

# Against Minimalist Responses to Moral Debunking Arguments

Daniel Z. Korman and Dustin Locke

Draft: 6/4/18

## 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.<sup>1</sup> 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 argument without having to affirm that moral facts explain or are explained by moral beliefs.<sup>2</sup> 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.

The first appearance of the minimalist response (to our knowledge) is in Robert Nozick’s *Philosophical Explanations*:

[E]thical behavior will serve inclusive fitness through serving or not harming others, through helping one’s children and relatives, through acts that aid them in escaping predators, and so forth; that this behavior is helpful and not harmful is not unconnected to why (on most theorist’s views) it is ethical. The ethical behavior will increase inclusive fitness *through the very aspects that make it ethical*, not as a side effect through features that only accidentally are connected with ethicality. (1981: 346, our emphasis)

Roughly, the idea is that, since evolutionary forces are bound to favor ethical beliefs that promote helping our children (e.g., the belief that helping them is good), and since such

---

<sup>1</sup> Such arguments have been advanced by Ruse (1985: ch. 6), Gibbard (2003: ch. 13), Lillehammer (2003), Kitcher (2005), Joyce (2006), and Street (2006).

<sup>2</sup> 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 (2009), 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), Moon (forthcoming), and Schechter (forthcoming: §6).

beliefs tend to be true, we are therefore bound to end up with by and large accurate ethical beliefs. Nozick goes on to emphasize that this account of our reliability on such matters is entirely compatible with granting that “the explanation of how ethical actions, beliefs, and principles originate and endure will turn out not to refer to the truth of the principles believed” (1981: 348). The same basic strategy has appeared in numerous forms in recent years, though typically without awareness of its Nozickian roots.

When spelled out in detail, the minimalist response is at once vexing and enticing. It is vexing because it seems *too easy*. The minimalist’s reasoning rests on assumptions about which behaviors are ethical—for instance that it’s true 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 irreducibly moral facts with the power to influence our beliefs, or introduce unnecessary complexity by working moral facts into the evolutionary 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 the response is 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 the absence of an explanatory connection could undermine moral beliefs only by way of demonstrating that the beliefs 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 explanatory revelations undermine belief. Indeed, as we will argue, minimalists have it exactly backwards. It’s only by way of revealing a lack of explanatory connection that the absence of tracking undermines beliefs.

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.<sup>3</sup> 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* (short for 'explanatorily connected') and *m-connected* (short for 'modally connected'). For any *p* that is about a domain *D*, we'll say that

The belief that *p* is *e-connected* iff it is explained by or explains some facts in *D*,

and that

The belief that *p* is *m-connected* iff it bears an epistemically significant counterfactual relation to some facts in *D*.<sup>4</sup>

The debunking argument can then be rendered as follows:

- (P1) Realists are rationally committed to believing that their moral beliefs are not *e-connected*.<sup>5</sup>
- (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.

---

<sup>3</sup> See Wielenberg (2016) on the variation in presentations of the argument.

<sup>4</sup> Cf. Enoch (2011: 174). Putative examples of epistemically significant counterfactual relations include safety and sensitivity (more on these below).

<sup>5</sup> That is to say, it is irrational, given the available evidence, to accept both realism and that one's moral beliefs are *e-connected*.

The idea behind P1 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 the 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).<sup>6</sup>

The idea behind P2 is that the absence of an e-connection ordinarily gives you strong reason to think that your beliefs aren't m-connected. When you discover that your magic 8-ball is just a toy, and that the facts about who does and doesn't have a crush on you are in no way influencing its verdicts about who has a crush on you, you should then believe that it is *at best* a "lucky coincidence" if its verdicts are correct. In other words, there is no epistemically significant counterfactual connection between its verdicts and the truth. *Mutatis mutandis* for the verdicts of your moral faculties.

The idea behind P3 is that you have no independent reason to think that you got lucky, so, 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.

Again, this is just one of many possible formulations of the debunking argument, and may not be everyone's preferred formulation. But we think that the minimalist is profitably understood as reacting to this formulation of the challenge, and there is every reason to expect that the minimalist response—if it works at all—can be adapted to handle alternative formulations of the challenge.

