14 Global Debunking Arguments

Andrew Moon

14.1 Introduction

I believe that blue whale tongues can weigh as much as an elephant.[1] If I learned that my source of information for this belief was unreliable, I would have a *defeater* for this belief. It would no longer be justified.

A *defeater* for a belief is a reason S has that makes S's belief unjustified.[2] A defeater can make many beliefs unjustified at once. S has a *global defeater* if and only if S has a defeater for *all* of her beliefs. A *global debunking argument* is an argument that concludes that S has a global defeater.[3] A *local debunking argument* concludes only that S has a defeater for some portion of her beliefs. The topic of this chapter is global debunking arguments. I aim to explore their features and how to respond to them.[4]

In Section 14.2, I explain how to respond to a global debunking argument after it has given one a global defeater. In Section 14.3, I distinguish three types of global debunking argument and explain some of their strengths and weaknesses. In Section 14.4, I focus on one of the three types of global debunking argument mentioned in Section 14.3 and explain how to respond to it.

14.2 Global Defeaters Cannot Be Defeated

In Section 14.2.1, I give examples of global debunking arguments. In Section 14.2.2, I explore ways to respond to them *after* they've given one a global defeater.

14.2.1 Global Debunking Arguments, the Evolutionary Argument against Naturalism, and the XX Pill Case

Some might think there are no global debunking arguments. Katia Vavova (2015: 105) writes: "The debunker's argument must ... be *targeted*. It should threaten only moral realists' moral beliefs. ... This distinguishes

debunkers from other, less modest, skeptics." Here's my reconstruction of Vavova's argument:

(P1) If *debunking arguments* are distinct from *skeptical arguments*, then debunking arguments must be local (or targeted).
(P2) Debunking arguments are distinct from skeptical arguments.
(C) Debunking arguments must be local.

Therefore, there are no global debunking arguments.[5] Vavova cares about distinguishing skeptical arguments from debunking arguments because the former are easier to dismiss than the latter. She writes:

Consider the Moorean dogmatist who dismisses the skeptic thus: *I might be a handless brain in a vat you say? But look — wriggles hand — I'm not! We cannot respond to the optometrist similarly: I might be colorblind you say? But look — points to a color — I'm not!*

(2015: 106)

I agree with Vavova that the Moorean response to the skeptic is more plausible than an analogous response to one's optometrist. This provides some support for (P2).

I'm skeptical of (P1). Vavova's discussion raises the interesting question of how to distinguish what philosophers typically call "debunking arguments" from what they typically call "skeptical arguments." But even if there is an important difference, it need not be based on locality or globality. Vavova helpfully suggests two more differences. First, debunking arguments are *empirically* driven (often appealing to considerations from evolutionary theory or psychology) while skeptical arguments are *a priori* (sometimes appealing to a *possible* evil demon). Second, debunking arguments claim *probable* error while skeptical arguments only claim mere *possible* error. These are better candidate explanations for why the debunker's argument (including the optometrist's) is more immune to a Moorean response than the skeptic's arguments because they both appeal to something about *the strength of the evidence* backing the debunker's argument. On the other hand, globality and locality are irrelevant to the strength of the evidence supporting the debunker's argument, and hence to their immunity to Moorean responses. So (P1) is false or, at best, unsupported.[6] I now set aside Vavova's concerns about the possibility of global debunking arguments and give some concrete examples in favor of their existence.

Probably, the most discussed global debunking argument is Alvin Plantinga's evolutionary argument against naturalism (EAAN).[7] I'll spend some time explaining EAAN and some of the literature on it because there are many insights about global debunking arguments that we can glean from that literature. Let 'naturalism' and 'N' denote the view that

there are no supernatural beings, 'E' denote the view that humans came to exist in the ways described by contemporary evolution, and 'R' denote the view that human cognitive faculties are generally reliable, i.e., they form mostly true beliefs over false beliefs. Plantinga also sometimes uses 'R' to refer to the faculties of a specific individual rather than the faculties of human beings as a whole. In this chapter, context (or my clarification) will determine which is being referred to. Here is Plantinga's (2011a: 344–345) formulation of the premises of the argument:

(1) P(R/N&E) is low.
(2) Anyone who accepts (believes) N&E and sees that P(R/N&E) is low has a defeater for R.
(3) Anyone who has a defeater for R has a defeater for any other belief she thinks she has, including N&E itself.
(4) If one who accepts N&E thereby acquires a defeater for N&E, N&E is self-defeating and can't rationally be accepted.
(5) Conclusion: N&E can't rationally be accepted.[8]

Plantinga's (2011a: 316–339) defense of premise (1) is the longest. Here's my brief summary. Given naturalism, materialism is probably true. Given materialism, a belief has both neurophysiological properties (NP properties) and the property of having a content. Given evolution, the properties of a creature are selected for because they bring about adaptive behavior. Given materialism, beliefs that bring about adaptive behavior do so because of their *NP properties*. Hence, beliefs (and the cognitive faculties that produce those beliefs) will be selected for because of the NP properties of those beliefs.

Now, if content properties *just are* NP properties (i.e., if *reductive materialism* is true), then beliefs will be selected for because of their content properties. And if content properties are *determined by*, but not reducible to, NP properties (i.e., if *nonreductive materialism* is true), then beliefs will not be selected for because of their content properties. The important thing is that in neither case is a belief selected for because its content is *true*. To illustrate, suppose an organism forms a belief in response to the presence of a predator. Suppose this belief has certain NP properties that result in adaptive behavior (e.g., fleeing). Plantinga writes,

But why think it is a *true* proposition that is determined by those NP properties? ... [T]he content generated by the NP properties of this structure, on this occasion, need have nothing to do with that predator, or with anything else in the environment. ... [A]s long as the NP properties are adaptive, it doesn't matter, for survival and reproduction, what content is determined by those NP properties. It could be true content; it could be false content; it doesn't matter.

(2011a: 330–331)

Given that the truth or falsity of the content of a belief plays no role in determining adaptive behavior, it is unlikely that evolution would select for cognitive faculties that produce mostly beliefs that are true (or have true contents).[9] So, P(R/N&E) is low.[10]

Plantinga's (2011a: 340–348; 2011b) defense of premise (2) rests on analogies like his well-discussed XX *pill case*:

You learn that a pill, called 'XX', destroys the cognitive reliability of 95 percent of those who ingest it. You take the pill and come to believe both that *I've ingested XX* and *P(R/I've ingested XX) is low*.

Intuitively, those beliefs give you a defeater for R. (This seems true even if you actually *are* one of the immune 5%; it is your *belief* about having taken XX that is the defeater.) Plantinga reasons that just as you have a defeater for R in the XX pill case, the naturalist also has a defeater for R.

Premise (3) states that anyone who has a defeater for R has, in my terminology, a global defeater. We therefore have two potential examples of global defeaters and two corresponding global debunking arguments: one for the naturalist reflecting on EAAN and one in the XX pill case. (Note that both appeal to defeaters that arise from empirical considerations and beliefs about low probabilities, which make them paradigm debunking arguments.) These examples will be a basis for discussion about global debunking arguments more generally in the next section.

14.2.2 Three Theses About Defeating Global Defeaters

I address two questions in this section. By answering them, I will have explained how one can respond to a debunking argument even after one has gained a global defeater. First, is global defeat possible? Second, supposing it is possible, how, if at all, could one have justified beliefs again after one has a global defeater?

My answer to the first question is 'yes':

Thesis 1
Global defeat is possible.

Here is an objection to Thesis 1. Suppose EAAN really is a global defeater for the naturalist. Then EAAN is a self-defeating argument. A *self-defeating argument* for S is one in which, by believing its conclusion, S is no longer justified in believing at least one of its premises. If a naturalist believes the conclusion of EAAN, then she will have a defeater for all of her beliefs, including her beliefs in the premises of EAAN. It is then tempting to conclude that since the argument is self-defeating, it no longer has defeating power. In slogan form, "Self-defeating arguments *can't*

defeat!" And what is true of EAAN is true of all global debunking arguments. Therefore, global defeat is impossible; any such defeater will defeat itself.

The view that *self-defeating arguments can't defeat* is implicitly endorsed by some in the debunking literature. Regarding debunking arguments against our *evaluative beliefs*, Katia Vavova writes: "But in aiming to debunk all of our evaluative beliefs, the debunker leaves us with nothing with which to evaluate whether those beliefs have been debunked" (2014: 89). Christos Kyriacou (2016) and Silvan Wittwer (2020) agree with Vavova; they also examine ways for the evaluative debunker to avoid self-defeat but to no avail. Now, these philosophers do not explicitly say, "Evaluative beliefs cannot be defeated by self-defeating debunking arguments," but this is implied by their subsequent dismissal of those debunking arguments on that basis. Jack Woods is more explicit:

I argue that certain species of belief, such as mathematical, logical, and normative beliefs, are insulated from a form of Harman-style debunking argument. ... [T]he very construction of Harman-style skeptical arguments requires the truth of significant fragments of our mathematical, logical, and normative beliefs. ... Given this property, Harman-style skeptical arguments against logical, mathematical, and normative beliefs are self-effacing; doubting these beliefs on the basis of such arguments results in the loss of our reasons for doubt.

(2018: 47)

By saying the beliefs are *insulated*, Woods is implying that they are not defeated by such arguments.

