Against Minimalist Responses to Moral Debunking Arguments
Daniel Z. Korman and Dustin Locke
Forthcoming in Oxford Studies in Metaethics
penultimate draft

1. Introduction

Evolutionary debunking arguments against moral realism aim to show that moral facts have no role to play in the evolutionary explanation of our moral beliefs, which in turn is meant to jeopardize the epistemic status of those beliefs, at least by realist lights.[1] Our aim here is to assess what has emerged as the dominant response to the debunking arguments. According to what we dub the 'minimalist response', one can resist the arguments without having to affirm that moral facts explain or are explained by moral beliefs.[2] Although the response can take a variety of different forms, the most common manifestation attempts to vindicate moral beliefs by showing how they could still (in a sense) track the truth even in the absence of the indicated explanatory connection.

When spelled out in detail, the minimalist response is at once vexing and enticing. It is vexing because it seems *too* easy. As we shall see, the minimalist's reasoning invariably rests on assumptions about which behaviors are in fact ethical—for instance that helping one's children is good—assumptions to which the respondent seems not to be entitled in the context of answering the debunker. But it is enticing insofar as it gets by with minimal metaphysical and explanatory commitments (hence the label 'minimalism'), promising to explain our moral reliability without having to abandon moral realism, reduce the moral to the natural, imbue irreducible moral facts with the power to influence our beliefs, or introduce unnecessary complexity by working moral facts into the evolutionary

---

[1] Such arguments have been advanced by Ruse (1985: ch. 6), Gibbard (2003: ch. 13), Lillehammer (2003), Kitcher (2005), Joyce (2006), Street (2006), and Braddock (2016).

[2] Minimalist responses—sometimes defended under the heading of "third-factors strategies" or "pre-established harmony"—have been advanced or defended by Nozick (1981: 342-348), Dworkin (1996: 117-126), Huemer (2005: 218-219), Schafer (2010, forthcoming), Enoch (2010: §§3-5, 2011: §7.4), White (2010: 588-589), Wielenberg (2010: §§4-8, 2014: ch. 4, 2016: §3), Brosnan (2011: 60-63), Parfit (2011: 532-533), Skarsaune (2011: §3), Berker (2014), Clarke-Doane (2015: §§4-6, 2016: §§2-4), Talbott (2015), Vavova (2015: §6), Baras (2017b), and Moon (2017).

explanations of our moral beliefs. If the response can make good on that promise, then we can evidently escape the debunking arguments entirely unscathed.

Our aim in this paper is two-fold. The first is to redirect the debate over minimalist responses, which has centered mainly on the charge that they are in one way or another "question-begging"—a charge that, we argue, is misguided. We think that the debate should instead be focused on a certain assumption about epistemic priority that lies at the heart of the minimalist response. Namely, that discovering that moral beliefs are not influenced by moral facts could undermine those beliefs only by way of demonstrating that they do not track the moral truth. Our second aim is to challenge this assumption by arguing that such explanatory information defeats *directly*: it is not in virtue of something else that such explanatory revelations undermine belief.

We conclude, not that the debunking argument succeeds, but rather that realists must disavow minimalism and embrace some account on which the moral facts explain our moral beliefs.

2. The Debunking Argument

To paraphrase an old Jewish joke: ask two philosophers, get three formulations of the evolutionary debunking argument. What follows may or may not be the best available to the debunker, but it will serve as an illuminating foil for the minimalist response.

Let's start by introducing two bits of terminology: *e-connected* ('e' for 'explanatory') and *m-connected* ('m' for 'modal'). One's moral beliefs are e-connected iff$_{def}$ moral facts either explain or are explained by one's moral beliefs. One's moral beliefs are m-connected iff$_{def}$ one's moral beliefs bear some epistemically significant modal relation to moral facts.[3] Putative examples of epistemically significant modal relations include safety, sensitivity, reliability, and non-accidental accuracy.

The debunking argument can then be rendered as follows:

(P1) Realists are rationally committed to believing that their moral beliefs are not e-connected.

---

[3] Cf. Enoch (2011: 174).

(P2) If one is rationally committed to believing that one's moral beliefs are not e-connected, then one is rationally committed to believing that one's moral beliefs are not m-connected.

(P3) If one is rationally committed to believing that one's moral beliefs are not m-connected, then one is rationally committed to withholding from moral beliefs.

(C) So, realists are rationally committed to withholding from moral beliefs.

P1 says that it is irrational, given the available evidence, to accept both realism and that one's moral beliefs are e-connected. The idea is that, because realists are committed to denying that moral beliefs explain the moral facts, they can affirm that moral beliefs are e-connected only by affirming that moral facts explain our moral beliefs. But (the idea goes) there is a broadly evolutionary explanation of our moral beliefs, and moral facts have no role to play within that evolutionary explanation. Strategies for resisting this premise include downplaying the role of evolutionary forces in shaping our current moral beliefs, reducing moral facts to the very natural facts cited in the debunker's evolutionary explanations, or finding a place for irreducible moral facts in the explanation of our moral beliefs (e.g., as proximate causes or as guiding the decisions of an intelligent designer).[4]

The idea behind P2 is that the absence of such explanatory connections ordinarily gives you strong reason to think that your beliefs aren't m-connected. Suppose that you've been using your magic 8-ball to find out who has a crush on you. Once you realize that it's just a toy, and that the facts about who does and doesn't have a crush on you are in no way influencing the beliefs you've formed about who has a crush on you, you should then believe that it is at best a "lucky coincidence" if those beliefs are correct. In other words, there is no epistemically significant modal connection between your crush beliefs and the crush facts. Mutatis mutandis, the idea goes, for your moral beliefs and the moral facts.

---

[4] See Street (2006: §§4-6) for an illuminating defense P1. See Huemer (2005: §8.6.3, 2016), Parfit (2011: 536), Fitzpatrick (2015: §3.2), and Isserow (forthcoming) on downplaying the influence of natural selection on our present moral beliefs; see Copp (2008) and Lott (2018) for reductionist responses; see Craig and Sinnott-Armstrong (2004: 20), Bogardus (2016), and Baras (2017a) on theistic responses; see Mogensen (2015) on invoking moral facts as proximate causes; and see Bengson (2015) for a rationalist approach.

Finally, the idea behind P3 is that you have no independent reason to think that you got lucky. Accordingly, once you acknowledge that you would have to have gotten lucky to end up with accurate moral beliefs, you shouldn't think that you did. So you should suspend your moral beliefs.