---

<sup>6</sup> See Street (2006) for an illuminating defense P1. See Huemer (2005: §8.6.3, 2016), Parfit (2011: 536), and Fitzpatrick (2015: §3.2) on downplaying the influence of natural selection on our present moral beliefs; see Copp (2008) and Lott (forthcoming) for reductionist responses; see Craig and Sinnott-Armstrong (2004: 20) and Bogardus (2016) for theistic responses; and see Mogensen (2015) on invoking moral facts as proximate causes.

### 3. The Minimalist Gambit

Minimalists resist the argument (so formulated) by denying P2.<sup>7</sup> 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 to the moral facts, all the while granting that moral beliefs are not e-connected to the moral facts. Call this *the minimalist gambit*.

For purposes of illustration, we'll focus on one particular moral belief—namely, the belief that feeding one's own children is good—and one way of being m-connected—namely, being safe. (S's belief that *p* is *safe* iff S could not easily have been wrong about whether *p*.) Seeing how the minimalist can secure the safety of this belief, and how that in turn is meant to undermine P2, will serve as a nice illustration of the minimalist gambit.<sup>8</sup>

To see how our target belief could be safe even without being e-connected, let's suppose that the moral facts are more or less as we take them to be. In particular, suppose that survival is (at least pro tanto) good and that actions that promote survival are themselves (at least pro tanto) good insofar as they promote survival. In that case, 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, these are not things that could easily have happened. In all the nearby worlds, feeding one's children is good.

Next, notice 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

---

<sup>7</sup> Though see note 13 for an alternative minimalist strategy that involves denying P3.

<sup>8</sup> This version of the strategy most closely resembles the one offered by Nozick (1981: 346-348) and Enoch (2010, 2011), though neither focuses on safety explicitly. Cf. Clarke-Doane (2015: 93, 2016: §2.3).

perform it, and performing this particular action helps keeps one's children alive. 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 good. Again, this is not to say it's impossible for us to have believed otherwise, only that that couldn't easily have happened.<sup>9</sup>

Putting the pieces together: On the assumption that moral reality is as we take it to be, feeding one's children is good in all nearby worlds. We also believe in all those nearby worlds that feeding one's children is good. So, if moral reality is as we take it to be, we *correctly* believe that feeding one's children is good in all nearby worlds—in other words, the belief is safe. Crucially, the reasoning does not at any point depend on the belief in question being e-connected. Let all ethical facts—including the fact that feeding one's children is good—be as abstract, inert, and mind-independent as you like. The reasoning still goes through. On the assumption that moral reality is as we take to be, the belief is safe even if it's not e-connected.

What we have just seen is that one moral belief could have one sort of m-connection without being e-connected. We have no doubt that structurally similar strategies can be deployed to cover other m-connections and other sorts of moral beliefs.<sup>10</sup> At any rate, our objections won't turn on the scope of the minimalist response. So let's simply grant the minimalist that something like the above line of reasoning can be adapted to show that every one of our moral beliefs can be m-connected in any way you like without having to be e-connected to the moral facts.

---

<sup>9</sup> Perhaps we could have easily failed to have any belief at all on the matter, but safety (as formulated above) requires only that one couldn't easily have been wrong, not that one couldn't easily have failed to be right.

<sup>10</sup> 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) for a minimalist strategy that turns on the goodness and adaptive value of finding survival-promoting things pleasurable. See Clarke-Doane (2015: §4, 2016: 26-27) on sensitivity and noncontingent moral truths. Though see Bedke (2014: §4) for an m-connection (“non-obliviousness”) that arguably cannot be captured, and see Braddock (2016: §6) for an argument that the minimalist explanation of our moral beliefs is inadequate.

Even so, the minimalist's work is not yet done. To see this, let's take an analogy. The menu says that today's special is minestrone. The menu's saying so obviously doesn't *entail* that today's special is minestrone. Still, the menu's saying so is good evidence that today's special is minestrone. So unless you have some countervailing evidence that either overrides or undercuts your evidence that today's special is minestrone, you have no grounds for denying that it's today's special, and no grounds for denying the material conditional: *if the menu says that minestrone is today's special, then minestrone is today's special*. Evidence that chicken soup is today's special or that a prankster tampered with the menus would fit the bill; but some such evidence is needed.