My response, in slogan form, is "Self-defeating arguments *can* defeat!" Consider the XX pill case, in which I believe that I took XX and that P(R/I took XX) is low. As a result, I gain a defeater for R. As a result of that, I gain a defeater for all my beliefs, including my belief that I took XX. Does this mean I am off the hook and no longer have to worry about having taken the XX pill? Of course not. Yes, the defeater for R is a global defeater, so I have a defeater for the argument's premises. But that does not imply that I am now somehow back to being *justified* in believing R. My belief in R, as well as my other beliefs, remain defeated.[11]

I conclude that Vavova, Wittwer, Kyriacou, and Woods provide a faulty way of responding to debunking arguments. I suspect they erred because they focused only on local debunking arguments. Global debunking arguments, like in the XX pill case, help us see that self-defeating arguments *can* defeat, which, in turn, helps block an objection to EAAN. So, the first objection to Thesis 1 fails. Furthermore, the XX pill case is strong support *for* Thesis 1.[12]

The second question of this section is how, if at all, one can have justified beliefs again after one has a global defeater. I agree with a point by Plantinga, which I formulate as follows:

Thesis 2
If S has a global defeater, then S cannot "get our" of defeat by reasoning or by a defeater-defeater.

This is because any potential defeater-defeater or basis for reasoning will have already been defeated by the original defeater for R. A person in this situation is epistemically screwed.

Some revisions of the XX pill case might make us question the conjunction of Thesis 1 and Thesis 2.[13] Plantinga (2002: 227–228) considers a case in which, shortly *after* I take XX (and, hence, come to believe *I took XX*), it seems to me that my trustworthy doctor tells me that I am one of the immune 5%. In this case, can any beliefs based on the apparent testimony of my doctor serve as defeater-defeaters? Plantinga thinks not, and I agree.[14] These new beliefs arising from the apparent doctor's testimony are defeated by my belief that *I took XX*.[15]

Here's another attempt at a defeater-defeater. Suppose that shortly after you take XX, you look around and note how *coherent* your experiences and beliefs seem.[16] Couldn't your belief that *my beliefs and experiences seem coherent* serve as a defeater-defeater? (Call the content of this belief 'STC'.) Plantinga (2011c: 169) replies that such a belief cannot be a defeater-defeater since P(R/STC & I took XX) is still low. To support this claim, Plantinga asks us to imagine a third-person case in which someone, Sam, has taken XX, and it also seems to Sam that his beliefs and experiences are coherent. We should think that the probability that Sam's faculties are reliable, given this information, would still be low. This reasoning applies to ourselves if *we* took XX and believed STC.

I am skeptical of Plantinga's claim that that probability is low. If it does seem to me (or Sam) that my beliefs and experiences are coherent, then this is evidence that I am one of the immune 5%, which, in turn, is strong evidence for R.[17] It seems far more likely that my beliefs and experiences would seem coherent, given that I am one of the 5%, than if I were not part of the 5%. The XX pill could render my faculties unreliable in a vast number of ways, with only a tiny number of those ways involving seemingly coherent beliefs and experiences. (My point here is even more plausible if we also specify that it seems to me that my doctor is telling me that I am one of the immune 5%.) So, I am skeptical of Plantinga's argument. My reason for thinking that belief in STC cannot serve as a defeater-defeater is simpler. As said before, it is already defeated by the original global defeater for R, so it is in no position to do any defeater-defeating.[18]

William Talbott suggests that coherence itself, not *beliefs about* seeming coherence, can help one avoid defeat. He writes:

> Coherent memories and experience help make our beliefs rational without our having to have any beliefs about the coherence. ... I believe that many nonhuman animals and young children have beliefs made rational by the coherence of their experience and memory, though they surely have no belief about that coherence.
>
> (2011b: 174)

A strength of this response is that, unlike the earlier responses, it does not claim that a *defeated belief* can be a defeater-defeater; it appeals directly to memories and experiences to play that role.

In reply, I argue that experiences alone cannot serve as defeater-defeaters. Suppose Watson is initially justified in believing that *Moriarty is the murderer*. Watson then hears testimony from Mary, whom he justifiedly takes to be a reliable source, that Moriarty is innocent. Watson now has a defeater for his belief that Moriarty committed the crime. Watson then has a series of experiences and memories that, if put together, would indicate that Mary is lying. However, it would take a good amount of reflection to put it all together before Watson has his "Aha, Mary was lying!" moment. Before this moment, it would not be reasonable to go back to believing that Moriarty is the murderer. *Only after* that moment would he have a defeater-defeater and go back to justifiedly believing that Moriarty is the murderer. The lesson to draw is that experiences and memories alone, if not "put together" in one's mind as the basis for belief, are unable to serve as defeater-defeaters in the way Talbott describes.

I admit that the view that a global defeater can never be defeated seems odd. Suppose it's been 30 years since I came to believe that *I took* XX. Suppose I actually *was* one of the immune 5%! I have to say that all the beliefs I formed through these decades—the beliefs that I have a family or that the earth is round or that I have a headache—are all unjustified. This seems like a bullet to bite.

I can soften the bullet by saying that global defeaters can still be eliminated; it just won't be by a defeater-defeater. Plantinga gives this example:

> I suffer from a delusion, thinking I've been captured and envatted by Alpha-Centaurian cognitive scientists; I then have a defeater for R. *But I recover from this deplorable condition*, recognizing it for the delusion it was; I no longer think I am a brain in a vat; now I accept R in perfect rationality.
>
> (2002: 229, my emphasis)

In this example, it is Plantinga's *recovering* or *healing* from mental illness, not reasoning, that saves him. Plantinga writes: "There is no way to reason oneself out of such a predicament; here salvation will have to be by grace rather than works" (2002: 230; cf. Plantinga 1993: 237).

Here's another example. Suppose Lenny takes LSD. His belief that he recently took LSD is a defeater for R. Now, he should not trust his memory, his perceptions, or his reasoning. What can he do? He shouldn't reason, since he can't trust his reasoning. Rather, he should go to sleep so he can wear off the effects of the drug. After waking up, his cognitive faculties will once again function properly, and he can know things again. His cognitive faculties will have, to some extent, *reset* after his good night's sleep, and the proper function of his cognitive faculties will now, on their own, delete the belief that *I recently took LSD* and replace it with the belief that *I took LSD last night but am now doing better*. It is these *non-reasoning* factors that rid Lenny of the defeater, not his *reasoning*. Furthermore, in this case, there is something Lenny can *do*—go to sleep—to aid in his eliminating the defeater for R.[19]

These points lead to the third thesis:

Thesis 3

There are cases in which S has a global defeater, and S can act in *epistemically good* ways to rid herself of the defeater.

In Plantinga's Alpha Centauri case, he just *recovers* and doesn't need to *do* anything. The epistemically good state of affairs just *occurs*. In the case of Lenny, however, he can still *act* so as to bring about epistemically good states of affairs. So, in some cases of global defeat, we can act so as to bring about an epistemically good state of affairs in which we eliminate the defeater. Hence, perhaps I should have said above that the person with a global defeater is epistemically screwed *if* he limits himself to using *reasoning*. Since our epistemic options are not always exhausted by what's available via reasoning, this should soften the bullet the proponent of Theses 1 and 2 must bite.

Here is some fruit from this discussion. We can think of general strategies for epistemically good (and healthy) ways of eliminating defeaters, and this need not be by way of reasoning. Sometimes, you might just need to take a nice fresh walk out into nature, properly hydrate yourself, or get a good night's rest. In many cases, you will need help from another person: Lenny's friend might help him get home to bed. Perhaps you can be that friend for another person. Epistemic flourishing might occur by way of community. Or perhaps, as Plantinga slyly hints, what is needed is God's intervention, and you must be saved by grace. There are plenty of healthy ways to nurture our human and epistemic flourishing, either by

our own deeds, the deeds of others, or other external factors. So, our epistemic flourishing depends heavily on factors outside ourselves and our agency, but we already knew that justified belief and knowledge depended on such external factors anyway.[20]

14.3 Three Types of Global Debunking Argument

So far, we've considered how to respond to a global debunking argument, *after* it has given one a global defeater. But how does one *avoid* getting defeaters from global debunking arguments in the first place? That question is the focus of Sections 14.3 and 14.4. In Section 14.3, I distinguish between three types of global debunking argument and devote a section to each of the three types. I explain some of their strengths and weaknesses. Along the way, I reply to Bergmann's (2002) criticism of EAAN.

14.3.1 Pure-Undercutter

One has a *rebutting defeater* for p only if one has evidence for ~p. One has an *undercutting defeater for p* if one has a significant reason to doubt that one's belief in p was formed reliably. While a rebutting defeater for p is evidence for ~p, an undercutting defeater for p is merely a reason to withhold p. (Note: strictly speaking, a defeater won't be for a *proposition* p, but for someone's *belief that* p, but I'll use the looser language for ease of expression.)

The first type of global debunking argument employs the following notion:

Pure-Undercutter

S's belief in R has a pure-undercutter = S's belief in R has an undercutting defeater and no rebutting defeater.

