasoning, Darwin draws our attention to what our behaviour, psychology, and moral beliefs would be like if 'men were reared under precisely the same conditions as hive-bees':

[T]here can hardly be any doubt that our unmarried females would, like worker-bees, think it a sacred duty to kill their brothers, and mothers would strive to kill their fertile daughters; and no one would think of interfering. (Darwin 1879/2004: 122)

Similarly, Rosenberg (2011) tells us that 'if the environment had been very different, another moral core would have been selected for ... But

<sup>15</sup> There are many ways of analysing the safety notion. For disagreeing perspectives, see Williamson (2000), Pritchard (2005, 2009), and Sosa (1999). For present purposes, I have chosen a bare-bones formulation that is neutral between the different analyses.

it wouldn't have been made right, correct, or true by its fitness in that environment'.

However, this suggestion is also problematic, since there are good reasons to think that safety is not necessary for either knowledge or justification. Consider, for example, a young girl who asks her mother where babies come from. The mother tells her the truth about 'the birds and the bees', and the daughter believes her. Now let's also imagine that the daughter easily could have ended up with false beliefs about human reproduction, since the mother had planned to tell her that storks deliver babies to married adults who love each other very much, and that she (i.e. the mother) actually was surprised to hear herself telling the young girl the truth about human reproduction despite her young age. Now even though the young girl easily could have ended up with false beliefs about where babies come from, her newly formed beliefs intuitively seem to be justified and prime candidates for qualifying as knowledge, since the threat of being told a white lie by the mother remains counterfactual. And since a belief easily could have been false without having to be unjustified or failing to constitute knowledge, debunking arguments cannot rely on the notion of safety in order to undermine the rational standing of our beliefs.

In this section, I have taken a brief look at some prominent suggestions about how debunking arguments should be formulated.<sup>16</sup> However, as we have seen, none of the suggestions satisfy both the validity desideratum and the informativeness desideratum. In the next section, I will therefore present a novel schema for debunking argumentation that is both valid and informative.

### III. DOXASTIC DEBUNKING AND REASONS FOR BELIEF

In order to flesh out the various ways in which debunking arguments actually can debunk our beliefs, it will be much more helpful to consider the reasons on which our beliefs are based, rather than the processes that caused or sustain them, or special debunking principles. As an alternative to the proposals mentioned in the previous section, I offer the following schema for debunking argumentation:

#### *Doxastic Debunking*

1. S believes that  $p$  on the basis of motivating reason R.<sup>17</sup>
2. R is not a normative reason to believe that  $p$ , or R is defeated, or  $p$  is not properly based on R.

<sup>16</sup> For other, more thorough critiques of the debunking criteria found in the literature, see White (2010) and Bogardus (2016).

<sup>17</sup> In order to avoid overdetermination that threatens the validity of the schema, R should be understood as the set of all of S's motivating reasons for believing that  $p$ .

3. Therefore, S's belief that  $p$  is unjustified.

(where S is any subject, and  $p$  is any proposition)

Before I explain how this generalized schema helps to specify the various ways in which debunking arguments can undermine the justificatory status of the target beliefs, let me first offer some clarifications about reasons for belief. A *motivating* reason is a reason for which someone believes or acts. This typically involves something which causes the subject to believe or act in a certain way. For example, my motivating reason for believing that Hamlet dies in the fifth act of Shakespeare's *Hamlet* is some factor, presumably some memory belief or memory-seeming, which is causally responsible for it. A *normative* reason, on the other hand, is a reason that counts in favour of belief or action. This typically involves something which makes it rational or justified for the subject to believe or act in a certain way. For example, my normative reason for believing that there are pieces of fruit on my desk is the fact that I currently have a perceptual experience of pieces of fruit on my desk. Moreover, motivating and normative reasons can come apart. First, not all motivating reasons are normative reasons: it is possible to believe or act for reasons which aren't any good (i.e. rationalizing or justifying). Secondly, not all normative reasons are motivating reasons: it is possible to have good reasons for belief or action without believing or acting on the basis of them.<sup>18</sup>