Similarly, our representative minimalists have conceded that the belief that feeding one's children is good is not e-connected. They rightly observe that the lack of e-connection doesn't *entail* a lack of safety; as we saw, there is at least one epistemically possible way for moral reality to be—namely, the way we take it to actually be—on which the belief turns out to be safe. Still, the absence of e-connection is good evidence of a lack of safety (remember the magic 8-ball from §2).<sup>11</sup> So, unless they have countervailing evidence that either overrides or undercuts the evidential import of the lack of e-connection, minimalists have no grounds for denying the conditional premise P2:

(P2) If one believes 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.

What evidence might they provide that, in this particular case, the beliefs are safe despite the lack of e-connection?

Not all minimalists have appreciated the need to answer this further question.<sup>12</sup> But others have, and the answer is always more or less the same: affirm (rather than merely

---

<sup>11</sup> To our knowledge, no minimalist has denied that the lack of e-connection is at least *prima facie* or *pro tanto* evidence of a lack of m-connection.

<sup>12</sup> Nor are all minimalists entirely clear about whether they are trying to answer this question or only demonstrating the mere possibility of m-connectedness without e-connectedness. Those seeking to establish only the weaker claim (e.g., Brosnan 2011) may be conceiving of the debunking arguments as resting on the needlessly strong premise that it is impossible for beliefs to be m-connected without being e-connected.

suppose for the sake of argument) the truth of the moral beliefs featuring in their vindicatory reasoning. Cast in terms of our representative example it goes like this:

Feeding one's children *is* good. And (for the reasons given above) we couldn't easily have failed to believe that it's good. So, we couldn't have been mistaken about whether feeding one's children is good. The belief is safe.

In other words, instead of holding merely that moral beliefs are safe *if we suppose* that moral reality is as we take it to be, one concludes that moral beliefs are safe because moral reality *is* as we take it to be. In this way, minimalists cinch the m-connectedness of moral beliefs, allowing them to deny P2.<sup>13</sup>

#### **4. Begging Questions and Default Entitlement**

Cue the balking. Surely, the minimalist 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 that are responsible for those beliefs.<sup>14</sup> Minimalists have been quick to dismiss the charge of begging the question, and we think they are right to dismiss it.<sup>15</sup> So, before turning to our own objection to the minimalist strategy, let us briefly explain why the charge of question-begging misses the mark.

---

<sup>13</sup> We have been focusing on P2, though in the interest of seeing the full range of possibilities, it is worth mentioning an alternative minimalist strategy that involves denying P3. This one is decidedly less attractive, but what it lacks in plausibility, it makes up for in sheerchutzpah. 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) and White (2010: 589).

<sup>14</sup> 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 (forthcoming: §2.2).

<sup>15</sup> For responses to charges of question-begging (and the like), see Schafer (2010: 475-476 and 487-488), White (2010: 588-592), Wielenberg (2010: 447, 2016: 506), Brosnan (2011: 62), Enoch (2011: 117-121), Bedke (2012: §3.3), Berker (2014: §8), Locke (2014:



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 specifically target our moral beliefs.<sup>16</sup> The question is thus whether the debunker can 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, as numerous epistemologists have observed, there is good reason to reject the sort of "evidential neutrality" principles that would prohibit relying on contested beliefs.<sup>17</sup> Among other things, accepting any such principle leads 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 reliance on the deliverances of their moral faculties isn't accompanied by independent evidence of their reliability—independent, that is, of the deliverances of those very faculties. But, as (again) is widely recognized by epistemologists, 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.<sup>18</sup> Plausible candidates for such sources are perception, introspection, testimony, and

---

§5), Clarke-Doane (2015: 89, 2016: 31), Korman (forthcoming: §5), and Moon (forthcoming).

<sup>16</sup> Cf. Vavova (2015: 105).

<sup>17</sup> See, e.g., Nozick (1981: 197-198), Pryor (2004: §7), Williamson (2007: §7.3), and Kelly (2008: 73-76).

<sup>18</sup> 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) and Wielenberg (2016: 506).

memory, and we see no good reason—nor have debunkers provided good reason—to exclude moral sources from the list.