Suppose Una is asked to consider R and then comes to believe R for the first time. A trustworthy scientist then notifies her that she ingested a pill, called the 'X pill', in her lunch earlier in the day. The X pill only affects the reliability of the faculties that are responsible for forming belief in R. It does not affect *other* faculties, like one's vision or memory. After learning this, it seems that Una gets an undercutting defeater for R. She also gets no rebutting defeater for R because she doesn't get any evidence for ~R. She just got evidence that she formed her belief in R unreliably. So, this is a *pure-undercutter* for R. It seems that she then gets a defeater for all her beliefs, and so she has a global defeater.

A benefit of this sort of global debunking argument is that it doesn't require showing that most or all of one's faculties are not reliable, only

that a specific subset of one's faculties is. But what are those? To explore this, we'll have to explore how R comes to be justified in the first place.

Now, some might object from the very start that belief in R can't be justified because such a belief is *epistemically circular*: one must use one's faculties to form the belief that one's faculties are reliable. In reply, I'll simply *assume* in this chapter that epistemic circularity does not preclude the possibility of justified belief in R. This assumption is justified for two reasons. First, the view that epistemic circularity does not disqualify a belief from being justified is already ably defended by a large number of philosophers; it's probably the majority view.[21] Second, I am currently exploring how belief in R might get a pure-undercutter *given that* R is already initially justified. The question of *whether* it can be initially justified is not the topic at hand.

So, supposing that belief in R *can* be justified, *how* is it justified? According to Bergmann (2002: 66–68; 2006: 206–211; 2021: 122–126), we have strong nonpropositional evidence for R. According to his view, inspired by Thomas Reid, a person first considers the possibility of ~R, perhaps by considering the possibility that she is beguiled by a demon or some other skeptical hypothesis. She then experiences what Reid calls an *emotion of ridicule*, where it *seems* to her that ~R is absurd. On the basis of this seeming, she comes to justifiedly believe (and know) R noninferentially. In this case, R is supported by a strong seeming.[22]

Bergmann's view is defended at length in the above references and worth taking seriously. I describe it not because I endorse it, but because it gives us a concrete idea of how to understand that the X pill targets Una's undercutters more generally. We can imagine that the X pill case and pure-undercutters more generally. We can imagine that the X pill targets Una's emotions of ridicule so that there is no correlation between ridiculous and false propositions and the presence of seemings of absurdity. When she learns about the properties of the pill, she will get an undercutting defeater for R.

Global debunking arguments featuring pure-undercutters for R are unexplored and worth considering. Here, I'll just briefly lay out what the general structure of such arguments should be and how to respond to them. They will include two things: (1) an empirical premise stating how S comes to believe R and (2) a premise that says that that process is unreliable. Bergmann gave us one story that describes (1), but future research should evaluate whether he is right and whether there are good competitors. Regarding (2), the assessment of reliability will depend on what the relevant process is. If Bergmann is right, then we should look into how much we can trust such emotions or seemings. And the way to avoid a pure-undercutter from such an argument would be to *either* show that the debunker is wrong that belief in R is produced the way the argument says *or* to show that the relevant process is reliable after all.

14.3.2 The Undercutter-Because-Rebutter

One might expect the next type of global debunking argument to feature a pure-rebutter for R:

Pure-Rebutter

S's belief in R has a *pure-rebutter* = S's belief in R has a rebutting defeater and no undercutting defeater.

But global debunking arguments won't feature pure-rebutters for R. Consider an example of a rebutting defeater for R: the XX pill case. In it, you get strong probabilistic evidence for ~R. But is this a *pure-rebutter*? No. Once you get the rebutting defeater for R, you *then* gain a significant reason to doubt, for each of your beliefs, that it was formed reliably. This includes your belief in R. You thereby have an undercutting defeater for R that depends on a rebutting defeater for R. So, you have no pure-rebutter.

I'll give this sort of defeater a different name:

Undercutter-Because-Rebutter

S's belief in R has an *undercutter-because-rebutter* = S's belief in R has both a rebutting defeater and an undercutting defeater, but R only gets an undercutting defeater *because* R gets a rebutting defeater.

It seems that any global debunking argument that features a rebutting defeater for R will end up featuring an undercutter-because-rebutter for R for the same reason it does in XX pill case. So, global debunking arguments that feature pure-rebutters are unlikely to exist.[23] On the other hand, we have now specified the second type of global debunking argument: one that features an undercutter-because-rebutter for R. This naturally raises the question of whether there are any *rebutters-because-undercutters* for R. I don't think so. An undercutting defeater for R is a reason to withhold R; that will result in a defeater for all of one's beliefs. This isn't evidence *for* ~R, and so, not a rebutting defeater for R.

What else can we say about the features of a global debunking argument featuring an undercutter-because-rebutter? We can learn some lessons about them from EAAN and Bergmann's criticism of EAAN. Given that the XX pill case is Plantinga's main analogy for the naturalist's defeater, it is natural to interpret EAAN as also featuring an undercutter-because-rebutter. The evidence from N&E gives the naturalist a rebutting defeater for R, which then results in an undercutting defeater for all of her beliefs, including belief in N. Plantinga doesn't explicitly say this, but from our earlier discussion, it is natural to conclude that the naturalist

would also have an undercutting defeater for R, and so the naturalist has an undercutter-because-rebutter for R (if EAAN succeeds).

Bergmann (2002: 73–76) criticizes EAAN by arguing that one's non-inferential, intuitive evidence *for* R outweighs the probabilistic evidence *against* R. He compares it to a case in which I have non-inferential, perceptual evidence that I *drew a 3-inch straw*. Even if I know *I drew the straw from a lottery of varyingly sized straws* and P(*I drew a 3-inch straw | I drew from a lottery of varyingly sized straws*) *is low*, I don't have a rebutting defeater for my belief that *I drew a 3-inch straw* because of my strong perceptual evidence. Bergmann (2002: 75–76) takes the intuitive evidence for R to be at least as strong as the perceptual evidence that I *drew a 3-inch straw*, and the probabilistic evidence against R to be weaker than the probabilistic evidence that the straw is not 3 inches.[24] And if the naturalist has no rebutting defeater for R, then the argument fails. (Using our terminology from earlier, with no rebutting defeater for R, she won't have an undercutting defeater for R, since this is an undercutter-*because*-rebutter for R.)

I now draw two lessons about global debunking arguments that feature undercutters-because-rebutters. First, they are relatively straightforward because they are simply a matter of weighing the evidence for and against R. Second, for a successful rebutting defeater for R, the evidence against R must be *very* strong to outweigh the positive justification for R. If a good and trustworthy friend told me that ~R, this testimonial evidence alone would not give me a rebutting defeater for R; I might laugh at the friend. A rebutting defeater for R requires something like the strong evidence I have regarding the XX pill.

14.3.3 Undercutters-While-Rebutters

In this section, I explain the third type of global debunking argument. This will result in a *new* version of EAAN, which helps avoid Bergmann's objection to Plantinga's version of EAAN, mentioned in the previous section.[25]

Omar Mirza (2008: 141–145) replies to Bergmann that if we understand the potential defeater for R in EAAN not as a rebutting defeater, but as an undercutting defeater, then it can undermine the intuitive evidence for R by being evidence that the faculties producing the intuitive evidence for R is not reliable. He thinks this would distinguish EAAN from Bergmann's straw case and protect it from his objection. However, Mirza doesn't explain why any undercutting defeater the naturalist has for R wouldn't just depend on the rebutting defeater for R and so still be susceptible to Bergmann's initial objection.[26] Still, Mirza's suggestion to appeal to an undercutting defeater points us in the right direction.

I'll now present my new version of EAAN. There's an important disanalogy between the believer in the XX pill case and the naturalist

consuming EAAN, as I said before, the former gets a standard undercut-ter-because-rebutter for R. The latter, I'll show, gets what I call an *under-cutter-while-rebutter* for R.

Undercutter-While-Rebutter

S's belief in R has an undercutter-while-rebutter = S's belief in R gets both a rebutting defeater and an undercutting defeater, but S has an undercutting defeater for R that does not depend on S's having a rebutting defeater for R.

Recall Plantinga's defense of his first premise. Given naturalism and materialism, evolution selects for beliefs (and whatever faculty that pro-duces them) regardless of whether those beliefs are true. Thus, the prob-ability that any *specific faculty* is reliable (or produces beliefs with true contents), given N&E, is low. In this case, there is not a rebutting defeater for R *first*, which then brings about an undercutting defeater for each belief (including belief in R). One does not *first* gain evidence for a propo-sition about one's *collection* of faculties, from which one gains reason to doubt the reliability of each *individual* faculty (including whatever facul-ties produce belief in R). So, this is not an undercutter-because-rebutter for R.

Instead, the same evidence from N&E *both* makes R improbable (this is a rebutting defeater for R), *and also* directly gives the naturalist reason to doubt, of each of her beliefs (including belief in R), that it was formed reliably (this is an undercutting defeater for R). Here, you have both a rebutter and an undercutter for R, but the latter doesn't *depend* on the former. So, my version of EAAN, at least initially, features undercutters-*while*-rebutters, not undercutters-*because*-rebutters. With my version of EAAN, one's initial evidence applies to the collection *and also* directly applies to the individuals.