Now let's take a closer look at the schema above and how it helps to specify the ways in which debunking arguments can undermine the justificatory status of the target beliefs. The first premise is a *causal* premise: it says what the subject's motivating reason for their belief is. The second premise is an *epistemic* premise: it says that the subject's belief fails to satisfy some necessary condition for justification. Moreover, it tells us that there are several such conditions which the subject's belief can fail to satisfy. More specifically, there are three different ways in which the subject who bases their belief on a motivating reason R can fail to be justified in holding that belief.<sup>19</sup>

First, R might not be a normative reason at all. This can be the case if, for example, what causes the subject to hold the belief has to do with wishful thinking, bias, or brain damage. In such a case, the subject's belief would be unjustified because the reason on which it is based isn't normative (i.e. rationalizing or justifying).

An example of this kind of debunking is provided by Freud's argument, briefly mentioned in Section I, against theism. This is what one of Freud

<sup>18</sup> For more about the distinction between motivating and normative reasons, see McNaughton and Rawling (2018).

<sup>19</sup> As we will see, only one of them has to do defeat. Thus, by claiming that all of them have to do with defeat, Sauer (2018: 29) and Kahane (2011: 106) fail to recognize the other two ways in which debunking arguments can demonstrate the unreasonableness of a subject's belief.

formulations looks like:

[O]ur attitude to the problem of religion will undergo a marked displacement. We shall tell ourselves that it would be very nice if there were a God who created the world and was a benevolent Providence, and if there were a moral order in the universe and an after-life; but it is a very striking fact that all this is exactly as we are bound to wish it to be. (Freud 1927/1961: 42)

Freud claims that belief in God has its aetiological origins in wishful thinking. And although the claim does not demonstrate that God does not exist, it nevertheless undermines the rational standing of theistic belief since it shows that the reason motivating such belief isn't normative (insofar as the claim is true).

Secondly, R might be a normative reason, albeit one that is defeated by some other reason.<sup>20</sup> This can be the case if, for example, the subject has some other normative reason which counts against their belief—either by undercutting the support it receives from the reason R on which the belief actually is based or by supporting some proposition which is inconsistent with the content of the belief.<sup>21</sup> In such a case, the subject's belief would be unjustified because the reason on which it is based is undercut or rebutted by some other normative reason which the subject has.<sup>22</sup>

Arguably, this is how most debunking arguments in the literature actually work. Take for example the argument against moral beliefs offered by Joyce. Joyce claims that our moral beliefs are produced by processes that do not presuppose that the contents of those beliefs are true, and that this defeats any reason one might have for holding them (2006: 211). If this claim is true, then it shows that moral belief is defeated by higher-order evidence exposing the unreliability of first-order evidence for such belief (e.g. moral intuitions).

Thirdly, R might be a normative reason, but the subject doesn't properly base their belief on that reason. This can be the case if, for example, the subject fails to satisfy whatever conditions are required in order to properly base one's belief on a certain reason. In such a case, the subject's belief would be unjustified because it isn't properly based on its motivating reason.

Whether this kind of case actually is possible will depend upon one's view of the basing relation. If one has a view according to which a motivating reason necessarily satisfies the conditions for proper basing, then this kind of case will be impossible. That said, there are conceptions of the basing relation in the

<sup>20</sup> The first two kinds of debunking correspond to White's (2010: 575) distinction between undermining and blocking debunkers.

<sup>21</sup> These two ways in which the subject's belief can be defeated correspond to Pollock's (1986: 38–9) distinction between undercutting defeat and rebutting defeat.