3. The Minimalist Gambit

A *minimalist* response to the debunking arguments is any response that involves granting P1 and denying one of the other premises. In practice, minimalists almost invariably resist the argument (so formulated) by denying P2. The exact form of the response will differ from one minimalist to the next, depending on which range of moral beliefs she is trying to vindicate and what sort of m-connection she is aiming to establish. But the core idea is always the same. The minimalist reasons from her antecedent moral beliefs to the conclusion that her moral beliefs are in one way or another m-connected, all the while granting that moral beliefs are not e-connected—that is, that they neither explain nor are explained by the moral facts. Call this *the minimalist gambit.*

(Some may worry that we have somehow "stacked the deck" against a certain strain of minimalism in how we have defined 'e-connection'. Many minimalists hold that the noncoincidental accuracy of our moral beliefs is secured by a more attenuated explanatory connection, a "third factor" that explains both why we have the moral beliefs that we do and why the moral facts are as they are. We vindicate our terminological choices below (§8) by arguing that postulating a third-factor does not help the minimalist escape our objections, which suggests that our taxonomy does "carve at the joints". In the meantime, take it as a matter of stipulation that these attenuated third-factor explanatory connections are not e-connections, and third-factor theorists do therefore count minimalists.)

To illustrate how the minimalist gambit is supposed to work, we'll focus on how the minimalist gambit is supposed to secure one particular m-connection for one particular moral belief; though structurally identical strategies can be (and have been) deployed to cover other m-connections and other sorts of moral beliefs. Let us see, then, how the minimalist gambit can be used to show that the moral belief that feeding one's own children

is good is a *safe* belief.[5] As we are understanding it, S's belief that p is *safe* iff S could not easily have been wrong about whether *p*. In other words, S is not wrong about whether p in any "nearby worlds".[6]

Step One: The minimalist maintains that it's no accident that feeding one's children is good. After all, the idea goes, survival is (at least pro tanto) good and, thus, actions that promote survival are themselves (at least pro tanto) good insofar as they promote survival; accordingly, since it's no accident that feeding one's children promotes their survival, it's no accident that feeding one's children is good. This is not to say that it is impossible for feeding one's children not to be good. There are a lot of worlds out there, including ones in which feeding your own children is a surefire way to kill them. But none of these are nearby worlds—that is to say, this is not something that could easily have happened. In all the nearby worlds, feeding one's children is good.

Step Two: The minimalist observes that it is no accident that we *believe* that feeding one's children is good. After all, aiming (as it does) at enhancing reproductive success, evolution is bound to favor beliefs that further that end by keeping our children alive. The belief that feeding one's children is good furthers this end because believing an action to be good motivates one to perform it, and performing this particular action helps keep one's children alive long enough to reproduce. Indeed, this belief enhances reproductive success not just in this world but also in the nearby worlds: feeding them couldn't easily have failed to keep them alive nor could believing that it's good to feed them easily have failed to motivate us to feed them. So it's no accident that we believe that feeding one's children is

---

[5] Our representative version of the gambit most closely resembles the one advanced by Nozick (1981: 346-348) and Enoch (2010, 2011), though neither focuses on safety explicitly; cf. Clarke-Doane (2015: 93, 2016: §2.3). See Huemer (2005: §8.6.4), Brosnan (2011), and Talbott (2015) for minimalist strategies that turn on the goodness and adaptive value of cooperation and which are poised to cover believed obligations to non-kin. See Wielenberg (2010, 2014: ch.4) on the reliability of beliefs about rights. See Skarsaune (2011) on the goodness of pleasure and Vavova (2015: §6) on the badness of pain. See Clarke-Doane (2015: §4, 2016: 26-27) on sensitivity and noncontingent moral truths. See Bedke (2014: §4) for an m-connection ("non-obliviousness") that arguably cannot be captured.

[6] The epistemic notion of 'safety' can has been defined in other ways (see Rabinowitz 2011). We have chosen a particularly strong notion of safety, so as not to rig things in our favor when we criticize the minimalist response below.

good. Again, this is not to say it's impossible for us to have believed otherwise, only that that couldn't easily have happened. In all nearby worlds, we believe that feeding one's children is good.

In all nearby worlds, feeding one's children is good. In all those nearby worlds, we believe that feeding one's children is good. Putting the pieces together (Step Three): we *correctly* believe, in all nearby worlds, that feeding one's children is good. In other words, the belief is safe. Crucially, the reasoning does not at any point invoke an e-connection between moral beliefs and moral facts. Let all moral facts be as abstract, inert, and mind-independent as you like. The reasoning still evidently goes through.[7]

4. Begging Questions and Default Entitlement

Cue the balking. Surely minimalists can't just take it for granted that feeding one's children is good! Why not? Because, the idea goes, relying on such moral beliefs *begs the question*. They're using their moral beliefs—the very beliefs that debunkers are calling into question—to vindicate the very faculties responsible for those beliefs. Minimalists have been quick to dismiss the charge of begging the question, and we think they are right to dismiss it.[8] So, before turning to our own objection to the minimalist strategy, let us briefly explain why the charge of question-begging misses the mark.

---

[7] It is worth mentioning a variation on the minimalist gambit, which involves denying P3. This one is decidedly less attractive, but what it lacks in plausibility, it makes up for in sheer chutzpah. Concede that it would take a massive stroke of luck to wind up with accurate moral beliefs. But then consult your moral intuitions, check whether it is good to feed one's children, and find that it is. And, introspecting, check whether you believe that it's good to feed one's children, and find that you do. Putting these together, conclude that you correctly believe that it is good to feed one's children. Repeat the process for other moral beliefs, and conclude that you have a great many accurate moral beliefs and that you must therefore have gotten miraculously lucky. Cf. Dworkin (1996: 125-127), White (2010: 589), Setiya (2012: ch. 2.2), Locke (2014b), Vavova (2014: 80-82), and McBee (2018) for discussion.

[8] For charges of question-begging (and the like), see Fraser (2014: 471), Vavova (2014: 81), Street (2008: §6, 2011: §6), Shafer-Landau (2012: §6), Crow (2016), Joyce (2016: 157-158), and Lott (2018: §2.2). For responses, see Schafer (2010: 475-476 and 487-488, forthcoming), White (2010: 588-592), Wielenberg (2010: 447, 2016: 506), Brosnan (2011: 62), Enoch (2011: 117-121), Bedke (2014: §3.3), Setiya (2012: ch. 2.3), Berker (2014: §8), Locke (2014a: §5), Clarke-Doane (2015: 89, 2016: 31), Korman (2015: §7.6.1, forthcoming: §5), Baras (2017b: 209-2010), and Moon (2017).

We take it for granted that we are justified in believing a great many things. We do not expect any disagreement with the debunker on this point: the debunking argument is not meant to be an argument for generic skepticism; it is meant to target our moral beliefs specifically.[9] Can the debunker make good on the (somewhat amorphous) charge of question-begging without relying on some generic skeptical principle? We don't think they can. To see why, let's consider three precisifications of the charge.