Let's return to Bergmann's objection. If he is wrong that the intuitive evidence for R is strong enough to outweigh the probabilistic evidence against R, then we have an undercutter-while-rebutter. We have both a rebutting defeater for R and an undercutting defeater for R, and the latter does not depend on the former. But suppose he is right that the intuitive evidence for R outweighs that probabilistic evidence against R. Then, even if there is no rebutting defeater for R, R can still have an undercut-ting defeater. So, my version of EAAN features an *undercutter-while-rebutter* for R if Bergmann is wrong about the evidence for R, and a *pure-undercutter* for R if he is right. Thus, there are two paths for R to get a defeater; the undercutter is a fail-safe in case the rebutter gets out-weighed by the evidence for R.

To illustrate these points, add to Bergmann's straw example that most of the straws are *tricky straws*, purchased from the local magic shop.[27]

Tricky straws appear or feel longer (or shorter) than they actually are. In the original case, Bergmann's probabilistic evidence that he didn't draw a 3-inch straw was easily overwhelmed by his perceptual evidence that he did. So, there was no rebutting defeater. However, in my new version of the straw example, he is probably not reliably perceiving the correct length of the straw, and so he still has an undercutting defeater for his belief that he drew a 3-inch straw.

Now, my strategy won't make it easier to block Bergmann's objection if the following conditions are met. Let the proposition that *S's belief in R was formed reliably* be picked out by 'R*'. S has an undercutting defeater for R, in the sort of case we're describing, only if S has sufficient evidence against R*. Now, if the evidence for R* is just as strong as the evidence for R, then it will be just as hard to undercut R as it is to rebut R. More specifically, if N&E cannot rebut R because of strong evidence for R, then N&E might not undercut R because of strong evidence for R*.

Fortunately, those conditions are not met. The evidence for R is signifi-cantly stronger than the evidence for R*. Suppose that R* is supported by a seeming, like R is.[28] ~R still seems much more absurd than ~R*. The probability of *just* the specific faculties producing belief in R being unreli-able, on a given occasion, seems much higher than the probability of our faculties in general being unreliable. We see individual operations of fac-ulties misfiring all the time; we rarely see our faculties being *generally* unreliable. Or think of the strength of evidence that would be required to show that you formed your belief in R unreliably. You might just need a good argument that your emotion of ridicule (or a specific faculty) was not reliable on a specific occasion. On the other hand, it seems that you would need much stronger evidence to convince you that your cognitive faculties are not *generally* reliable. So, it will be easier for N&E to under-cut R (by rebutting R*) than to rebut R directly.

Note that I am not arguing that my version of EAAN *does* successfully refute Bergmann's objection. That depends on whether the evidence for R* is overwhelmed by the evidence against it from N&E. And that depends on the strength of Plantinga's defense of the first premise of EAAN; settling that is beyond the scope of this chapter. So, I'll just say that my new version of EAAN, which appeals to undercutters-while-rebutters, is significantly more immune to Bergmann's criticisms than Plantinga's version.

I end this section by noting how subtle the difference can be between an undercutter-because-rebutter and an undercutter-while-rebutter. The XX pill case features the former but is easily revised to feature the latter. Suppose you learn of the XX* pill, which has a 95% chance of rendering your perception unreliable, a 95% chance of rendering your reasoning unreliable, a 95% chance of rendering your emotion of ridicule unreli-able, and so on. Believing you took this pill will give you an undercutter-while-rebutter for R.

Note the differences. In the XX pill case, R is defeated only if the evidence against R overwhelms the evidence for R. If the evidence for R overwhelms the evidence against R, then there is no basis for either rebutting R or undercutting R (i.e., no basis for rebutting R*). In the XX* pill case, on the other hand, one's evidence can directly undercut the evidence for R (by rebutting R*). Now, in both of these cases, there *is* sufficient evidence against R to overcome evidence for R and, hence, to rebut R. So, defeat of R happens in both cases, and their difference is easily missed. However, in EAAN, the potential rebutter for R comes not in the form of clear scientific evidence but a subtle philosophical argument. Even if the evidence against R cannot overcome the evidence for R (so there is no undercutter-because-rebutter), it might overcome the evidence for R* (so that there is a pure-undercutter). Here, the prospects for defeat depend on the distinction.

14.4 The Conditionalization Problem

14.4.1 Explaining the Problem

Global debunking arguments that feature undercutters-while-rebutters face a philosophical problem that has been discussed in the EAAN literature. Recall EAAN's second premise:

Defeater Thesis
Anyone who believes *N&E* and *P(R/N&E)* *is low* has a defeater for R.

One might argue that even if the naturalist grants that R's probability is low, given some things she believes (*viz.* N&E), there could be *other* propositions she believes, on which R is probable (Plantinga 2002: 223–224; 2011b). More formally, even if P(R/N&E) is low (and so N&E is a *potential* defeater for R), a naturalist might believe some other proposition Y, such that P(R/N&E&Y) is high. According to the objector, Y then serves as a *defeater-deflector* for the potential defeater, i.e., *Y prevents* P(R/N&E) is low and N&E from being a defeater for R.[29]

What are some possible defeater-deflectors in the case of EAAN? John Perry suggests that Y could be the proposition that *we have won the evolutionary lottery*, where P(R/N&E&*we have won the evolutionary lottery*) is high.[30] Carl Ginet (1995: 407) suggests that Y could be R itself, where P(R/N&E&R) is 1. Note that Ginet's suggestion is something the naturalist presumably *knows* to be true; since we are assuming that this is *pre*-defeat, it seems that there should be nothing wrong with conditionalizing on R.

Although this move might seem fishy, Ginet suggests that theists make a similar move. Suppose we take 'austere theism' or 'A' to name the proposition that humans have been created by a powerful and knowledgeable being. Suppose that P(R/A) is low. Presumably, theists also believe other propositions, such as that we have been created by a perfect being who has created us in his image so that we are knowers as God is a knower. Call the conjunction of these propositions 'theism' or 'T'. Since P(R/A&T) is high, the theist might claim that T is a defeater-deflector. But, Ginet asks the following:

Why isn't it just as reasonable for the naturalist to take it as one of the tenets of naturalism that our cognitive systems are on the whole reliable ...? If it is cheating to include such a thing in naturalism, for the purposes of deciding whether it is reasonable to believe in naturalism, then why isn't it cheating to include the corresponding thing in theism, for purposes of deciding whether it is reasonable to accept it?

(1995: 407)

These are good questions. However, something is puzzling about these moves. Plantinga notes that the strategy, if generalized, allows for too much. For *any* time that R might get defeated by a probabilistic defeater, where P(R/X) is low, the person could just say P(R/X&R) is high; hence, the defeater is deflected! This is no good. In the XX pill case, one clearly cannot use R as a defeater-deflector. In that case, even if I believe that R *and P(R/I took XX & R) is high*, I still have a defeater for R.[31] Plantinga formulates what he calls the *conditionalization problem*:

Which beliefs Y are such that if P(R/N&E&Y) is high, then Y prevents *N&E&P(R/N&E) is low* from being a defeater for R?[32]

Plantinga writes: "That is a difficult question indeed; I certainly don't know how to give a complete and rigorous (or even a complete and unrigorous) answer to it" (2002: 224). My goal in Section 14.4 is to solve the conditionalization problem. However, instead of focusing specifically on it, I will focus on this related, more general question:

Conditionalization Problem*
Which beliefs Y are such that if P(R/X&Y) is high, then Y prevents *X&P(R/X) is low* from being an *undercutter-while-rebutter* for R?

By replacing 'N&E' with 'X', I can focus on global debunking arguments beyond just EAAN. And by replacing 'a defeater' with 'an undercutter-while-rebutter', I still focus on EAAN. Now, Plantinga's version of EAAN

featured an undercutter-because-rebutter for R, so I might not be attacking the problem exactly how Plantinga conceived it. However, my version of EAAN, which features an undercutter-while-rebutter for R, is a more powerful version of EAAN since it helps avoid Bergmann's objection.[33] Regardless, conditionalization problem* will be my focus.

Notice that conditionalization problem* is asking which propositions are *preventing* a proposition from being a defeater. So, the problem does not apply to cases in which there is no *potential* defeater for R to begin with. This might be in cases in which the believer is a little child or has never reflected on any global debunking arguments. The relevant proposition might also not be a potential undercutter-while-rebutter for R* if it is not strong enough to outweigh the evidence for R*. So, the question of defeater-deflection arises only if there is a potential defeater to be deflected in the first place.

14.4.2 Epistemic Origin Stories

My answer to the conditionalization problem* will employ the following concept:

S's *epistemic origin story* is a set of propositions that explains how it is that S's cognitive faculties have their degree of reliability.[34]

An *actual* epistemic origin story should be distinguished from a *believed* epistemic origin story, the epistemic origin story that one believes applies to her cognitive faculties.

Next, there is the *reliability-promoting epistemic origin story*, which is an epistemic origin story such that, conditional on it, there is a high probability that S's cognitive faculties are reliable. I will *not* believe a reliability-promoting epistemic origin story if I believe that I have been created by a demon that is intent on deceiving me. Similarly, suppose Plantinga is right about EAAN. Then the naturalist's believed origin story is not reliability-promoting; conditional on it, the probability that her faculties are reliable is low. Of course, most naturalists will disagree and think the probability is high that their faculties evolved so as to form mostly true beliefs about ordinary objects (flora, fauna, other humans, rocks) and their properties (whether they are dangerous, friendly, tasty, heavy, etc.). So, there will be disagreement about which origin stories are actually reliability-promoting.