<sup>22</sup> In the case of undercutting defeat, the (other) normative reason in question will plausibly be a piece of higher-order evidence—i.e. evidence about one's evidence—which indicates that the lower-order evidence doesn't provide sufficient support to the subject's belief. The term 'higher-order evidence' is borrowed from Kelly (2005).

literature, such as the doxastic conception offered by Joseph Tolliver (1982), that open up for this kind of debunking. According to Tolliver, a person's belief is properly based on a reason if and only if the person believes that the reason provides evidence for the belief, and also that the probability that the belief is true likely increased when the reason was accepted. If this view were to be correct, then any subject who has a belief that is motivated by a normative reason, but who either does not believe that the reason provides evidence for the belief or does not believe that it is likely that the probability that the belief is true increased when the reason was accepted, would be vulnerable to a debunking argument.

What all these ways of debunking the target belief have in common is that they show that the belief in question fails to satisfy some necessary condition for justification, which typically will justify an objection along the familiar lines of 'you just believe that because...';<sup>23</sup> or, in the case where the belief isn't properly based on the motivating reason, 'you don't really believe that because...'<sup>24,25</sup>

<sup>23</sup> The phrase is borrowed from the title White's (2010) paper.

<sup>24</sup> Moreover, it should be noted that the kind of justification which the debunking arguments undermine is *doxastic* justification (cf. Thurrow 2013, 2014). Doxastic justification has to do with the justifiably held beliefs (or other doxastic attitudes) someone has. *Propositional* justification, by contrast, has to do with the beliefs (or other doxastic attitudes) you have justification to hold, regardless of whether or not you actually hold them (cf. Firth (1978), who was the first to introduce the distinction). Doxastic justification is thus normally analysed as propositional justification plus proper basing. An analogous argument about the debunking of propositional justification could of course be formulated—perhaps as follows:

*Propositional Debunking*

1. S believes that  $p$  on the basis of motivating reason R.
2. S has no undefeated normative reason, including R, to believe that  $p$ .
3. S does not have justification to believe that  $p$ .

(Strictly speaking, the first premise is redundant. However, it might be helpful in order to target a particular belief which the subject has in order to see whether the subject has justification for it.)

However, such an argument would be more ambitious insofar as it would require some insight not just into the subject's motivating reason for belief, but also into *all* of the subject's normative reasons in order to determine what the subject's evidence or reasons on balance support believing.

<sup>25</sup> An anonymous referee asked whether the schema developed in this paper might be present just under the surface in the debunking literature. The idea, more specifically, is that the papers that focus on undercutting defeaters or higher-order defeat present similar-looking proposals, so long as we restrict the second premise somewhat. By way of response, I would point out that the originality of the schema lies precisely in the formulation of the second premise. It is indeed that premise which opens up for the three aforementioned ways in which the schema can be used in order to undermine some target belief—two of which (i.e. the first and third) the papers that focus on defeaters cannot account for. Moreover, even though debunking in virtue of showing that R is not a normative reason is quite common—see, e.g. Peter Singer (2005) and Joshua Greene's (2008) argument against belief in deontological values or principles—it is rarely acknowledged that *this* actually is how these debunking arguments function. My schema also rectifies this issue by making this explicit.

## IV. MOTIVATING THE ACCOUNT

Given that this general schema for doxastic debunking respects our intuitions and plausibly unpacks the various ways in which debunking arguments can undermine the target beliefs' justificatory status, I will briefly mention three reasons as to why I believe that it should be considered the leading candidate of how debunking arguments actually function. First, in contrast to the other proposals in the literature (*cf.* Section II), it satisfies both the validity desideratum and the informativeness desideratum. Not only will any debunking argument formulated in accordance with the schema successfully undermine the justificatory status of the target belief whenever the premises are true, but the schema also tells us exactly how the argument debunks said belief. More specifically, the schema satisfies the validity desideratum, since the premises specify conditions that virtually everyone agrees are necessary for doxastic justification. So whenever both premises are true, meaning that the target belief fails to satisfy some necessary condition for its doxastic justification, it invariably becomes vulnerable to a debunking argument. And, moreover, the schema also satisfies the informativeness desideratum, since it tells us exactly what the conditions are that debunked beliefs fail to satisfy.