On a first precisification, one begs the question when one treats as evidence the very claims that one's interlocutor is calling into question. The minimalist does do that. She takes for granted, in the face of the debunking arguments, that feeding one's children is good. But there is good reason to reject the sort of "evidential neutrality" principles that would prohibit relying on contested beliefs.[10] For such principles lead straight to sweeping skeptical results. All it would take is one encounter with a global skeptic or an unruly philosophy major to render all your beliefs unjustified.

On a second precisification, one begs the question when one relies on the deliverances of some source of information without independent evidence of the source's reliability. Minimalists do that as well: their appeals to the deliverances of their moral faculties isn't accompanied by any independent evidence of the reliability of those faculties—independent, that is, of the deliverances of those very faculties. But one had better not insist that independent evidence of the reliability of a source is always needed before relying on that source, on pain of a fairly obvious regress. We need to allow that, in some cases, we enjoy a default (albeit defeasible) entitlement to believe the deliverances of a source, even absent independent evidence of its reliability.[11] Plausible candidates for such sources are perception, introspection, testimony, and memory, and we see no good reason—nor have debunkers provided good reason—to exclude moral sources from the list.

---

[9] Cf. Vavova (2015: 105).

[10] See, e.g., Nozick (1981: 197-198), Pryor (2004: §7), Williamson (2007: §7.3), and Kelly (2008: 73-76).

[11] See Burge (1993), Wright (2004), Field (2005), and White (2006) on the general need for something like default entitlement. For minimalist appeals to default entitlement, see Schafer (2010: 476, forthcoming) and Wielenberg (2016: 506).

On a third precisification, one begs the question when one engages in a certain kind of circular reasoning: reasoning from some beliefs to the conclusion that those beliefs have some desirable epistemic feature. The minimalist does that too, insofar as she relies on her moral beliefs in the course of establishing their m-connectedness. But, as many have observed, prohibiting this sort of epistemic circularity across the board leads to sweeping skeptical results.[12] Moreover, once we recognize the default entitlement to our moral beliefs, it is hard to see what could be illicit about reasoning from those beliefs to any conclusion one likes, so long as they entail or otherwise support that conclusion.

All of us, minimalists included, enjoy a default entitlement to certain of our beliefs, even absent a non-question-begging defense of those beliefs. However, default entitlement does not amount to *indefeasible* entitlement, and our objection to the minimalist gambit in what follows is that their explanatory concession undermines any default entitlement they have to rely on moral beliefs. So the charge of question-begging is on to something; the minimalists' reliance on their moral beliefs is indeed illicit. But it is illicit, not because they fail to meet some dubious requirement of evidence neutrality or independent evidence or noncircularity, but because they have a defeater for those beliefs.

5. Defeat and Epistemic Priority

When one withholds belief about whether a certain range of beliefs are e-connected—either believing that they aren't or at least suspending belief about whether they are—let's call this an *explanatory concession*.[13] Explanatory concessions typically serve as defeaters for the associated beliefs. Recall the magic 8-ball (from §2). The realization that your beliefs about who has a crush on you aren't explained by the facts

---

[12] See Alston (1986), Van Cleve (2003), Bergmann (2004), Vogel (2008), Titelbaum (2010), and Alexander (2011) in defense of circular reasoning. It has also been observed that debunkers themselves need to engage in some such circular reasoning if they are to prevent their debunking arguments from overgeneralizing to our perceptual beliefs; see Sosa (2002: 375), Gibbard (2003: ch. 13), Schafer (2010: 475-476), and Bedke (2014: 107-108).

[13] Above, we told you what it is for *moral* beliefs to be e-connected but we haven't yet told you what it is in general for a belief to be e-connected. Roughly, a belief is e-connected iff it explains or is explained by the sorts of facts it purports to be about. More on this in §8.

about who has a crush on you, and that believing it's so doesn't make it so, defeats those crush beliefs.

Why is it that such explanatory concessions defeat, when they do? Put another way: is there something in virtue of which they serve as defeaters, and if so what is it? This is a question of epistemic priority. And it is this question that will take us to (what should be) the heart of the debate between the minimalist and her opponents. To see why we think that the debate turns on this question, let us consider two possible answers: a minimalist-friendly answer and a minimalist-unfriendly answer.

The minimalist-friendly answer to the priority question is that explanatory concessions have epistemic import only to the extent that the absence of the relevant e-connection indicates the absence of one or another m-connection, and that it is the latter absence that ultimately does the defeating. Something like this is implicit in the usual presentations of the debunking arguments, in which subversive explanations are portrayed as doing their debunking work by way of revealing moral beliefs to be unsafe, or insensitive, or unreliable, or at best accidentally or coincidentally accurate. And if it's true that explanatory concessions have merely derivative epistemic import, defeating only by way of revealing the absence of some m-connection, then this is a boon to minimalists. For as we saw, by relying on their antecedent moral beliefs—to which they are defeasibly entitled—minimalists are able to assure themselves that the lack of e-connection did not prevent them from forming m-connected beliefs.

The minimalist-*un*friendly answer is that explanatory concessions have nonderivative epistemic import. They undermine beliefs directly, and it is not in virtue of revealing the beliefs to be unsafe or unreliable or in some other way deficient that the concessions undermine those beliefs. If that's right, then the minimalist gambit is a non-starter. For in that case, the minimalist's vindication of her beliefs proceeds from moral beliefs for which she already possesses a defeater, namely, her explanatory concession.[14]

Thus, the fate of the minimalist strategy turns on a delicate question of epistemic priority. If explanatory concessions defeat (when they do) only in virtue of what they rationally commit us to saying about m-connections, then the minimalist gambit is in good

---

[14] See Moon (2017) for more on how already having a defeater can preclude "defeater-deflection".

epistemic standing. But if explanatory concessions defeat directly, then the minimalist gambit can't get off the ground.

6. Sensitivity First

The minimalist gambit, as we just saw, presupposes that explanatory concessions defeat only in virtue of rationally committing one to the absence of some m-connection. Which m-connection? While we illustrated the minimalist gambit with *safety* above, there are other m-connections one might think are relevant here. We'll return to safety below (§7), but let's first consider an alternative answer: sensitivity.

S's belief that p is sensitive iff: were it not the case that p, S would not have believed that p. Suppose that minimalists identify sensitivity as the operative m-connection. That is, suppose that they embrace Sensitivity First:

> *Sensitivity First*
> Believing that one's belief that p is not e-connected rationally commits one to withholding on one's belief that p only by virtue of rationally committing one to believing that one's belief that p is insensitive.[15]

Choosing sensitivity as the operative m-connection would require minimalists to revise the vindicatory story from §3 somewhat, but we see no obstacle to doing so. If feeding one's children weren't good, this would be for some (relatively) mundane reason, for instance because people are prone to dangerously overfeeding their own children, making feeding something best left to nannies or medical professionals. If that were the case, we would no doubt be well aware that feeding one's own children isn't good. So the belief is sensitive. This line of reasoning does implicitly draw on antecedent moral beliefs.[16] But this is just business as usual for the minimalist.