An *explanans* can either entail or merely make probable its *explanandum*. Suppose Jane believes that she is a human and that God created and conserves the reliability of *all* humans' faculties. Then Jane believes a reliability-promoting origin story. Conditional on it, it is probability 1 that she has reliable faculties; it *entails* that she has reliable faculties. On the other hand, suppose Jane believes that although God *nearly always*

creates humans with reliable faculties, he sometimes allows them to have unreliable faculties. Then Jane would still believe a reliability-promoting origin story. Conditional on it, it is merely *highly probable* that her faculties are reliable.

14.4.3 A Solution to the Conditionalization Problem*

Here is my solution to the conditionalization problem*:

Epistemic Origin Story Solution (EO-Solution)

Even if S's belief in X *and P(R/X) is low* is a potential undercutter-while-rebutter for R, the belief that Y *and P(R/X&Y) is high* is a defeater-deflector for that potential undercutter-while-rebutter for R if and only if

(a) S is justified (prior to considering the argument at hand) in believing that Y and P(R/X&Y) is high.
(b) Y is part of S's believed epistemic origin story.[35]

In other words, a justifiedly believed *reliability-promoting epistemic origin story* will deflect a potential undercutter-while-rebutter for R. In what follows, I show how the EO-Solution gives plausible verdicts about a wide number of cases.

First, condition (b) of the EO-Solution explains why R itself cannot be a defeater-deflector. R is not part of one's believed origin story. It does not explain why S's faculties have their degree of reliability. Rather, R is the explanandum that the epistemic origin story—the explanans—is supposed to explain.

Second, the EO-Solution explains why Perry's I have won the evolutionary lottery is not an admissible deflector in the case of EAAN.[36] It does not actually explain (or is not part of an explanation of) why one's cognitive faculties are reliable. To say that something happened by chance is not to give an explanation at all. As such, one should not believe that it is part of one's epistemic origin story, and so condition (b) is not met.[37]

Third, Plantinga (2011b) considers whether various theories in philosophy of mind (Dretske's theory of content, functionalism, or Millikan's teleosemantics) could potentially serve as defeater-deflectors. Here's how such a theory, call it 'M', could serve as a deflector according to the EO-Solution. M, conjoined with N&E, might provide an explanation of how evolution would select for faculties because they produce beliefs with true contents. M would then figure crucially in an explanation of the degree of reliability of one's faculties. Then condition (b) could be met because M would be a part of the naturalist's believed epistemic origin story. Condition (a) could also then be met if the naturalist is justified in believing (prior to considering the argument at hand) both N&E&M and

that P(R/N&E&M) is high. Plantinga ultimately argues that there is no theory M in philosophy of mind that can play this role, but that is beside the point.[38] The point is that this is how a theory could be a defeater.[39]

This is a good place to make a structural point about EAAN. I am interpreting the appeal to M as an attack on EAAN's second premise. So, even if P(R/N&E) is low, R remains undefeated because M deflects the potential defeater. Note that one could also take the appeal to M to be an argument against the first premise, concluding that P(R/N&E) is not low after all.[40] I don't think it matters which way we interpret it, and the EO-Solution helps us see why. Plausibly, any potential undercutter-while-rebutter, X, for R will be a part of S's believed epistemic origin story. EO-Solution proposes that what is required to avoid defeat is other parts of that epistemic origin story, Y, which make R probable. Whether the appeal to P(R/X&Y) being high is to show that Y is a deflector (even if P(R/X) is low) or that P(R/X) is not low after all, what really matters for avoiding the potential undercutter-while-rebutter is a high conditional probability on one's justifiedly believed epistemic origin story.

Let us return to the earlier discussion between Plantinga and Ginet. Suppose Lenore justifiedly believes A (austere theism) and that P(R/A) is low. Can Lenore use belief in T (which includes as a conjunct that *God created me in his image*) to deflect the potential defeater? Note first that it's not clear that A and *P(R/A) is low* is even a *potential* undercutter-while-rebutter for R. But suppose it is. What does EO-Solution say about this case? Ginet makes it sound as if Lenore *arbitrarily* forms belief in T in order to deflect the potential defeater. But then condition (a) of the EO-Solution is not met because belief in T is not independently justified. So, T cannot be a deflector. However, unlike R or *we have won the evolutionary lottery*, T is *the sort* of proposition that could be a deflector if it was antecedently justified because T could be part of one's believed, epistemic origin story. So, suppose Lenore is *antecedently* justified in believing T.[41] Then even if Lenore believes P(R/A) is low, she is justified (prior to considering the argument at hand) in believing *T and P(R/A&T) is high*, and T is part of her justifiedly believed epistemic origin story. In such a case, EO-Solution can determine her to have a defeater-deflector. This seems right.[42]

The EO-Solution has the correct verdict about Talbot's (2011a: 161) ZZ example. Suppose that decades ago, a chemical compound, ZZ, entered the food chain, and exposed all human fetuses to a disease that would render each of their faculties unreliable. Fortunately, a strain of bacteria underwent a mutation that, if taken by a child, would render ZZ inert. The bacteria were cultured and given to every child. Decades later, an adult learns this bit of history and believes that *I ingested ZZ* and P(R/I ingested ZZ) is low; fortunately, she also believes that she ingested the friendly bacteria and that P(R/I ingested ZZ & I ingested the friendly bacteria) is high. Both Talbot and Plantinga (2011c: 171) agree that the potential defeater is deflected in this case, but neither develops a theory why. The EO-Solution explains why propositions about the friendly bacteria can serve as defeater-deflectors: they are part of her justifiedly believed epistemic origin story.

14.4.4 EO-Solution and Two Objections to EAAN

In this final section, I explain how EO-Solution helps answer two objections to EAAN. Then I make some concluding remarks.

First, Hendricks and Anderson (2020) question whether *God* would have an EAAN-style defeater for thinking his own faculties are reliable. Let 'D*' denote the proposition that "God, or anything at all like him, did not intervene (or order the world from the beginning) to ensure that his (God's) cognitive faculties are reliable" (2020: 494). Where 'R' in this case is about the reliability of *God's* faculties, Hendricks and Anderson suggest that P(R/D*) is low, and hence, that God has a defeater for R, and thereby, for all of his beliefs. They explore possible responses a theist could give to this challenge (2020: 494–498). (Although this is more of a question than an objection, the addition of the premise that God is omniscient results in *reductio* against the sort of reasoning used in EAAN.)

Here is my response. Let's suppose that D* and *P(R/D*) is low* was a potential undercutter-while-rebutter for God's belief in R. Could God get a defeater-deflector, one that appeals to an epistemic origin story? Admittedly, nothing is *temporally* prior to God's having reliable faculties because God is eternal. Fortunately, an epistemic origin story for God requires only something *explanatorily* prior. God doesn't just *happen* to have reliable faculties; God is the sort of being that *couldn't fail* to have reliable faculties, and this is because God *couldn't* be less than perfect.[43] So, God's having reliable faculties is explained by his essential perfection.[44] An eternal being that, perhaps *per impossibile*, has no explanation by way of a creator or designer and that is not essentially perfect, might consider R (as applied to itself) and find itself with a defeater for R and, thus, for all its beliefs. But *I am an essentially perfect being* is part of God's believed epistemic origin story, which can deflect the potential defeater.

Second, Erik Wielenberg objects to EAAN as follows:[45]

Plantinga argues, from the armchair, that any belief-forming creature produced by unguided evolution will probably tend to have lots of false beliefs. I'd say that if you want to know what sort of widgets a widget-maker is likely to produce, then you should consider the widgets it has made. Similarly, if you want to know what sort of creatures unguided evolution is likely to make, then you should consider the creatures it has made. It seems to me that all creatures on earth

have been produced by unguided evolution, and inspection of these creatures casts doubt on Plantinga's probability claim.

(Craig and Wielenberg 2021: 58)

According to Wielenberg, we can look around and *see* that unguided evolution has produced creatures around us with reliable faculties. Given this information, there is a high probability that our faculties are reliable. This calls EAAN's first premise—that P(R/N&E) is low—into doubt. Wielenberg takes himself to be attacking EAAN's first premise. As I noted earlier, we could also see him as attacking the second premise. Let 'O' denote the proposition that creatures on Earth have evolved so as to have reliable cognitive faculties; Wielenberg thinks the naturalist comes to know O through observation and science. His considerations, if correct, will show that P(R/N&E&O) is high, that O is a defeater-deflector and, hence, that even if P(R/N&E) is low, the naturalist fails to have a defeater. Wielenberg, on the other hand, takes O to be evidence that P(R/N&E) is not low and so evidence against EAAN's first premise.

Regardless of what is the better way to think about it, the following analogy will help us see that Wielenberg's criticism goes awry. Suppose Wilhelm is an early modern scholar who, after reading and studying his parents' favorite philosopher Leibniz, has come to believe that souls are causally disconnected from bodies, that bodies on Earth came to exist in the ways described by contemporary evolution, and that the devil shoots his psychophysical disharmony ray gun at every creature on earth to disrupt the psychophysical harmony between bodies and minds. (Let's stipulate that Wilhelm's beliefs about the devil's actions were somehow due to some sort of justifiedly believed scripture. Wilhelm had initially believed in Leibniz's famous *pre-established harmony* but came to reject that view.)[46] Call the conjunction of those beliefs 'DISHARMONY'. Plausibly, the P(R/DISHARMONY) is low, and it seems that Wilhelm has a defeater for R.[47]

Now, the following is *not* how Wilhelm should try to avoid the defeater. He should not say, "If you want to know what sort of creatures the devil-tampered process is likely to make, then you should consider the creatures it has made." He should not observe the creatures around him, see that they seem to have reliable faculties, think that this devilish evolutionary process has a good track record of producing reliable faculties, and conclude that P(R/DISHARMONY) is high. *Or* he should not try to conditionalize on O—the proposition that creatures on earth have evolved so as to have reliable cognitive faculties—and try to deflect the potential defeater by saying that P(R/DISHARMONY & O) is high. Both ways of avoiding the defeater seem illegitimate. Analogously, appealing to O seems to be an illegitimate way to avoid EAAN's defeater for the naturalist.