Secondly, the schema is consistent with the claim that there is no special condition or criterion that enables some (but not all) debunking arguments to undermine the target beliefs—a claim which receives inductive support from the fact that there is no consensus on what such a condition or criterion looks like, despite that debunking arguments have been discussed and forwarded at least since the ancient Greeks (*cf.* Section I). Given that we accept the plausible inductive inference that there most likely does not exist any special debunking principle that unifies all valid debunking arguments, the vast majority of proposals in the literature as to how such arguments function must be wrong. However, it does not imply that the schema for doxastic debunking developed in the previous section must be wrong, as it is consistent with the claim that there is no special debunking principle that unifies all valid debunking arguments.

Thirdly, the schema reinforces the importance of the accounts offered, whether genealogical, evolutionary, or sociological, in order to motivate the premises.<sup>26</sup> It is by providing an account which somehow explains why the subject has the target belief, and which does so in a way that highlights how the target belief fails to satisfy some necessary condition for doxastic justification, that the debunker justifies the second premise of their argument. If there is

<sup>26</sup> It is common in the literature to assume that all such accounts must be of a genealogical nature. However, as I plan to argue against this assumption in a future paper where I will examine Nietzsche's (1887/2000) use of the term 'genealogy', I will instead only assume that some such accounts are genealogies.

no special debunking principle that unifies all valid debunking arguments, then this gives rise to an explanatory challenge: What, if anything, unifies all valid debunking arguments? Here I suggest that what unifies valid debunking arguments and separates them from other kinds of argument does not have anything to do with their form or with special debunking principles, but rather with the way in which they are motivated: in general, debunking arguments rely on aetiological explanations demonstrating that a target belief fails to satisfy some necessary condition for doxastic justification, but without in any way showing or purporting to show that the content of the target belief is false (on pain of committing the genetic fallacy). What is special about debunking argumentation is thus the manner in which its premises call for justification, and not some special, hitherto undiscovered principle that figures in any and all debunking arguments that have the formal structure required in order to undermine the rational standing of the beliefs they target.<sup>27</sup>

## V. CONCLUSION

In this article, I have argued for a novel account of debunking argumentation. Section II examined some alternative accounts from the literature and argued that they fail to satisfy both the validity desideratum and the informativeness desideratum. Section III presented the schema for doxastic debunking, and it highlighted three different ways in which it can be used in order to undermine the rational standing of some target belief. Section IV motivated the schema from Section III by arguing (1) that it satisfies the aforementioned desiderata; (2) that it is consistent with the inductively supported claim that there is no special debunking principle; and (3) that it coheres with the plausible claim that what makes debunking arguments unique is that they rely on aetiological explanations (so-called genealogies) for the justification of their premises.

## REFERENCES

- Bedke, M. (2009). 'Intuitive Non-Naturalism Meets Cosmic Coincidence', *Pacific Philosophical Quarterly*, 90: 188–209.
- Benacerraf, P. (1973). 'Mathematical Truth', *Journal of Philosophy*, 70: 661–79.
- Bogardus, T. (2016). 'Only All Naturalists Should Worry About Only One Evolutionary Debunking Argument', *Ethics*, 126: 636–61.
- Braddock, M. (2017). 'Debunking Arguments from Insensitivity', *International Journal for the Study of Skepticism*, 7: 91–113.

<sup>27</sup> A fourth advantage mentioned by an anonymous referee, and that I too believe should be noted, is that the schema (in contrast to many others in the literature) is neutral with respect to different theories of justification and knowledge. Thanks to the referee for pointing this out to me.