On a third precisification, one begs the question by engaging in a 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 (yet again) is widely recognized by epistemologists, prohibiting this sort of epistemic circularity across the board leads to sweeping skeptical results. Moreover, once we recognize the default entitlement to our moral beliefs, it is hard to see what could be illicit about reasoning from those to any conclusion one likes, so long as it is entailed or otherwise supported by them.<sup>19</sup>

We must recognize that all of us, minimalists included, enjoy a default entitlement to certain of our beliefs, even when we have no independent evidence of their accuracy, even when one is in the middle of assessing a debunking argument, and even when face to face with the debunker. 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 their default entitlement to rely on moral beliefs. So the charge of question-begging is on to something; the minimalist's 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*. Typically, explanatory concessions serve as defeaters for the associated beliefs. Acknowledging, for instance, that a magic 8-

---

<sup>19</sup> 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 will 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).

ball's verdicts about your crushes are not explained by the "crush facts" defeats any beliefs based on its verdicts. 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?

One possible answer 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. This suggests that the defeating power of an explanatory concession derives from an associated missing m-connection: believing that one's belief that p is not e-connected rationally commits one to withholding on one's belief that p (when it does) only by virtue of rationally committing one to believing that one's belief that p is not m-connected.

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 their beliefs are indeed m-connected, in which case the lack of e-connection is no cause for concern.

Suppose, on the other hand, that explanatory concessions have nonderivative epistemic import. Suppose, in particular, that explanatory concessions undermine beliefs directly: it is not in virtue of revealing some 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.<sup>20</sup>

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

---

<sup>20</sup> See Moon (forthcoming) for more on how already having a defeater can preclude "defeater-deflection".

rationally commit us to saying about m-connections, then the minimalist gambit is in good epistemic standing. But if explanatory concessions defeat directly, then the minimalist gambit is deeply problematic.

This concludes the first part of the paper, the aim of which was to redirect the debate over minimalism, setting aside the charge that it's question-begging and focusing instead on the question of whether e-connections have to derive their epistemic import from their relation to m-connections. We turn now to our second aim: defending the nonderivative import of explanatory concessions and thereby undermining the minimalist gambit.

## 6. Sensitivity First

The minimalist gambit, as we have argued, requires insisting that explanatory concessions defeat only in virtue of rationally committing one to accept the absence of some m-connection. This isn't yet to say *which* m-connection it is whose absence ultimately does the defeating. We'll consider two possibilities: sensitivity and safety.

S's belief that p is sensitive iff: were it not the case that p, S would not have believed that p.<sup>21</sup> 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.<sup>22</sup>

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

---

<sup>21</sup> The formulation really ought to include a restriction to methods, to handle well-known counterexamples (see Nozick 1981: 179-185). We'll ignore this complication; nothing in what follows hangs on this.

<sup>22</sup> 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 (§7).

because people are prone to dangerously overfeeding their own children, making feeding something best left to nannies or medical professionals. But 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. The reasoning implicitly draws on antecedent moral beliefs, but this is just business as usual for the minimalist.<sup>23</sup>

We'll raise two problems for Sensitivity First. Both are adaptations of well-known problems with sensitivity constraints on knowledge (that S knows that p only if S's belief that p is sensitive).

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  $A \rightarrow B$  is true so long as all the nearest possible A-worlds are B-worlds; and 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).<sup>24</sup> 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 hypnotist gave her this intuition for reasons having nothing to do with whether Goldbach's conjecture is in fact true. (Let's suppose she sees him flip 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

---

<sup>23</sup> For instance, one implicitly assumes that one in fact has robust obligations to one's children in order to rule out the *actual* world from being the closest world in which it's not good to feed one's children. Cf. Clarke-Doane (2015: §4, 2016: §2), channeling Sturgeon (1988: §3).

<sup>24</sup> See Nolan (2013: §2.2) for discussion.

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

Our second complaint against Sensitivity First also draws on familiar counterexamples to sensitivity conditions on knowledge, but these counterexamples have nothing to do with counterpossibles.. Here is a representative example, drawn from Vogel (2012: 130-131). Suppose you know in some completely ordinary way that Omar has new shoes. You can also thereby know—from the most trivial inference—*that you aren’t mistaken in believing that Omar has new shoes*. But this belief is not sensitive. Had you mistakenly believed that Omar has new shoes, you’d still think you weren’t mistaken in believing he had new shoes. In other words, this belief amounts to knowledge despite not being sensitive.

