Modal Security and Evolutionary Debunking

Daniel Z. Korman and Dustin Locke

Forthcoming in Midwest Studies in Philosophy

1. Introduction

Evolutionary moral debunking arguments aim to show that certain evolutionary considerations undermine our moral beliefs, at least if those beliefs are realistically construed.[1] Formulations vary, but as an illustration, a debunker might argue as follows: We have the particular moral beliefs we do (e.g., the belief that we are morally obligated to feed our children) because those beliefs motivated our ancestors to perform actions that enhanced reproductive success, but the putative moral facts themselves (e.g., the fact that feeding your children really is morally obligatory) have no role to play in explaining why those actions enhanced reproductive success.[2] On the basis of such information, the debunker claims, we are justified in believing the following thesis:

> **Moral Disconnect**
> No moral beliefs are explained by moral facts.

Once one comes to be justified in believing Moral Disconnect, the debunker argues, one's moral beliefs (realistically construed) are rendered unjustified. Notably, justified belief in Moral Disconnect is meant to render those beliefs unjustified without giving one reason to believe that they are false. It does not "rebut" them; it "undermines" them.

Justin Clarke-Doane, both independently and in co-authored work with Dan Baras, has attempted to block evolutionary moral debunking arguments by advancing a principle he calls "Modal Security". The basic idea behind Modal Security is that evidence undermines only when it calls into question the "security" of one's belief, where a belief is secure iff$_{\text{def}}$ it is both "safe" and "sensitive". Very roughly, S's belief that p is safe if and only if S could not easily have been

---

[1] See especially Joyce (2006) and Street (2006). See Vavova (2015) for an overview of the debate. By "realistically construed", we mean: construed in the way that moral realists construe them (as, at least in some cases, representing mind-independent truths); cf. Clarke-Doane (2012: §1).

[2] Cf. Street (2006: §6).

wrong about whether p (using the method by which S actually forms a belief about p), and S's belief that p is *sensitive* if and only if had it been the case that not-p, S would not still have believed that p (using the method by which S actually forms a belief about p).[3] Hence, Modal Security says that evidence undermines only if it calls into question certain purportedly important modal connections between one's beliefs and the truth.

It's one thing to state the rough idea behind Modal Security, as we've just done, and another to formulate a precise principle that is capable of both withstanding counterexamples and blocking moral debunking arguments. Clarke-Doane's formulation of Modal Security has undergone considerable evolution in the short time since he introduced it. Here is the earliest formulation:

> "Information, E, cannot undermine our D-beliefs without giving us some reason to believe that our D-beliefs are not both safe and sensitive." (2015: 97)

In subsequent papers, the appeal to "reason to believe… not…" is replaced with "reason to doubt":

> "If information, E, undermines all of our beliefs of a kind, D, then it does so by giving us reason to doubt that our D-beliefs are both sensitive and safe." (2016a: 31)[4]

Notice that this second formulation requires not only that undermining evidence provide reason to doubt sensitivity or safety, but also that it undermine *by way of* providing such reasons. In later work with Dan Baras, this explanatory language ("by") disappears, as does the reference to entire classes of beliefs ("D-beliefs"), and a new notion appears ("direct reason to doubt"):

> "If evidence, E, undermines our belief that P, then E gives us direct reason to doubt that our belief is sensitive or safe *or* E undermines the belief that <the belief that P is safe and sensitive>" (2021: 178)[5]

---

[3] We ignore a further qualification in the definition of safety to the effect that one also couldn't easily have been mistaken about propositions "similar enough" to p; see Clarke-Doane and Baras (2021: 164) on the need for this qualification.
[4] Cf. his (2016b: 30) and (2017: 208).
[5] Cf. Clarke-Doane (2020: 117).

The latter principle involves a number of vexing components, not least the "recursive clause" in its consequent. We will eventually discuss this principle directly, but to appreciate what it says, why it says what it says, and what we want to say about it, it is best to start simple.