- Carroll, N. (2001). *Beyond Aesthetics: Philosophical Essays*. Cambridge: CUP.
- Clarke-Doane, J. (2012). 'Morality and Mathematics: The Evolutionary Challenge', *Ethics*, 122: 313–40.
- Cofnas, N. (2020). 'A Debunking Explanation for Moral Progress', *Philosophical Studies*, 177: 3171–91.
- Cohen, S. (1984). 'Justification and Truth', *Philosophical Studies*, 46: 279–95.
- Cooper, W. S. (2003). *The Evolution of Reason: Logic as a Branch of Biology*. Cambridge: CUP.
- Darwin, C. (1879/2004). *The Descent of Man, and Selection in Relation to Sex*. London: Penguin.
- Dennett, D. C. (2006). *Breaking the Spell: Religion as a Natural Phenomenon*. New York: Penguin.
- Field, H. (1989). *Realism, Mathematics and Modality*. Oxford: Basil Blackwell.
- Firth, R. (1978). 'Are Epistemic Concepts Reducible to Ethical Concepts?', in A. Goldman and J. Kim (eds), *Values and Moral*, 215–29. Dordrecht: Kluwer.
- FitzPatrick, W. J. (2015). 'Debunking Evolutionary Debunking of Ethical Realism', *Philosophical Studies*, 172, 883–904.
- Foucault, M. (1975/1995). *Discipline & Punish: The Birth of the Prison*, trans. A. Sheridan. New York: Vintage.
- Freud, S. (1927/1961). *The Future of an Illusion*. New York: Norton.
- Gibbard, A. (1990). *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press.
- Goldman, A. I. (1979). 'What is Justified Belief', in G. Pappas (ed.), *Justification and Knowledge*, 1–25. Boston: D. Reidel.
- Greene, J. (2008). 'The Secret Joke of Kant's Soul', in W. Sinnott-Armstrong (ed.), *Moral Psychology, vol. 3. The Neuroscience of Morality: Emotion, Brain Disorders, and Development*, 35–80. Cambridge, MA: MIT Press.
- Griffiths, P. and Wilkins, J. (2013). 'Evolutionary Debunking Arguments in Three Domains: Fact, Value, and Religion', in J. Maclaurin and G. Dawes (eds), *A New Science of Religion*. University of Chicago Press.
- (2015). 'When Do Evolutionary Explanations of Belief Debunk Belief?', in P. R. Sloan, G. McKenny and K. Eggleston (eds), *Darwin in the Twenty-First Century: Nature, Humanity, and God*, 201–31. University of Notre Dame Press.
- Handfield, T. (2016). 'Genealogical Explanation of Chance and Morals', in U. D. Leibowitz and N. Sinclair (eds), *Explanation in Ethics and Mathematics: Debunking and Dispensability*, 58–81. Oxford: OUP.
- Harman, G. (1977). *The Nature of Morality: An Introduction to Ethics*. New York: OUP.
- Hawthorne, J. (2002). 'Deeply Contingent A Priori Knowledge', *Philosophy & Phenomenological Research*, 65: 247–69.
- Huemer, M. (2008). 'Revisionary Intuitionism', *Social Philosophy & Policy*, 25: 368–92.
- Hume, D. (1757/1992). *The Natural History of Religion*, in J. Fieser (ed.). New York: Macmillan.
- Joyce, R. (2006). *The Evolution of Morality*. Cambridge, MA: MIT Press.
- Kahane, G. (2011). 'Evolutionary Debunking Arguments', *Nous*, 4: 103–25.
- Kelly, T. (2005). 'The Epistemic Significance of Disagreement', in J. Hawthorne and T. Gendler (eds), *Oxford Studies in Epistemology, Volume 1*, 167–96. Oxford: OUP.
- Kitcher, P. (2005). 'Biology and Ethics', in D. Copp (ed.), *Oxford Handbook of Ethical Theory*. Oxford: OUP.
- Knobe, J. and Nichols, S. (2008). *Experimental Philosophy*. Oxford: OUP.
- Ladyman, J. and Ross, D. with Spurrett, D. and Collier, J. (2007). *Everything Must Go: Metaphysics Naturalized*. Oxford: OUP.
- Laudan, L. (1981). 'A Confutation of Convergent Realism', *Philosophy of Science*, 48: 19–49.
- McNaughton, D. and Rawling, P. (2018). 'Motivating Reasons and Normative Reasons', in D. Star (ed.), *Oxford Handbook of Reasons and Normativity*. Oxford: OUP.
- Mitova, V. (2018). 'Introduction', in V. Mitova (ed.), *The Factive Turn in Epistemology*, 1–12. Cambridge University Press.
- Nichols, S. (2015). *Bound: Essays on Free Will and Responsibility*. Oxford: OUP.
- Nietzsche, F. W. (1888/1982). *The Antichrist*, trans. W. Kaufmann, in W. Kaufmann (ed.), *The Portable Nietzsche*. New York: Penguin Books.
- (1887/2000). *On the Genealogy of Morals*, trans. W. Kaufmann, in W. Kaufmann (ed.), *Basic Writings of Nietzsche*. New York: Modern Library.
- Plantinga, A. (1993). *Warrant and Proper Function*. New York: OUP.