---

[15] Perhaps minimalists will wish amend the principle to say: "…only by virtue of rationally committing one to *withholding belief* that one's belief that p is sensitive." This amendment won't affect what we go on to say below. Mutatis mutandis for Safety First (below).
[16] For instance, one must implicitly assume that we in fact has robust obligations to our children in order to rule out the *actual* world from being the closest world in which it's not

We'll raise two problems for Sensitivity First. The first involves counterfactuals with necessarily false antecedents, otherwise known as "counterpossibles". It was, for a time, widely believed that counterpossibles were all vacuously true. After all, the idea went, a counterfactual A→B is true so long as all the nearest possible A-worlds are B-worlds. If A is necessarily false, then there are no possible A-worlds and thus none that fail to be B-worlds. Never mind whether this is the right view of counterpossibles (it's not).[17] Certainly its proponents were not *irrational* in accepting it. They had their reasons.

Suppose that Lois, one such rational advocate of vacuous counterpossibles, finds herself having a powerful intuition that Goldbach's conjecture is true. She believes that it's true, and believes moreover that Goldbach's conjecture (like other mathematical truths) is necessarily true. She then remembers that she was recently hypnotized and—after watching the video—realizes that the intuition is the result of a post-hypnotic suggestion. Further, she is convinced that the hypnotist gave her this intuition for reasons having nothing to do with whether Goldbach's conjecture is in fact true. She sees in the video that he flipped a coin to decide whether to give her a pro-Goldbach or an anti-Goldbach intuition.

The reasonable thing for Lois to do at this point is to suspend belief about Goldbach's conjecture; the explanatory concession defeats the belief. But suppose that she attempts to assure herself that the belief is nevertheless in good standing, by noting its sensitivity. "Goldbach's conjecture *is* true"—she says, helping herself to her intuitive belief—"and so it's vacuously true that if it were false I wouldn't have believed it." Starting from a belief to which she is default entitled, she reasons her way to the sensitivity of the belief. Clearly, though, it is not rational for her to stand by her Goldbach belief, and her reasoning does nothing to improve her epistemic situation.

Why is it not rational for her to stand by her belief? The natural answer is that she has a defeater, and that defeater is the explanatory concession: she knows that she believes that Goldbach's conjecture is true for reasons having nothing to do with how mathematical reality in fact is. If that's right, then it must be that explanatory concessions have epistemic

---

good to feed one's children. Cf. Clarke-Doane (2015: §4, 2016: §2) and Korman (forthcoming: §6), both channeling Sturgeon (1988: §3).
[17] See Nolan (2013: §2.2) for discussion.

import that's independent of what they reveal about sensitivity. Lois is a counterexample to Sensitivity First.

Our second complaint draws on familiar counterexamples to sensitivity constraints on knowledge. Here is a representative example, from Jonathan Vogel (2012: 130-131). Suppose you know in some completely ordinary way that Omar has new shoes. You can also thereby know—via a trivial inference—*that you aren't mistaken in believing that Omar has new shoes*. But this further belief is not sensitive. Had you mistakenly believed that Omar has new shoes, you'd still think you weren't mistaken in believing he has new shoes. In other words, this belief amounts to knowledge despite not being sensitive. (It's worth noting that this case was designed to work—and does work—even against more sophisticated versions of sensitivity that hold fixed the particular belief-forming method the believer uses.)

Even acknowledged insensitivity does not defeat. It's not irrational for you to stand by your belief that you're not mistaken about Omar's new shoes—you double checked!—even while appreciating the arcane philosophical point that you would have believed this even if (for some bizarre reason) it were false. But if that's right, it's just not plausible that explanatory concessions defeat by way of showing the associated beliefs to be insensitive. Recognized insensitivity isn't, in itself, a threat to beliefs.

To be sure, this isn't a counterexample to Sensitivity First. Minimalists might grant the point, conceding that recognized insensitivity doesn't *always* defeat, but insist that it does defeat when accompanied by explanatory concessions. But what could account for why it defeats only in these cases? Presumably what accounts for it is that it's the explanatory concessions that are ultimately doing the defeating, *pace* Sensitivity First.

The foregoing also has implications for disjunctive accounts of how explanatory concessions defeat (when they do). For instance, Justin Clarke-Doane (2015: §6, 2016: §4) advances the following account, which he dubs *Modal Security*.

Modal Security
Information, E, cannot undermine our D-beliefs without giving us some reason to believe that our D-beliefs are not both safe and sensitive.

Rendered as a claim about epistemic priority, the idea would be:

> *Security First*
> Believing that one's belief that p is not e-connected rationally commits one to withholding on one's belief that p only by virtue of rationally committing one to believing *either* that one's belief that p is unsafe *or* that one's belief that p is insensitive.

What we have seen is that the safety disjunct would have to be doing all the heavy lifting. As we saw in the Lois case, reasoning one's way to sensitivity doesn't preclude explanatory concessions from defeating. And as we saw in the Omar case, getting reasons (indeed: conclusive reason) to think that one's beliefs are insensitive doesn't by itself undermine them. What this means is that, if Security First is true, then it must be because Safety First is true:

> *Safety First*
> Believing that one's belief that p is not e-connected rationally commits one to withholding on one's belief that p only by virtue of rationally committing one to believing that one's belief that p is unsafe.

So let's turn now to Safety First.

7. Safety First

Safety First doesn't fall victim to the Lois counterexample. Recall that S's belief that p is *safe* iff S could not easily have been wrong about whether p. Lois may be able to reason her way to the sensitivity of her Goldbach belief but not to its safety. She could easily have been wrong about whether Goldbach's conjecture is true, because the hypnotist's coin could easily have landed tails, in which case he would have hypnotized her to believe that the Goldbach's conjecture is false. And, while your belief about Omar's new shoes is insensitive, it isn't unsafe: it couldn't easily have happened that you

mistakenly believed that he has new shoes. So Safety First doesn't face the same problems as Sensitivity First. It faces different problems.

Here is a counterexample to Safety First. Jack sees a streak in a cloud chamber and believes that the streak was caused by a proton. But Jack has not received the training of an ordinary physics student. Rather, has believes it because some Martians—after convincing him of their superior intellect—told him that protons cause those kinds of streaks. Moreover, they decide to tell him this, not because they themselves had done any physics, but simply because they liked the sound of the English word 'proton'. You may even suppose, if you like, that there is some deep law of Martian psychology that makes them like the sound of the word 'proton', and so it could not easily have happened that the Martians told Jack that such streaks were caused by something else. Finally, let us suppose that after forming the belief that protons cause those streaks, Jack learns all these details about the origins of his belief, and concedes that his belief that the streaks are caused by protons is not explained by the facts about what causes them.[18]

The reasonable thing for Jack to do at this point is to abandon his belief that such streaks are caused by protons. But suppose Jack retains the belief and attempts to vindicate it with the following line of reasoning:

> Yes, my belief that such streaks are caused by protons is not explained by the facts about what causes those beliefs. Still, given what I have just learned about Martian psychology, I could not easily have formed a different belief about whether such streaks are caused by protons. Moreover, such streaks *are* caused by protons, and—since this interaction is surely underwritten by natural laws—it could not easily have failed to be the case that they are caused by protons. So the belief is safe: it's true in all nearby worlds and I believe it in all nearby worlds.