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 (in some bizarre way) 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 *at least* when accompanied by explanatory concessions. But what could account for why it defeats only in these cases? Presumably what account for it is that it’s the explanatory concessions that are ultimately doing the defeating, *pace* Sensitivity First. (Which incidentally is exactly what we think is happening. **More on this in §9.**)

The foregoing also raises some concerns about 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 had 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.<sup>25</sup> But Jack has not received the training of an ordinary physics student. Rather, Martians tell Jack that those kinds of streaks are caused by protons. 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 the Martians could not easily have told Jack that such streaks were caused by something else. Finally, let us suppose that after forming the belief that a proton has gone by, 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.

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 e-connected. 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.<sup>26</sup> Clearly, though, it is irrational for him to stand by his proton belief, and the reasoning does nothing to improve his epistemic situation.

---

<sup>25</sup> This case is drawn from Locke (2014). Clarke-Doane (2016: 32-33) addresses a similar case—framed as a counterexample to Modal Security—but it is unclear that his treatment of that case can be adapted to handle the cases we present here.

<sup>26</sup> See Burge (1993: 485) on default entitlement to testimonial beliefs.



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.<sup>27</sup>

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 building a simulation inhabited by conscious cognizers but that—through a great deal of trial and error—they found that they could achieve this result only by rendering the inhabitants 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.<sup>28</sup> Despite believing all this, she doesn't abandon her belief in an all-powerful deity.

Surely it is irrational for her to retain this belief. Why? Because she has conceded that it is not e-connected. But her belief is safe by her lights. Her intuitions are very plausibly a source of default entitlement. (And minimalists, who take *moral* intuitions to be a source of default entitlement, are hardly in any 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

---

<sup>27</sup> To be clear, this is not to say that Jack's belief in the safety of his proton beliefs is rational. It's not. And there's a natural explanation of why it's not: because the reasoning behind his safety belief relied on beliefs that had already been defeated by his explanatory concession. More on this in §9 below.

<sup>28</sup> In particular, she doesn't think that the deity is ultimately responsible for the existence of the simulation or for the designers' intentions. Contrast this with the theistic response to the moral debunking argument, where the theist insists that the deity, guided by the moral facts, is ultimately responsible for the evolutionary process that produced our moral beliefs. Whatever objections this response might face, it is not an instance of the minimalist gambit, and so does not face the particular objection we raise here.

otherwise, she reasons, since (as she learned from Smith) the only nearby worlds in which the simulation is functional 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 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 vindictory reasoning is exactly analogous to the minimalists' vindictory reasoning. It isn't. Rather, the point is that these are cases in which explanatory concessions defeat without revealing the defeated beliefs to be unsafe by the believer's lights—which makes them counterexamples to Safety First.

One might worry that our counterexamples fail if we slightly tweak 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 his own lights, Jack's belief that protons cause the streaks isn't safe\*. After all, the 'proton'-loving Martians might easily have told him instead that protons cause the expansion of the universe, in which case he wouldn't have had any belief about the streaks (let alone a correct belief). Likewise, even by her own lights, Neora's belief in an all-powerful deity isn't safe\*. After all, the scientists could easily have chosen not to implant the religious intuitions, in which case she wouldn't have had any beliefs about the deity. She wouldn't have had any beliefs. She wouldn't have existed at all. 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.<sup>29</sup> 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. As with sensitivity, it just isn't plausible that explanatory concessions derive their explanatory import from a lack of safety\*.

---

<sup>29</sup> Adapting an example from White (2010: 597).

## 8. Never Mind the Gap

Minimalists might think that the lesson of the Lois and Jack and Neora cases is that Safety First or Sensitivity First need some tweaking, or that they should be invoking some other m-connection to bridge the gap between explanatory concessions and defeat. But there is a more natural lesson to draw from these cases.

The more natural lesson is that explanatory concessions defeat directly, and not by way of indicating the absence of some m-connection. There is no “gap” to be filled in explaining how explanatory concessions generate defeat. This yields a straightforward explanation of what’s illicit about Lois’s and Jack’s and Neora’s reasoning: their reasoning proceeds from beliefs that have already been defeated by their explanatory concessions. Furthermore, it would undermine the minimalist gambit, which (as we saw in §5) crucially involves reasoning from antecedent moral beliefs that minimalists themselves concede are not e-connected to the moral facts. More tendentiously: the minimalist gambit is manifestly problematic, and the thesis that explanatory concessions defeat directly yields a better diagnosis of the problem than the usual complaint that the gambit is question-begging.