What *would* be a legitimate way of avoiding defeat? Suppose that Wilhelm, due to another justifiedly believed scripture, believes that God sent a multitude of angels to shield Earth's creatures from the devil's psychophysical harmony disrupter ray gun. (Call this proposition 'SHIELD'.) Wilhelm's beliefs in SHIELD and that P(R/DISHARMONY & SHIELD) is high could serve as deflectors. This would protect Wilhelm's belief in R from defeat. The case mirrors Talbott's ZZ example or an example in which a naturalist uses some justifiedly believed theory in philosophy of mind as an admissible defeater-deflector.

Why *can't* the naturalist appeal to O? O meets condition (a) of EO-Solution since it is justified by simple observation and science. So, the culprit must be (b). *Suppose that* Plantinga's considerations are correct. Then, on naturalism, a creature's having reliable faculties is not a part of its epistemic origin story. And this is because a belief's having true content (and thus, a faculty's being disposed to produce beliefs with true content) is not part of the explanation of why the naturalist's faculties have the degree of reliability they do. So, O will not be part of a naturalist's justifiedly believed epistemic origin story. One might object that Plantinga's considerations in favor of his first premise are *not* correct, but that would be to criticize a different aspect of Plantinga's argument. So, EO-Solution helps the defender of EAAN respond to Wielenberg's objection and helps reveal where the issue really lies.[48]

Suppose EO-Solution has correct intuitive verdicts about a wide variety of cases. *Why* is it correct? *Why* is the justifiedly believed epistemic origin story what R should be conditionalized on to determine whether R gets defeated? Recall that I've restricted our inquiry specifically to potential *undercutters-while-rebutters*. Given the powerful undercutting power of undercutters-while-rebutters, most bits of evidence one would like to appeal to to deflect the potential defeater will get undercut. To prevent defeat, one must prevent the undercutting from happening in the first place. The appropriate deflector to such a potential defeater will be the set of justifiedly believed propositions in one's justifiedly believed epistemic origin story, propositions that make probable and *explain* why R is *true*. This will prevent undercutting from happening in the first place. So, the EO-Solution not only gets the cases right, but there is an explanation for why it is right.

With the EO-Solution in place, I end this chapter by presenting the structural features of a global debunking argument featuring an undercutter-while-rebutter for R. The first premise will state that S's belief in X and P(R/X) is low is a potential undercutter-while-rebutter for R. The second premise will say that this undercutter-while-rebutter has no deflector, i.e., S's justifiedly believed epistemic origin story is not such that, conditional on it, the probability of R is high. The conclusion will be that S's belief in R has an actual undercutter-while-rebutter. It will follow

from this that S has a global defeater. The way to attack such an argument, naturally, will be to argue against the premises.[49]

Notes

1　See https://www.rd.com/list/weird-facts/.

2　See Bergmann (2006: 153–178) for a defense of the existence of defeaters from recent criticisms.

3　A global debunking argument could also conclude that S lacks justification for all her beliefs not because of a defeater, but because the beliefs are not *prima facie* justified. Roger White (2010: 575) calls these *blocking debunkers*, to contrast with *undermining debunkers*, which debunk by way of defeat. My focus in this chapter is undermining debunkers.

4　In particular, I will focus on global, *reliability* debunking arguments, which conclude that S has a global defeater because S has a defeater for believing her beliefs were formed reliably. What I say about global, reliability debunking arguments should apply to other global debunking arguments, which conclude that S has a global defeater because S has a defeater for believing her beliefs are *sensitive* or that they are *safe* or that they lack some other positive epistemic property. See Korman (2019) for different ways of formulating debunking arguments.

5　Of course, there are famous *local* skeptical arguments, such as Hume's arguments against induction or those targeting knowledge of the external world. Perhaps Vavova means to say that while skeptical arguments *cannot* be global, debunking arguments *can*, and this distinguishes them.

6　Here are some concerns about Vavova's suggestions. She focuses on *Cartesian* skeptical arguments, which aim to show *possible* error, but prominent traditional skeptical arguments (e.g., regress arguments or underdetermination arguments) imply more than mere possible error: they imply a *lack of justified belief*, just as many debunking arguments do. Yet, a Moorean response to them seems reasonable to many. Second, underdetermination arguments appeal to how things *appear* to us, and thus, seem to be empirical in some sense. At the end of the day, perhaps there is no clear dividing line between skeptical arguments (which are susceptible to Moorean responses) and debunking arguments (which are not so susceptible). In each case, we just need to examine the strength of evidence; arguments with empirical premises or that show a high probability of error *typically* contain stronger evidence than the typical skeptical argument. My point is that globality and locality are irrelevant.

7　See Plantinga (1993, 1996, 2000, 2002, 2011a, 2011b), Beilby (2002), Moon (2017b) and the references therein.

8　This statement of the argument has formal problems. Premise (1) plays no role in the inference to the conclusion. Also, the argument is invalid. At most, all that could be validly inferred from the premises is, "N&E can't rationally be accepted by anyone who accepts (believes) N&E and sees that P(R/N&E) *is low.*" Fortunately, no philosophically interesting issue hangs on this, and the argument can be easily made valid. (See Rea's [2002: 182] approach.)

9　Thanks to Dan Korman for helpful discussion. Michael Deem (2018) interprets Plantinga as reasoning from *possible* cases of evolution resulting in cognitive unreliability (*c. elegans* and a hominid race) to evolution *actually* resulting in human cognitive unreliability. He writes, "[S]imply imagining cases in which reliable cognitive processes would not be as favored … will not tell us whether selection actually favored cognitive reliability in the historical evolutionary lineages of hominins" (223). He says that just because a trait, such as *showy plumage*, is not *sometimes* selected for, it doesn't imply that it wasn't actually selected for (as in peacocks). Deem seems to miss that Plantinga's considerations about the implications of N&E, if correct, will apply directly to *any* evolving organism. This would make the selection of *reliable faculties* unlikely in *any* naturally evolving species, unlike *showy plumage*, which might be selected for in some species but not others.

10　See Plantinga (2002: 253–254, 258–261, 265) for an alternative defense of premise (1), which does not appear in EAAN's more recent defenses (Plantinga 2011a, 2011b).

11　An interesting corollary is that *unjustified beliefs* can be defeaters. Plantinga (2002: 272–275) defends this in the context of EAAN in response to an objection by Alston (2002). The view is also defended by Pryor (2004), Bergmann (2006), and especially Alexander (2017). My view might conflict with Alexander's, but exploring that is beyond the scope of this paper. Thanks to Luke Kallberg for helpful discussion.

12　Consider Richard Fumerton's (1995: 50–51) case in which you ask a Magic-8 ball whether it is reliable, and it says 'no'. Any basis for believing its testimony gets undermined by its own testimony. But clearly, you ought not now go on believing it's reliable. Rather, you still have a defeater for believing it is reliable. See Bergmann's (2021: 59–60, 65–73) excellent discussion of self-undermining global *skeptical* arguments, which appeal to Fumerton's case.

13　In what follows, I draw from cases discussed in a back-and-forth between Talbott (2002, 2011a, 2011b) and Plantinga (2002, 2011c). Their discussion is complex, and my own discussion would be overly long if I tried to capture all the parries and thrusts of the exchange. I recommend interested readers to consult these references for details.

14　Plantinga's (2002: 228) view about this case is subtler. He thinks the doctor's testimony would give him a *proper function* defeater-defeater (since a human with properly functioning faculties would still believe the doctor), but this would not be a purely alethic (truth-aimed) defeater-defeater (since a human with only truth-aimed properly functioning faculties at work would not believe the doctor). I take it that the alethic defeater-defeater is what's relevant here.

15　For an exploration of why you should trust the apparent doctor's testimony before believing you took XX but not after, see Moon (2017b: 212–213).

16　The appeal to coherence to avoid defeat is first suggested by Talbott (2011a: 158). Plantinga interprets Talbott as suggesting a belief *about* coherence (or seeming coherence) to be the defeater-defeater. Since Talbott (2011b) denies that interpretation, I do not here attribute this objection to him. I discuss Talbott's (2011b) clarification of his view in the following paragraphs.

17　Put probabilistically, P(I am one of the immune 5% / STC & I took XX) is rather high, and P(R/ STC & I took XX & I am one of the immune 5%) is very high. Using the Law of Total Probability, if these values are sufficiently high, then P(R/STC & I took XX) is high, regardless of the low probabilities of the remaining values (i.e., P[R/STC & I took XX & ~(I am one of the immune 5%)] and P[~(I am one of the immune 5%) / STC & I took XX]).