- Plato. *Republic*, trans. G. M. A. Grube & rev. C. D. C. Reeve, in J. M. Cooper and D. S. Hutchinson (eds), *Plato: Complete Works* (1997). Indianapolis: Hackett Publishing Company.
- Poincaré, H. (1905/1952). *Science and Hypothesis*. New York: Dover.
- Pollock, J. (1986). *Contemporary Theories of Knowledge*. Savage, MD: Rowman & Littlefield.
- Pritchard, D. (2005). *Epistemic Luck*. Oxford: OUP.
- (2009). 'Safety-Based Epistemology: Whither Now?', *Journal of Philosophical Research*, 34: 33–45.
- Rini, R. A. (2016). 'Debunking Debunking: A Regress Challenge for Psychological Threats to Moral Judgment', *Philosophical Studies*, 173: 675–97.
- Rosenberg, A. (2011). *The Atheist's Guide to Reality: Enjoying Life without Illusions*. New York: W. W. Norton.
- Ruse, M. and Wilson, E. O. (1986). 'Moral Philosophy as Applied Science', *Philosophy*, 61: 173–92.
- Sauer, H. (2018). *Debunking Arguments in Ethics*. New York: CUP.
- Singer, P. (2005). 'Ethics and Intuitions', *Journal of Ethics*, 9: 331–52.
- Sosa, E. (1999). 'How to Defeat Opposition to Moore', *Philosophical Perspectives*, 13: 141–54.
- Srinivasan, A. (2015). 'The Archimedean Urge', *Philosophical Perspectives*, 29: 325–62.
- Street, S. (2006). 'A Darwinian Dilemma for Realist Theories of Value', *Philosophical Studies*, 127: 109–66.
- (2008). 'Reply to Copp: Naturalism, Normativity, and the Varieties of Realism Worth Worrying About', *Philosophical Issues*, 18: 207–28.
- Thurrow, J. (2013). 'Does Cognitive Science Show Belief in God to Be Irrational? The Epistemic Consequences of the Cognitive Science of Religion', *International Journal for Philosophy of Religion*, 74: 77–98.
- (2014). 'Some Reflections on Cognitive Science, Doubt, and Religious Belief', in J. B. R. Trigg (ed.), *The Root of Religion*. London: Ashgate.
- Tolliver, J. (1982). 'Basing Beliefs on Reasons', *Grazer Philosophische Studien*, 15: 149–61.
- Vavova, K. (2018). 'Irrelevant Influences', *Philosophy & Phenomenological Research*, 96: 134–52.
- White, R. (2010). 'You Just Believe That Because. . .', *Philosophical Perspectives*, 24: 573–615.
- Williams, B. (2004). *Truth and Truthfulness: An Essay in Genealogy*. Princeton: Princeton University Press.
- Williamson, T. (2000). *Knowledge and Its Limits*. Oxford: OUP.

*University of Stavanger, Norway*