Why think that there *must* be some gap to be filled, something in virtue of which explanatory concessions defeat when they do? The closest we have found to an attempt to answer this question is due to Clarke-Doane (2015: 96-97). In defending his modal security principle (mentioned in §6 above), Clarke-Doane insists that once you have assured yourself that certain of your beliefs are m-connected (e.g. safe and sensitive), then they are, by your lights, *bound to be true*: you couldn’t easily have been mistaken about them and you wouldn’t have believed them if they hadn’t been the case. To think that some information (explanatory or otherwise) can undermine these beliefs without in any way challenging your belief 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 virtue of giving us reason to think that the beliefs are unsafe or insensitive.<sup>30</sup>

---

<sup>30</sup> Relatedly, one might insist that, 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 an explanatory and a

We think the reasoning goes wrong at the very last step. Indeed, 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 second is by undercutting it, that is, by giving us reason to think that our reasons fails 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 none of this gives us (positive) reason to think that the moral beliefs aren't bound to be true. Rather, it leaves us in the dark entirely, depriving us of any way of telling whether they are or aren't 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.

Perhaps another strand of resistance comes from the concern that any kind of explanatory constraint on rational belief will be plagued by all the same problems as the long-ago-abandoned causal theory of knowledge. For instance, such theories and constraints rule out inductive knowledge. 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.<sup>31</sup>

This indeed is a counterexample to flat-footed proposals for explanatory constraints, 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.

---

correlative reading of "tracking". (Cf. Bedke 2014: 105.) On the former, beliefs track some facts just in case there is an explanatory connection between them. On the latter, they track them just in case there is a counterfactual covariance between them. 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.

<sup>31</sup> The example is due to Goldman (1967: 365-366).

But it is no problem for the sort of explanatory constraint we have been working with here. For a belief to count as e-connected in our sense, it suffices that it be explained by *some* facts in the domain that it is about. This gives us:

(EC\*) If p is about domain D, and S believes that her belief that p is neither explained by nor explains some D-facts [i.e., is not e-connected], 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 observations of smoke coming from chimneys. So EC\* doesn't rule out the belief that smoke is 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—are not explained by facts about what is right or wrong. So EC\* will still entail that the minimalist is rationally required to withhold on her moral beliefs.<sup>32</sup>

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

(EC\*\*) If S believes that (i) her belief that p is not e-connected 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 on her 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* prescribe withholding belief that feeding one's children is good. Here, there is a "third factor" that is poised to explain both the belief that feeding them is good and the fact that it is. The fact that feeding one's children promotes their survival both explains 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).

---

<sup>32</sup> EC\* admittedly still needs some work, e.g., to account for beliefs about logical truths (e.g. that all aliens are aliens). One way forward would be to amend it to say: If p is about domain D, and (i) S believes that her belief that p neither explains nor is explained by D-facts, *and* p is not a logical consequence of beliefs meeting condition (i), then...

The problem with EC\*\* is that it simply isn't strong enough to account for the sorts of cases presented above (e.g., it cannot explain why Jack's belief is defeated by his explanatory concession). Indeed, it's so weak that it hardly rules out any beliefs at all. Virtually all worldly matters and beliefs about them will be at least partly explained by the Big Bang.<sup>33</sup>

## 9. Explanation First?

This completes our defense of the claim that explanatory concessions have epistemic import that does not derive from the absence of 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 concession.

We now wish to defend, somewhat more tentatively, the stronger conclusion that minimalists have things precisely backwards. M-connectedness derives its epistemic import from what it indicates about e-connectedness. 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. Our case for this rests on four families of examples: sensitivity without relief, safety without relief, insensitivity without concern, and unsafeness without concern.

First, there are cases of *sensitivity without relief*, in which finding that your belief (by your own lights) is sensitive intuitively does nothing to convince you that it is in good standing. Typically, seeing that you wouldn't have believed something if it were false is reassuring. But no sensible person in Lois's shoes would feel relieved upon reaching the conclusion that they wouldn't have believed Goldbach's conjecture if it weren't true—but only for the trivial reason that (if true) there are no worlds in which it's false. Indeed, many of us once were and perhaps still are in the grips of this view of counterpossibles. But no one, if they are being honest with themselves, ever found it plausible that their moral, modal, or mathematical beliefs were vindicated by the doctrine.