Starting from a belief to which he is default entitled—the testimonial belief that such streaks are caused by protons—he reasons his way to the safety of the belief.[19] Clearly,

---

[18] This case is drawn from Locke (2014a). Clarke-Doane (2016: 32-33) addresses a similar case—framed as a putative counterexample to Modal Security—but it is unclear that his treatment of that case can be adapted to handle the cases we present here.
[19] See Burge (1993: 485) on default entitlement to testimonial beliefs. If you doubt that testimony is a source of default entitlement, you may instead suppose that Jack was

though, it is irrational for him to stand by his proton belief, and the reasoning does nothing to improve his epistemic situation.

Why is it irrational? The natural answer is that he has a defeater for the belief, and that defeater is his explanatory concession: he accepts that he believes such streaks are caused by protons for reasons having nothing to do with whether they're in fact caused by protons. If that's right, then it must be that explanatory concessions have epistemic import that's independent of what they reveal about safety. Jack is a counterexample to Safety First.

Here is a second counterexample. On the basis of clear and distinct intuitions, Neora believes in an all-powerful deity. Later, Agent Smith convinces her that she is part of a computer simulation. He tells her that the designers had a terrible time getting consciousness and cognition to arise in the simulation, but—through endless trial and error—found that they could achieve this result only by programming the inhabitants to be strongly disposed to believe in an all-powerful deity. Without such beliefs, the simulations would break down before they even got going. Neora believes everything he tells her. And she believes that the deity (if it does exist) had nothing to do with her religious intuitions and associated beliefs. Despite having now accepted all this, she doesn't abandon her belief in an all-powerful deity.

Surely it is irrational for Neora to retain this belief. Why? Because she has conceded that it is not e-connected. But her belief is safe by her lights. How so? Her intuitions are very plausibly a source of default entitlement. (Those who take moral faculties to be a source of default entitlement, at any rate, are hardly in a position to deny that religious intuitions enjoy the same status.) Starting from her intuitive belief that there is an all-powerful deity, she concludes that this belief couldn't easily have been false: after all, nothing could easily have stood in the way of such a being's existing. Nor could she easily have believed otherwise, she reasons, since, as she learned from Smith, the only nearby worlds in which a simulation inhabited by her even exists are ones in which she has these religious beliefs. By her lights, then, she believes it in all nearby worlds and it's true in all

---

brainwashed to have an intuition or disposition to believe that protons cause those streaks, or to have experiences that (richly) present those streaks as caused by protons.

nearby worlds, so she isn't mistaken in any nearby worlds: the belief is safe. And yet it's defeated. Safety First is false.

Note that our claim is not that Jack's or Neora's vindicatory reasoning is exactly analogous to the minimalists' vindicatory reasoning. Rather, the claim is that these are cases in which explanatory concessions defeat without revealing the defeated beliefs to be unsafe by the believers' lights—which makes them counterexamples to Safety First.

One might worry that our counterexamples fail if we slightly alter our formulation of safety. S's belief that p is safe, we said, iff S couldn't easily have been mistaken about p. But consider *safety**, where S's belief that p is safe* iff S couldn't easily have failed to have a correct belief about p. Even by their own lights, Jack's proton belief and Neora's deity belief aren't safe* (exercise for the reader). So, perhaps minimalists can affirm Safety* First, which claims that explanatory concessions defeat by virtue of revealing beliefs to lack safety*.

But it's hardly plausible that recognizing beliefs to be unsafe* undermines those beliefs. Suppose I pull a reputable encyclopedia off the shelf, flip to a random page, and read that Hume was born in 1711.[20] The belief is unsafe*—I could easily have flipped to a different page and would have had no beliefs one way or the other about when Hume was born. But realizing this does not in any way jeopardize my newly formed Hume belief. It just isn't plausible that explanatory concessions derive their explanatory import from a lack of safety*.[21]

8. Never Mind the Gap

Minimalists might take the lesson of the Lois and Jack and Neora cases to be that Safety First or Sensitivity First just need to be tweaked in some other way, or that some other m-connection should be wheeled in to bridge the gap between explanatory concessions and defeat. But there is a more natural lesson to draw from these cases, namely

---

[20] Adapting an example from White (2010: 597).
[21] Alternatively, one might prefer a formulation of safety that is in one way or other relativized to methods. We leave it as a challenge to Safety Firsters to find some such relativization that does not fall victim to variants of our objections or to variants of known counterexamples to method-based safety constraints (see, e.g., Bogardus 2014 and Zhao forthcoming).

that explanatory concessions defeat directly, and not by way of indicating the absence of some m-connection, and that this is why Lois and Jack and Neora's reasoning is illicit. There is no "gap" to be filled in explaining how explanatory concessions generate defeat. In that case, the minimalist gambit is a non-starter, for (as we saw in §5) the gambit crucially involves reasoning from antecedent moral beliefs that minimalists themselves concede are not e-connected.

Why think that there *must* be some gap to be filled, something in virtue of which explanatory concessions defeat when they do? To our knowledge, the only minimalist attempt to answer this question is due to Clarke-Doane (2015: 96-97). In defending his modal security principle (see §6 above), Clarke-Doane insists that once you have assured yourself that certain of your beliefs are safe and sensitive, then they are, by your lights, *bound to be true*. To think that some information, explanatory or otherwise, can undermine some beliefs without in any way challenging your conviction that they are bound to be true is, he says, "dubiously coherent" (2015: 96). In his words: "How could information obligate us to give up our beliefs of a kind while failing to threaten our judgment that they were (all but) bound to be true?" (2015: 97). So, the idea goes, if some explanatory information does undermine our beliefs, it must be by way of giving us reason to think that the beliefs are unsafe or insensitive.

We think the reasoning goes wrong at the very last step. True, explanatory information cannot rationally commit us to abandoning beliefs without *threatening* our judgment that those beliefs are bound to be true. But there is more than one way to threaten a judgment. One is by rebutting it, that is, by giving us reason to think that it is false. The other is by undercutting it, that is, by giving us reason to think that our reasons fail to support it. The envisaged explanatory concessions do undercut the minimalist's stated reasons for thinking that our moral beliefs are safe, sensitive, or otherwise bound to be true, insofar as they undercut the antecedent moral beliefs used (in the minimalist gambit) in reasoning one's way to an m-connection.