18　Given how clear and obvious the introspective belief in STC is, it might seem that it *can't get defeated*. In reply, note that I could still believe that the XX pill could make even my introspective beliefs unreliably formed so that I sometimes believe things seem a certain way when they don't, and *vice versa*. It seems that this could defeat my belief in STC. Second, even if such introspective beliefs *are* immune

19 I've been told that LSD doesn't work the way I say it does in this paragraph, and I couldn't think of an actual drug that has the exact effects needed for my case. I ask the reader to forgive my inexperience with and ignorance of psychedelic drugs and to imagine that LSD has some of the properties I am attributing to it here.

20 For a defense of this final claim, see any defense of externalism.

21 I have found Bergmann's (2002: 76–82; 2006: 206–211; 2021: 176–190) defenses the most convincing, especially as it relates to R. For more, see the references in footnote 4 in Bergmann (2002: 78), footnote 16 in Bergmann (2021: 182), and footnote 14 in Moon (2021: 792).

22 Keith DeRose (2017: 227–230), also inspired by Reid, affirms pretty much this same story as a way to come to know that one is not a brain in vat (and that other skeptical hypotheses are false), which is just a step away from justifiedly believing R.

23 You could get a rebutting defeater for R without getting an undercutting, global defeater in the following way: you knowingly take a pill that you know negatively affects the reliability of the faculties producing all your beliefs *except* your introspective beliefs. You would then get a rebutting defeater for R (since you have a reason to think that your faculties are not *generally* reliable), but you'd only get an undercutting defeater for your non-introspective beliefs. So, this would not be a *global* defeater. Thanks to Jeff Tolly for pointing out this sort of case to me.

24 Bergmann thinks the probabilistic evidence is weaker because he thinks Plantinga has only persuasively argued that P(R/N&E) is *inscrutable* (2002: 74). At the time of Bergmann's writing, Plantinga's (2011a, 2011b) more developed arguments that P(R/N&E) *is low* were unpublished.

25 This version is new, not because it *replaces* any of Plantinga's official premises. It is instead a new way to defend some of his premises.

26 Mirza thinks an undercutter for R in the case of EAAN would be an instance of what he calls "the process characteristic conception" (2008: 131–132). According to this conception, roughly, the defeater for R comes from S's reasons to doubt that her faculties were caused by factors that would make those faculties reliable. Mirza seems to think this is not a rebutting defeater for R because "in these situations none of her beliefs give S anything like strong evidence *against* her belief in the reliability" of her faculties (2008: 132). He then says that such reasons "cannot support anything more than a weak inference to the conclusion not-R." But the fact that S cannot *infer* ~R or does not have *strong* evidence against R doesn't imply that there's no rebutting defeater for R. Having a reason to doubt that your faculties were caused by factors that would make your faculties reliable still seems like evidence for ~R (i.e., a *rebutting defeater* for R).

27 Thanks to Jamie Fritz and Philip Swenson for help putting this case together.

28 See Bergmann (2021: 221). I am grateful to Bergmann for helpful conversation about these points.

29 Defeater-deflectors are first introduced in Plantinga (2002: 224). We should distinguish these from defeater-*defeaters* since, as we saw in the previous section, defeaters for R cannot get defeated.

30 Plantinga (2002: 221) notes that this is from discussion with Perry.

31 Some, perhaps inspired by Williamson (2000), might think S can never get a probabilistic defeater for *p* if S *knows p*. The XX pill case is a counterexample

32 I quote here from Plantinga (2002: 224), although he uses 'B' where I use 'Y', a view.

33 It is more accurate to say that it, at least initially, features an undercutter-while-rebutter for R, which might actually be a pure-undercutter if Bergmann is right about the evidence for R. In the paper, I'll stick with the less accurate characterization I just used in the main text for brevity's sake.

34 Two points. First, I stipulatively exclude from 'epistemic origin story' what we might call *constitutive explanations*. For example, <That a human is disposed to form true beliefs about medium-sized objects> *partly constitutes* or *is partly what it is for* that human's faculties to have their degree of reliability. I exclude such propositions from that person's epistemic origin story. Second, the explanation here is an answer to what Josh Schechter (2010) calls the *etiological question* (how we came to have reliable faculties) and not the *operational question* (how it is that the faculties are reliable). Thanks to Dan Korman for drawing my attention to Schechter's distinction.

35 When Plantinga (2000: 223) asks what is relevant to determining the probability of R, he says, "presumably the relevant facts would be facts about how these faculties originated; whether they were designed; if so, by whom and with what end in view …" But surprisingly, he does not develop this thought into a solution to the conditionalization problem. And Plantinga (2002: 224), as I quoted earlier, says he doesn't have even an unrigorous solution to it. So, in what follows, I develop and apply some of Plantinga's earlier thoughts to the conditionalization problem.

36 Thanks to Dan Korman for helpful discussion.

37 Consider Dan Baras's (2017: 200) example in which two coins, independently flipped, have matching sequences of heads and tails. Saying that they "landed in matching sequences merely by coincidence" is no explanation at all. Someone might object that the appeal to chance (or coincidence) *is* an explanation. In reply, I would just rule it out by stipulation, defining 'epistemic origin story of S's faculties' as a *non-coincidental* explanation of how it is that S's faculties came to have their degree of reliability. Here I was helped by Baras (2020: 1504 n. 7). Thanks to Bill Vincent for pushing me to address this sort of objection.

38 See Plantinga (2011b). For criticism of Plantinga's arguments, see Ye (2011) and Leahy (2013).

39 I take Stephen Law (2012) and Collin (2013) to be proposing such deflectors from philosophy of mind. See Calum Miller (2015) and note 46 in this chapter for responses to Law.

40 Collin (2013) takes himself to be doing this. Thanks to Erik Wielenberg for helpful discussion on whether these sorts of criticisms apply to premise (1) or premise (2) of EAAN.

41 Perhaps, as Plantinga (2000) argues, it is properly basic. See Moon (2016) for a survey of some recent literature on whether belief in God can be properly basic.

42 Due to the problem of evil, theistic belief alone might not be able to be a deflector. (See Moon [2017a], Mirza [2011: 83–84] and especially Roeber [2009].) However, note that Plantinga (2000) thinks that his *Christian* belief is also properly basic. I argue in Moon (2021) how Christian belief itself could be a deflector for many debunking arguments targeting it.

43 Peter van Inwagen (1993: 102) writes that "there could hardly be a more satisfying explanation for the existence of a thing than that its non-existence was impossible." This applies to God's reliable faculties. Thanks to Donald Smith for this reference.

to defeat, I would also have a defeater for thinking that I can *reason reliably*, which makes grim the prospects of using my belief in STC as a defeater-defeater in the ways just described in the main text. Thanks to Philip Swenson for helpful discussion.

to such a claim. See Brown (2018: Chapter 5) for extended criticism of such

44 Thanks to Philip Swenson, who first suggested to me that God's omniscience is explained by his being the greatest possible being. Thanks to him, Nevin Climenhaga, and Donald Smith for additional helpful discussion.

45 See also this excellent discussion of EAAN by Wielenberg (2002).

46 Stephen Law (2012) presents an argument against EAAN, which depends on the assumption that there are certain *conceptual connections* between belief and behavior. (For example, believing that water is south, conjoined with desiring water, has some conceptual connection to moving one's body south.) My Wilhelm case shows that no such conceptual connection exists. (Perhaps there *is* a conceptual connection between believing that water is south, desiring water, and *intending or trying* to move south, but this isn't a connection to bodily behavior but to another mental event.) For a more comprehensive attack on conceptual connections belief and behavior, see Smithies (2019: Chapter 4).

47 In an ordinary human, the intuitive evidence for both R and R* would likely overwhelm any potential undercutting-while-rebutting of R, leaving him no longer justified in believing DISHARMONY. This might be true, but it's beside the point. What matters is the sort of move Wilhelm may not make, described in the next paragraph, and how that is analogous to Wielenberg's reasoning. We can also just *stipulate* that Wilhelm's evidence for DISHARMONY is incredibly strong and strong enough to overwhelm R and R*.

48 Wielenberg uses another analogy where one has an independent *a priori* argument that Bic lighters are unreliable, but the "evidence provided by actually flicking the Bic 100 times swamps that independent argument" (Craig and Wielenberg 2021: 69). Suppose we take EAAN to be giving the naturalist an *undercutter-because-rebutter* for R. Then Wielenberg's analogy is apt. Furthermore, Craig's criticism of the analogy that "it doesn't include a defeater of one's perceptions that the lighter is aflame" (Craig and Wielenberg 2021: 75) fails. However, suppose we understand EAAN as giving an *undercutter-while-rebutter* for R. Then the undercutter-while-rebutter would also undercut O. Then, for the analogy, we *should* take that independent argument to also be giving an undercutter for the perceptual beliefs about the lighter. Now we can see that Wielenberg's analogy is not apt, and Craig's criticism of the analogy is vindicated.