---

<sup>33</sup> Thanks to Daniel Story for helpful discussion.

Here is our diagnosis. Sensitivity—like counterfactual correlations generally—is a decent but fallible indicator of the presence of an e-connection. That’s why finding out that some belief is sensitive is typically reason to think that your beliefs are in good standing. But they can potentially come apart, with beliefs turning out sensitive (if true) for reasons having nothing to do with any e-connection. And it is in these latter cases that putative sensitivity brings no relief. This is exactly what we would expect if m-connections like sensitivity are of epistemic interest only derivatively, owing to being indicative of an e-connection.

Second, there are cases of *safety without relief*. Typically, seeing that you couldn’t easily have been mistaken about something is reassuring. But no sensible person in Jack’s shoes would feel relieved upon reaching the conclusion that they couldn’t easily have been mistaken about the protons, all the while granting that they believe what they do about protons for reasons having nothing at all to do with the facts about protons. Again, our diagnosis is that safety is a decent but fallible indicator of the presence of an e-connection, and the former derives its epistemic significance from the latter. Thus, when you see that your belief is safe (if true) for reasons having nothing to do with an e-connection, the putative safety brings no relief.

Third, there are cases of *insensitivity without concern*. Typically, seeing that you would have held some belief even if it were false is reason to abandon that belief. But in some cases, insensitivity is intuitively no threat at all. For instance, the belief that you aren’t wrong in thinking that Omar has new shoes. Why is it perfectly sensible to retain these beliefs even after seeing a cut and dry demonstration of their insensitivity? Our diagnosis is that concessions of insensitivity derive their epistemic import from their relation to e-connectedness. Typically, insensitivity is a good but fallible indicator of an explanatory disconnect. That’s why discovering that your beliefs are insensitive typically serves as a defeater. But in these other cases, one can see that they are insensitive for reasons having nothing to do with an explanatory disconnect. Your belief is explained by the facts about Omar’s shoes, and its insensitivity gives you no reason to think otherwise. Because, in this case, there is reason to doubt that the insensitivity is the result of an absent e-connection, the insensitivity is no cause for concern. This is what one would expect if m-connections derived their epistemic import from their relation to e-connections.

Finally, there are cases of *unsafeness without concern*. Typically, discovering that your beliefs are unsafe is cause for concern. Upon discovering that the menus have been tampered with by a prankster, you realize that there are nearby worlds in which you falsely believe that today's special is minestrone, and you lose your justification for believing that that's today's special. But there are other cases in which recognized lack of safety doesn't undermine your beliefs. Here is one such case:

The world's most accurate clock hangs in Smith's office, and Smith knows this. Its accuracy is due to a radiation sensor. This radiation sensor is very delicate, however, and could easily malfunction if a radioactive isotope were to decay in the vicinity. This morning, against the odds, someone did in fact leave a small amount of a radioactive isotope near the clock. The isotope will decay at any moment, disrupting the clock's sensor and freezing the clock on "8:22". The clock is still running normally at 8:22 when Smith enters her office. Smith takes a good hard look at the clock and forms the true belief that it is 8:22.<sup>34</sup>

First things first: she does know that it is 8:22, even though her belief is unsafe. So safety is not a precondition for knowledge. Second, and more importantly for our purposes, *discovering* that the belief is unsafe is also no cause for concern. Suppose Smith later learns that the isotope was present but that it hadn't yet decayed. She can then recognize that her belief that it is 8:22 is unsafe: there are nearby worlds where she falsely believed that it is 8:22. But surely—especially now knowing that decay has not occurred—she is not rationally required to withhold belief.

Here is our diagnosis of why the discovery that a belief is unsafe sometimes is a defeater and other times isn't. The absence of safety is a decent but fallible indicator of the absence of an e-connection. The Recognized lack of safety defeats in the menu case in virtue of revealing the absence of an e-connection. But it doesn't defeat Smith's belief about the time because there is a known e-connection in this case: Smith believes it was 8:22 when she walked in because it indeed was 8:22, and she knows this.

What these last two families of cases suggest is that Safety First and Sensitivity First and other such minimalist-friendly priority theses have things exactly backwards. Absences of m-connectedness undermine in virtue of what they reveal about e-connectedness, not vice versa. In other words:

---

<sup>34</sup> The case is due to Tomas Bogardus (2016: 649-650).