Clarke-Doane is right that explanatory revelations fail to give us positive reason to think that the moral beliefs aren't bound to be true. Rather, they leaves us in the dark entirely, depriving us of any way of telling whether they are bound to be true. And that is

more than enough to "threaten" the judgment that they're bound to be true, which opens the door to the explanatory concession obligating us to give up the beliefs.[22]

Alternatively (but relatedly), minimalists might insist that explanatory concessions cannot defeat directly because, once you assure yourself that your beliefs are tracking the truth, it would be absurd to demand more than that. This does have some prima facie plausibility, which we suspect arises from an equivocation between a correlative and an explanatory reading of "tracking".[23] On the correlative reading, beliefs track some facts so long as there is a counterfactual covariance between them. On the explanatory reading, beliefs track some facts when there is a covariation underwritten by an e-connection. We agree that it would be overly demanding to ask for more than *explanatory* tracking. But assuring yourself of mere correlative tracking, we maintain, is not enough.

Another possible strand of resistance comes from the concern that the sort of explanatory constraint on rational belief that we are proposing will be plagued by all the same problems as the long-ago-abandoned causal theory of knowledge. For instance, such theories and constraints threaten to cut against inferential beliefs. You observe the fire in the fireplace and are justified in believing that there is smoke coming out the chimney. Of course, the fact that smoke is coming out of the chimney does not explain (causally or otherwise) the belief that it is. But this realization surely does not undermine the belief.[24]

This indeed is a counterexample to flat-footed formulations of the explanatory constraint, like the following:

(EC) If S believes that her belief that p neither explains nor is explained by the fact that p, then S is thereby rationally committed to withholding belief that p.

But it is not a problem for a more refined explanatory constraint, like the one we have been working with throughout. We said that a moral belief is e-connected if and only if it explains or is explained by *some* fact in the domain of moral facts, and we've argued that believing that your moral beliefs are not e-connected defeats those beliefs. Extrapolating from this, we get something like the following, more permissive explanatory constraint:

---

[22] See Woods (2018: §3.2) for further discussion of Clarke-Doane's modal security principle.

[23] Cf. Bedke (2014: 105).

[24] The example is due to Goldman (1967: 365-366).

(EC*) If p is about domain D, and S believes that her belief that p is neither explained by nor explains some D-facts, then S is thereby rationally committed to withholding belief that p.

Your belief that smoke is coming from the chimney is explained by facts about smoke and chimneys: it is the result of an inductive inference from past observations of smoke coming from chimneys with active fireplaces. So EC* doesn't prescribe withholding belief that smoke is now coming out of the chimney.

What EC* does rule out, however, is the minimalist's vindicatory reasoning. The minimalist concedes that her moral beliefs—her beliefs about what's right and wrong—neither explain nor are explained by facts about what's right or wrong. So EC* will entail that the minimalist is rationally required to withhold on her moral beliefs.

Admittedly, EC* faces problems of its own, and a full defense lies outside the scope of the paper. Even absent such a defense, we think that what the Jack and Neora cases clearly suggest is that there must be some such explanatory constraint on rational belief. The only question is what exactly it is.[25]

Minimalists might grant the need for *some* explanatory constraint on rational belief, but suggest an alternative, minimalist-friendly replacement for EC. For instance:

(EC**) If S believes that (i) her belief that p neither explains nor is explained by the fact that p and (ii) there is no single fact that (at least partially) explains both her belief that p and the fact that p, then S is thereby rationally committed to withholding belief that p.

This will handle the smoke case, since the fire in the fireplace explains both the fact that there is smoke coming out of the chimney and your belief that there is. But it *won't*

---

[25] One immediate worry for EC* is that it faces a sort of "generality problem" in specifying the relevant domains. No plausible constraint should prescribe withholding belief that the sun will rise tomorrow. But if that belief counts as being "about" the domain of facts about the future, then EC* does have this unwanted implication. So the proponent of EC* must supply some account of which domains are relevant to assessing whether a belief satisfies EC*. Thanks to David Killoren and several others for pressing us on this point. Your two authors are of two minds about how troubling this generality problem is. See Locke's Cognition Defeat (2014b: 232), McCain's EF (2014: §4.4 and 6.4), Lutz's EAD (2017: §2), Schechter (2018: §3), and Korman's EC5 (forthcoming: §8) for alternative approaches to formulating explanatory constraints, also tailor-made to handle inductive beliefs (including beliefs about the future).

prescribe withholding belief that feeding one's children is good. For the fact that feeding one's children promotes their survival serves as a "third factor", both explaining why we believe that it's good (together with the fact that survival-promoting moral beliefs are adaptive) and why it is good (together with the fact that survival is good).

The problem with EC** is that it cannot account for trivial variations of the cases from §7. For instance, suppose that Jack learns (in some entirely normal way) that a proton has just been fired through a cloud chamber. Drawing on his Martian belief that protons cause streaks in cloud chambers, and without looking at the cloud chamber, he concludes that there is a streak in the chamber. Indeed there is a streak in the chamber. Moreover, the fact that there is a streak and his belief that there is a streak have a common cause—the fact that a proton was fired through the chamber—and, relying on his antecedent proton beliefs, Jack can reason his way to that third-factor explanation. Yet, just as in the original Jack case, Jack clearly ought to withhold belief about whether there is a streak in the chamber, knowing what he does about the origins of his belief that protons cause streaks. EC** therefore fails to deliver the correct verdict, that Jack is required to withhold belief. (You had *one job*, explanatory constraint!)

Here is the more general lesson to draw about such third-factor explanations. Their apparent value—and, indeed, their advertised value in the work of Enoch and others—lies in their ability to explain the noncoincidental accuracy of moral beliefs. In other words, the value lies in their ability to secure an m-connection between facts and associated beliefs. But they are of little to no value in the present context, where we are trying to account for cases in which reasoning one's way to an m-connection fails to vindicate the relevant beliefs.

9. Upshots

We have argued that explanatory concessions do not derive their epistemic import from what they reveal about safety, sensitivity, or other such m-connections. If we're right about that, then it is neither here nor there that minimalists can reason from their antecedent moral beliefs to the m-connectedness of those beliefs, for those beliefs have already been undermined by minimalists' explanatory concessions. Let us close by drawing out four further upshots of the foregoing.

First, the failure of minimalist strategies shows us what it *would* take to resist the evolutionary debunking argument: one must reject P1 and affirm some sort of e-connection between moral beliefs and moral facts. Reductive views on which the moral facts just are the very natural facts that ultimately explain our moral beliefs are still in the running. So are theistic views on which the moral facts influence our moral beliefs by way of making themselves known to an intelligent designer who ensures that evolutionary processes yield reliable moral faculties. So are rationalist views on which moral facts influence our moral beliefs via some sort of quasi-perceptual apprehension. These all have problems of their own, to be sure. But they are the sorts of responses that aren't ruled out by what we have shown.