49 Thanks to participants of the 2020 Princeton Project in Philosophy and Religion Inaugural conference, the 2021 Evolutionary Debunking Arguments workshop, and the 2021 summer philosophy workshop at Virginia Commonwealth University. Thanks also for helpful discussion to Michael Bergmann, Kenny Boyce, Kevin Brosnan, Nevin Climenhaga, Sarah Colquhoun, Jamie Fritz, Jordan Hampton, Perry Hendricks, Luke Kallberg, Dustin Locke, Eugene Mills, Hayoung Moon, Daniel Rubio, Donald Smith, Cathy Sutton, Philip Swenson, Jeff Tolly, Miles Tucker, Bill Vincent, and Dean Zimmerman. Thanks to Bill Craig, Dan Korman, Christos Kyriacou, and Diego Machuca for helpful written comments and extended discussion.

References

Alexander, David. 2017. "Unjustified Defeaters," *Erkenntnis* 82: 891–912.

Alston, William. 2002. "Plantinga, Naturalism, and Defeat." In Beilby 2002, 176–203.

Baras, Dan. 2017. "Our Reliability is in Principle Explainable," *Episteme* 14: 197–211.

Baras, Dan. 2020. "A Strike Against the Striking Principle," *Philosophical Studies* 177: 1501–1514.

Beilby, James (ed.). 2002. *Naturalism Defeated?* Ithaca: Cornell University Press.

Bergmann, Michael. 2002. "Commonsense Naturalism." In Beilby 2002, 61–90.

Bergmann, Michael. 2006. *Justification without Awareness*. New York: Oxford University Press.

Bergmann, Michael. 2021. *Radical Skepticism and Epistemic Intuition*. Oxford: Oxford University Press.

Brown, Jessica. 2018. *Fallibilism: Evidence and Knowledge*. Oxford: Oxford University Press.

Collin, James. 2013. "Semantic Inferentialism and the Evolutionary Argument against Naturalism," *Philosophy Compass* 8: 846–856.

Craig, William Lane and Erik Wielenberg. 2021. *A Debate on God and Morality*. Edited by A. L. Johnson. New York: Routledge.

Deem, Michael. 2018. "A Flaw in the Stich-Plantinga Challenge to Evolutionary Reliabilism," *Analysis* 78: 216–225.

DeRose, Keith. 2017. *The Appearance of Ignorance*. Oxford: Oxford University Press.

Fumerton, Richard. 1995. *Metaepistemology and Skepticism*. Lanham, MD: Rowman and Littlefield.

Ginet, Carl. 1995. "Comments on Plantinga's Two-Volume Work on Warrant," *Philosophy and Phenomenological Research* 55: 403–408.

Gordon, Bruce and William Dembski (ed.). 2011. *The Nature of Nature: Examining the Role of Naturalism in Science*. Wilmington: ISI Books.

Hendricks, Perry and Tina Anderson. 2020. "Does the Evolutionary Argument against Naturalism Defeat God's Beliefs?" *Sophia* 59: 489–499.

Korman, Dan. 2019. "Debunking Arguments," *Philosophy Compass* 14: 1–17.

Kyriacou, Christos. 2016. "Are Evolutionary Debunking Arguments Self-Debunking?" *Philosophia* 44: 1351–1366.

Law, Stephen. 2012. "Naturalism, Evolution, and True Belief," *Analysis* 72: 41–48.

Leahy, Brian. 2013. "Can Teleosemantics Deflect the EAAN?" *Philosophia* 41: 221–238.

Miller, Calum. 2015. "Response to Stephen Law on the Evolutionary Argument against Naturalism," *Philosophia*, 43, pp. 147–152.

Mirza, Omar. 2008. "A User's Guide to the Evolutionary Argument against Naturalism," *Philosophical Studies* 141: 125–146.

Mirza, Omar. 2011. "The Evolutionary Argument against Naturalism," *Philosophy Compass* 6: 78–89.

Moon, Andrew. 2016. "Recent Work in Reformed Epistemology," *Philosophy Compass* 11: 879–891.

Moon, Andrew. 2017a. "Plantinga's Religious Epistemology, Skeptical Theism, and Debunking Arguments," *Faith and Philosophy* 34: 449–470.

Moon, Andrew. 2017b. "Debunking Morality: Lessons from the EAAN Literature," *Pacific Philosophical Quarterly* 98: 208–226.

Moon, Andrew. 2021. "Circular and Question-Begging Responses to Religious Disagreement and Debunking Arguments," *Philosophical Studies* 178: 785–809.

15 Global Evolutionary Arguments

Self-Defeat, Circularity, and Skepticism about Reason

Diego E. Machuca

15.1 Introduction

Evolutionary arguments have primarily been used either to debunk or to vindicate our beliefs in specific domains. For example, some authors have employed them to debunk our beliefs in objective moral facts or properties.[1] Others have employed them to vindicate our commonsense beliefs[2] or our logical beliefs.[3] But evolutionary arguments could in principle also be used to debunk or to vindicate the belief in the reliability of our belief-forming processes or mechanisms as a whole and, hence, the epistemic credentials of our beliefs in general.[4]

In this chapter, written in an exploratory mode, I would like to focus on those wide-ranging evolutionary arguments. My interest in them lies in what they *might* tell us about the *possibly* aporetic nature of reason. Whereas the evolutionary debunking arguments (EDAs) that call into question the belief in the reliability of our cognitive faculties seem to fall prey to crippling self-defeat, the evolutionary vindicating arguments (EVAs) that seek to support that belief seem to fall prey to vicious circularity. If we take both kinds of arguments to consist of true or plausible premises and valid inferences at which we arrive through a meticulous use of reason, then their falling victim to either crippling self-defeat or vicious circularity *might* be regarded as a sign that, when we push rational reflection on the reliability, or lack thereof, of our belief-forming capacities to the limit, we end up in a situation of *aporia* from which there seems to be no escape.[5] I hasten to emphasize that I do not assert that human reason is of such a nature that it inexorably gives rise to *aporiai*. Rather, I limit myself to observing that this is a possibility that cannot be dismissed out of hand but only, if at all, after careful consideration. For this reason, the skepticism adopted in this chapter is of a Pyrrhonian stripe.[6]

The layout of the chapter is as follows. In Section 15.2, I focus on an EDA that targets the justification of the belief in the reliability of our belief-forming capacities and on the charge that such an argument is self-defeating. In so doing, I look at Alvin Plantinga's self-defeat argument

...tinga, Alvin. 1993. *Warrant and Proper Function.* New York: Oxford University Press.

Plantinga, Alvin. 1996. "Respondeo." In J. Kvanvig (ed.), *Warrant in Contemporary Epistemology,* 307–378. Lanham, MD: Rowman and Littlefield.

Plantinga, Alvin. 2000. *Warranted Christian Belief.* New York: Oxford University Press.

Plantinga, Alvin. 2002. "Reply to Beilby's Cohorts." In Beilby 2002, 204–275.

Plantinga, Alvin. 2011a. *Where the Conflict Really Lies: Science, Religion, and Naturalism.* Oxford: Oxford University Press.

Plantinga, Alvin. 2011b. "Content and Natural Selection," *Philosophy and Phenomenological Research* 83: 435–458.

Plantinga, Alvin. 2011c. "It's No Illusion!" In Gordon and Dembski, 166–172.

Pryor, James. 2004. "What's Wrong with Moore's Argument?" *Philosophical Issues* 14: 349–378.

Rea, Michael. 2002: *World without Design: The Ontological Consequences of Naturalism.* New York: Oxford University Press.

Roeber, Blake. 2009: "Does the Theist Have an Epistemic Advantage Over the Atheist?" *Journal of Philosophical Research* 34: 305–328.

Schechter, Joshua. 2010. "The Reliability Challenge and the Epistemology of Logic," *Philosophical Perspectives* 24: 437–464.

Smithies, Declan. 2019. *The Epistemic Role of Consciousness.* New York: Oxford University Press.

Talbott, William. 2002. "The Illusion of Defeat." In Beilby 2002, 153–164.

Talbott, William. 2011a. "More on the Illusion of Defeat." In Gordon and Dembski 2011, 152–165.

Talbott, William. 2011b. "The End of an Illusion?" In Gordon and Dembski 2011, 172–178.

van Inwagen, Peter. 1993. *Metaphysics.* Boulder, CO: Westview Press.

Vavova, Katia. 2014. "Debunking Evolutionary Debunking," *Oxford Studies in Metaethics* 9: 76–101.

Vavova, Katia. 2015. "Evolutionary Debunking of Moral Realism," *Philosophy Compass* 10: 104–116.

White, Roger. 2010. "You Just Believe That Because...," *Philosophical Perspectives* 24: 573–615.

Wielenberg, Erik. 2002. "How to Be an Alethically Rational Naturalist," *Synthese* 131: 81–98.

Williamson, Timothy. 2000. *Knowledge and Its Limits.* Oxford: Oxford University Press.

Wittwer, Silvan. 2020. "Evolutionary Debunking, Self-Defeat and All the Evidence." In M. Klenk (ed.), *Higher-Order Evidence and Moral Epistemology,* 31–53. New York: Routledge.

Woods, Jack. 2018. "Mathematics, Morality, and Self-Effacement," *Noûs* 52: 47–68.

Ye, Feng. 2011. "Naturalized Truth and Plantinga's Evolutionary Argument against Naturalism," *International Journal for Philosophy of Religion* 70: 27–46.

EVOLUTIONARY DEBUNKING ARGUMENTS

ETHICS, PHILOSOPHY OF RELIGION, PHILOSOPHY OF MATHEMATICS, METAPHYSICS, AND EPISTEMOLOGY

Edited by
Diego E. Machuca

ROUTLEDGE