*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.

This is the lesson we ought to draw from the above cases.

More generally, it is the lesson we should draw from the troubled history of developing and defending sensitivity and safety constraints on knowledge. Such attempts have been foiled by counterexample after counterexample, involving cases of insensitive or unsafe beliefs that nevertheless count as knowledge. And there is something that these counterexamples almost invariably in common: the beliefs in question, despite being unsafe or insensitive, do (by the believer's lights) bear an e-connection to the associated facts. *Explanation First* nicely accounts for this.

What would be trouble for our account is if there could be cases of explanatory concessions without concern. Yet it is hard to see how one would get cases of this sort. The only putative cases we are aware of are those offered up by the minimalist: moral beliefs that are by-one's-own-lights safe and sensitive despite not being e-connected. But even the minimalist surely won't claim that these *intuitively* count as unproblematic explanatory concessions. The only grounds for thinking that these beliefs remain in good standing despite the explanatory concession is the deeply suspicious minimalist's gambit. But that gambit takes for granted the very priority thesis that we have argued against: that explanatory concessions derive their epistemic import from what they reveal about m-connections.

## 10. 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 (and that if anything the reverse is true). If that is right, the minimalist gambit is dead on arrival. Defeat cannot be averted by reasoning from the beliefs that one concedes are not e-connected to the conclusion that they bear some modal "tracking" relation to the associated facts. That is because the beliefs from which the reasoning begins have already been defeated by the explanatory concession. Thus, while minimalists are entirely correct that there are

possibilities according to which our moral beliefs are m-connected without being e-connected, they cannot rationally *believe* any such possibility is actually the case.

The failure of minimalist strategies also shows us what it takes to resist the evolutionary challenge: 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 somehow 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.

Additionally, the foregoing shows that 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. A superior formulation, relying directly on an explanatory constraint on rational belief, might run as follows:

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

(P2\*) If S believes that her belief that p is not e-connected, then S is thereby rationally committed to withholding belief that p.

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

Finally, although we won't pursue the idea here, we suspect that the objections we have raised against minimalist responses to moral debunking arguments 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.<sup>35</sup> In each of these cases, the strategy is to grant that our beliefs are not e-connected, but then reason from the

---

<sup>35</sup> See also Korman (2014: §§3-4, 2015: §7.3, forthcoming: §7.2) on minimalist responses to debunking arguments against perceptual beliefs about ordinary objects.

disconnected beliefs to some account of our reliability. But, if our argument above is successful, then these minimalist strategies fail for the same reasons as moral minimalism. The only viable responses to these challenges are “explanationist” responses on which our beliefs (or those from which they are inferred) bear an appropriate explanatory connection to their subject matter.

## 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.
- Bedke, Matthew (2014), ‘No Coincidence?’, *Oxford Studies in Metaethics* 9: 102-125.
- 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 (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, Willian 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.
- 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 (2014), 'Darwinian Normative Skepticism', in Michael Bergmann and Patrick Kain (eds.), *Challenges to Moral and Religious Belief* (Oxford: Oxford University Press), pp. 220-236.
- Lott, Michael (forthcoming), 'Must Realists Be Skeptics? An Aristotelian Reply to a Darwinian Dilemma', *Philosophical Studies*.
- Mogensen, Andreas L. (2015), 'Evolutionary Debunking Arguments and the Proximate/Ultimate Distinction', *Analysis* 75: 196-203.
- Moon, Andrew (forthcoming), 'Debunking Morality: Lessons from the EAAN Literature', *Pacific Philosophical Quarterly*.
- 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.
- Ruse, Michael (1986), *Taking Darwin Seriously* (Prometheus).
- Schafer, Karl (2010), 'Evolution and Normative Scepticism', *Australasian Journal of Philosophy* 88: 471-488.

- Schechter, Joshua (forthcoming), 'Explanatory Challenges in Metaethics', in Tristram McPherson and David Plunkett (eds.), *Routledge Handbook of Metaethics* (Routledge).
- 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 Priority', 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 Sceptics* (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).
- Wright, Crispin (2004), 'Warrant For Nothing (and Foundations for Free)?', *Aristotelian Society* 78: 167-212.