Second, the formulation of the debunking arguments with which we began is at best misleading, insofar as it suggests that explanatory concessions undermine one's beliefs by way of revealing something about m-connections. Here is superior formulation, relying directly on an explanatory constraint on rational belief:

(P1) Realists are rationally committed to believing that their moral beliefs are not e-connected.

(P2*) If p is about domain D, and S believes that her belief that p is neither explained by nor explains some D-facts, then S is thereby rationally committed to withholding belief that p.

(C) So, realists are rationally committed to withholding from moral beliefs.

Third, the objections we have raised against minimalist responses to moral debunking arguments seem to apply equally to minimalist responses to reliability challenges that arise in other literatures, for instance Mark Balaguer's (1995) response to the Benacerraf challenge, Ernest Sosa's (2002) and Joel Pust's (2004) accounts of modal reliability, and Michael Bergmann's (2002) response to Plantinga's evolutionary argument against naturalism.[26] In each of these cases, the strategy is to grant that our beliefs are not e-connected, but then reason from the disconnected beliefs to some account of our reliability. But, if our argument above is successful, then these minimalist strategies fail

---

[26] See also Korman (2014: §§3-4, 2015: §7.3, forthcoming: §7.2) against minimalist responses to debunking arguments against perceptual beliefs about ordinary objects.

for the same reasons as moral minimalism. The only viable responses to these challenges are "explanationist" responses on which our beliefs bear an appropriate explanatory connection to their subject matter.

Finally, and more speculatively, we think that the considerations raised above can be adapted to show that minimalists have things exactly backwards. Safety and sensitivity are epistemic virtues because, and to the extent that, they are indicative of e-connectedness; becoming convinced that some beliefs are not sensitive, safe, or in some other way m-connected undermines those beliefs (when it does) by virtue of indicating the absence of an e-connection. In other words, we should accept Explanation First:

> *Explanation First*
> Believing that one's belief that p is not m-connected rationally commits one to withholding on one's belief that p only by virtue of rationally committing one to believing that one's belief that p is not e-connected.

Explanation First draws support from cases of sensitivity without relief (e.g. Lois)—where reasoning one's way to the sensitivity of one's beliefs does not seem to vindicate those beliefs—as well as cases of safety without relief (e.g. Jack and Neora) and cases where the absence of insensitivity or safety seems to be no cause for concern (e.g. Omar's shoes and Bogardus's atomic clock).[27] The lack of relief can be explained by the known absence of an e-connection, and the lack of concern can be explained by the known presence of an e-connection. This suggests that, at bottom, it is e-connections doing the heavy epistemic lifting.[28]

References

Alexander, David J. (2011), 'In Defense of Epistemic Circularity', *Acta Analytica* 26: 223-241.

Alston, William P. (1986), 'Epistemic Circularity', *Philosophy and Phenomenological Research* 47: 1-30.

Balaguer, Mark (1995), 'A Platonist Epistemology', *Synthese* 103: 303-325.

Baras, Dan (2017a), 'A Reliability Challenge to Theistic Platonism', *Analysis* 77: 479-487.

Baras, Dan (2017b), 'Our Reliability is in Principle Explainable', *Episteme* 14: 197-211.

Bedke, Matthew (2014), 'No Coincidence?', *Oxford Studies in Metaethics* 9: 102-125.

Bengson, John (2015), 'Grasping the Third Realm', *Oxford Studies in Epistemology* 5: 1-38.

Bergmann, Michael (2002), 'Commonsense Naturalism', in James Beilby (ed.), *Naturalism Defeated?* (Ithaca: Cornell University Press), pp. 61-90.

Bergmann, Michael (2004), 'Epistemic Circularity: Malignant and Benign', *Philosophy and Phenomenological Research* 69: 709-727.

Berker, Selim (2014), 'Does Evolutionary Psychology Show That Normativity is Mind-Dependent', in Justin D'Arms and Daniel Jacobson (eds.) *Moral Psychology and Human Agency* (Oxford: Oxford University Press), pp. 215-252.

Bogardus, Tomas (2014), 'Knowledge Under Threat', *Philosophy and Phenomenological Research*, 88: 289-313.

Bogardus, Tomas (2016), 'Only All Naturalists Should Worry About Only One Evolutionary Debunking Argument', *Ethics* 126: 636-661.

Braddock, Matthew (2016), 'Evolutionary Debunking: Can Moral Realists Explain the Reliability of our Moral Judgments', *Philosophical Psychology* 29: 844-857.

Brosnan, Kevin (2011), 'Do the Evolutionary Origins of Our Moral Beliefs Undermine Moral Knowledge', *Biology and Philosophy* 26: 51-64.

Burge, Tyler (1993), 'Content Preservation', *The Philosophical Review* 102: 457-488.

Clarke-Doane, Justin (2015), 'Justification and Explanation in Mathematics and Morality', *Oxford Studies in Metaethics* 10: 80-103.

Clarke-Doane, Justin (2016), 'Debunking and Dispensability', in Neil Sinclair and Uri Leibowitz (eds.), *Explanation in Ethics and Mathematics* (Oxford: Oxford University Press), pp. 23-36.

Copp, David (2008), 'Darwinian Skepticism about Moral Realism', *Philosophical Issues* 18: 186-206.

Craig, William L. and Sinnott-Armstrong, Walter (2004), *God?: A Debate Between a Christian and an Atheist* (Oxford: Oxford University Press).

Crow, Daniel (2016), 'Causal Impotence and Evolutionary Influence', *Ethical Theory and Moral Practice* 19: 379-395.

Dworkin, Ronald (1996), 'Objectivity and Truth: You'd Better Believe It', *Philosophy and Public Affairs* 25: 87-139.

Enoch, David (2010), 'The Epistemological Challenge to Metanormative Realism', *Philosophical Studies* 148: 413-438.

Enoch, David (2011), *Taking Morality Seriously* (Oxford: Oxford University Press).

Field, Hartry (2005), 'Recent Debates about the A Priori', *Oxford Studies in Epistemology* 1: 69-88.

Fraser, Benjamin (2014), 'Evolutionary Debunking Arguments and the Reliability of Moral Cognition', *Philosophical Studies* 168: 457-473.

Gibbard, Allan (2003), *Thinking How to Live* (Cambridge: Harvard University Press).

Goldman, Alvin I. (1967), 'A Causal Theory of Knowing', *The Journal of Philosophy* 64: 357-372.

Huemer, Michael (2005), *Ethical Intuitionism* (New York: Palgrave).

Huemer, Michael (2016), 'A Liberal Realist Answer to Debunking Skeptics', *Philosophical Studies* 173: 1983-2010.

Isserow, Jessica (forthcoming), 'Evolutionary Hypotheses and Moral Skepticism', *Erkenntnis*.

Joyce, Richard (2006), *The Evolution of Morality* (Cambridge: MIT Press).

Joyce, Richard (2016), *Essays in Moral Skepticism* (Oxford: Oxford University Press).

Kelly, Thomas (2008), 'Common Sense as Evidence: Against Revisionary Ontology and Skepticism', *Midwest Studies in Philosophy* 32: 53-78.

Kitcher, Philip (2005), 'Biology and Ethics', in David Copp (ed.), *The Oxford Handbook of* Lewis, David (1973), *Counterfactuals* (Malden: Blackwell).

Korman, Daniel Z. (2014), 'Debunking Perceptual Beliefs about Ordinary Objects', *Philosophers' Imprint* 14: 1-21.

Korman, Daniel Z. (2015), *Objects: Nothing out of the Ordinary* (Oxford: Oxford University Press).

Korman, Daniel Z. (forthcoming), 'Debunking Arguments in Metaphysics and Metaethics', in Brian McLaughlin and Alvin I. Goldman (eds.), *Cognitive Science and Metaphysics*.

Lillehammer, Hallvard (2003), 'Debunking Morality: Evolutionary Naturalism and Moral Error Theory', *Biology and Philosophy* 18: 567-581.

Locke, Dustin (2014a), 'Darwinian Normative Skepticism', in Michael Bergmann and Patrick Kain (eds.), *Challenges to Moral and Religious Belief* (Oxford: Oxford University Press), pp. 220-236.

Locke, Dustin (2014b), Review of *Knowing Right from Wrong*, *Ethics* 124: 649-656.

Lott, Michael (2018), 'Must Realists Be Skeptics? An Aristotelian Reply to a Darwinian Dilemma', *Philosophical Studies* 175: 71-96.

McBee, Joshua D. (2018), 'Circularity in Setiya's *Knowing Right from Wrong*', *Social Theory and Practice* 44: 349-375.

McCain, Kevin (2014), *Evidentialism and Epistemic Justification* (New York: Routledge).

Mogensen, Andreas L. (2015), 'Evolutionary Debunking Arguments and the Proximate/Ultimate Distinction', *Analysis* 75: 196-203.

Moon, Andrew (2017), 'Debunking Morality: Lessons from the EAAN Literature', *Pacific Philosophical Quarterly* 98: 208-226.

Nolan, Daniel (2013), 'Impossible Worlds', *Philosophy Compass* 8: 360-372.

Nozick, Robert (1981), *Philosophical Explanations* (Cambridge: Harvard University Press).

Parfit, Derek (2011), 'On What Matters: Volume Two' (Oxford: Oxford University Press).

Pryor, James (2004), 'What's Wrong With Moore's Argument', *Philosophical Perspectives* 14: 349-378.

Pust, Joel (2004), 'On Explaining Knowledge of Necessity', *Dialectica* 58: 71-87.

Rabinowitz, Dani (2011), 'The Safety Condition for Knowledge', *The Internet Encyclopedia of Philosophy*.

Ruse, Michael (1986), *Taking Darwin Seriously* (Prometheus).

Schafer, Karl (2010), 'Evolution and Normative Scepticism', *Australasian Journal of Philosophy* 88: 471-488.

Schafer, Karl (forthcoming), 'Evolutionary Debunking Arguments, Explanatory Structure, and Anti-Realism', in *Taking Sentimentalism Seriously* (Cambridge University Press).

Schechter, Joshua (2018), 'Explanatory Challenges in Metaethics', in Tristram McPherson and David Plunkett (eds.), *Routledge Handbook of Metaethics* (Routledge), pp. 443-459.

Setiya, Kieran (2012), *Knowing Right from Wrong* (Oxford: Oxford University Press).

Shafer-Landau, Russ (2012), 'Evolutionary Debunking, Moral Realism and Moral Knowledge', *Journal of Ethics and Social Philosophy* 7: 1-37.

Skarsaune, Knut (2011), 'Darwin and Moral Realism', *Philosophical Studies* 152: 229-243.

Sosa, Ernest (2002), 'Reliability and the A Priori', in Tamar S. Gendler and John Hawthorne (eds.), *Conceivability and Possibility* (New York: Oxford University Press), pp. 369-384.

Street, Sharon (2006), 'A Darwinian Dilemma for Realist Theories of Value', *Philosophical Studies* 127: 109-166.

Street, Sharon (2008), 'Reply to Copp: Naturalism, Normativity, and the Varieties of Realism Worth Worrying About', *Philosophical Issues* 18: 207-228.

Street, Sharon (2011), 'Mind-Independence Without the Mystery: Why Quasi-Realists Can't Have it Both Ways', *Oxford Studies in Metaethics* 6: 1-32.

Sturgeon, Nicholas (1988), 'Moral Explanations', in Geoffrey Sayre-McCord (ed.), *Essays on Moral Realism* (Ithaca: Cornell University Press), pp. 229-255.

Talbott, William J. (2015), 'How Could a 'Blind' Evolutionary Process Have Made Human Moral Beliefs Sensitive to Strongly Universal, Objective Moral Standards', *Biology and Philosophy* 30: 691-708.

Titelbaum, Michael G. (2010), 'Tell Me You Love Me: Bootstrapping, Externalism, and No-Lose Epistemology', *Philosophical Studies* 149: 119-134.

Van Cleve, James (2003), 'Is Knowledge Easy – Or Impossible? Externalism as the Only Alternative to Skepticism', in Stephen Luper (ed.), *The Skeptics* (Burlington: Ashgate), pp. 45-59.

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.

Vogel, Jonathan (2008), 'Epistemic Bootstrapping', *Journal of Philosophy* 105: 518-539.

Vogel, Jonathan (2012), 'The Enduring Trouble with Tracking', in Kelly Becker and Tim Black (eds.) *The Sensitivity Principle in Epistemology*, (Cambridge: Cambridge University Press), pp. 122-151.

White, Roger (2006), 'Problems for Dogmatism', *Philosophical Studies* 131: 525-557.

White, Roger (2010), 'You Just Believe That Because', *Philosophical Perspectives* 24: 573-615.

Wielenberg, Erik J. (2010), 'On the Evolutionary Debunking of Morality', *Ethics* 120: 441-464.

Wielenberg, Erik J. (2014), *Robust Ethics* (Oxford: Oxford University Press).

Wielenberg, Erik J. (2016), 'Ethics and Evolutionary Theory', *Analysis* 76: 502-515.

Williamson, Timothy (2007), *The Philosophy of Philosophy*, (Malden: Blackwell).

Woods, Jack (2018), 'Mathematics, Morality, and Self-Effacement', *Noûs* 52: 47-68.

Wright, Crispin (2004), 'Warrant For Nothing (and Foundations for Free)?', *Aristotelian Society* 78: 167-212.

Zhao, Haicheng (forthcoming), 'Knowledge without Safety', *Synthese*.