We begin by reviewing Clarke-Doane and Baras' motivations for endorsing a security principle—that is, some principle that requires undermining evidence to in one way or another call into question the security of one's belief. We then examine some security principles that are more straightforward than those mentioned above, showing that some of these are too strong, failing to accommodate clear cases of undermining, while others are too weak, failing to do their advertised work of blocking evolutionary moral debunking arguments. This discussion will pave the way for our consideration of Clarke-Doane and Baras' most recent formulation and our argument that it is likewise too weak to block evolutionary moral debunking arguments. We conclude that if there is security principle that slips between the horns of this dilemma—one that is both viable and debunker-blocking—it remains to be formulated.

2. Why Security?

Before looking at specific security principles, let us ask: why think that undermining has anything to do with calling into question the security of one's beliefs? Clarke-Doane and Baras (2021) offer two motivations.

Clarke-Doane and Baras' first motivation is that evidence that undermines a belief must provide some reason to doubt that the belief has "some important epistemic feature" and "safety and sensitivity, along with truth, have emerged from recent epistemology literature as among the most important epistemological features of beliefs" (2021: 165). For the record, we, along with virtually all contemporary epistemologists, view sensitivity as epistemically insignificant.[6] But we grant, at least for now, that (some suitably sophisticated version of) safety may be epistemically important.

Clarke-Doane and Baras' second motivation is that paradigmatic cases of undermining "conform to" the idea that evidence undermines only if it calls into question the security of one's belief (2021: 165). Suppose, for example, that you believe p but then learn that your belief was

---

[6] See Vogel (2012) for a catalog of concerns about sensitivity constraints on knowledge, many of which can be adapted to show that challenges to the sensitivity of beliefs cannot undermine those beliefs. Cf. our (2020: §6).

caused by a pill that unpredictably gives rise to belief in p or belief in ~p. It seems that learning that your belief was caused by this pill undermines your belief that p. Security principles can accommodate this result, since what you learn seems to give you reason to believe that your belief that p is not safe—that is, that you easily could have formed, via the same belief forming method, a false belief about whether p.

How exactly should we view these two "motivations" for adopting a security principle? The first thing to note is the obvious: neither is a deductive argument. Rather, what each of these motivations claims is that there is a certain minimum bar that any necessary condition on undermining must meet and that security principles meet that bar. For example, the first motivation does not assert that safety and sensitivity (and truth) are the *only* epistemically important features of belief. Rather, it asserts that evidence that undermines must be evidence that bears on *some* epistemically important feature of belief, and safety and sensitivity (along with truth) are "among" the most important epistemic features of belief. Likewise, the second motivation does not assert that paradigm cases of undermining can be understood *only* in terms of safety and sensitivity.[7] Rather, it merely asserts that paradigmatic cases of undermining "conform to"—that is, are not counterexamples to—the idea that undermining evidence is evidence that bears on safety and sensitivity.

The point of the preceding paragraph was not to criticize Clarke-Doane and Baras' two motivations for endorsing a security principle, but simply to get clear about what those motivations do and do not show. In what follows, we will keep these two motivations in mind as we survey various security principles, with an eye to which security principles they motivate, and which are simply unmotivated.

3. Basic Security

So far, we have formulated the main idea behind security principles like this: E undermines one's belief that p only if E in one way or another "calls into question" the security of one's belief. A loose formulation like this is likely to land the proponent of such a principle in equivocation,

---

[7] We have argued elsewhere (Korman and Locke 2020 and 2023), that it is not revelations about the safety or sensitivity of one's beliefs but rather the revelation that there is no such explanatory connection that accounts for why undermining occurs in paradigm cases of undermining (like the pill case).

relying on a (weaker) reading of the principles when arguing for its plausibility and another (stronger) reading when wielding the principle against debunking arguments. What's needed is a more precise formulation.

A natural place to begin is with a principle we'll call

Basic Security

E undermines S's belief that p only if E renders S unjustified in believing that her belief that p is secure.

Like the other security principles to come, Basic Security is what we might call an 'order-shifting' principle: on the left-hand side we find a statement about E's bearing on the belief that p, and on the right-hand side we find a statement about E's bearing on a certain higher-order *belief about* one's belief that p.

We think that Basic Security is highly plausible. It does indeed seem that if E does not render you unjustified in believing that your belief that p is secure, then E cannot possibly undermine your belief that p. As Clarke-Doane (2015: 96-97) puts the idea, if a belief is both safe and sensitive, then it is in a certain sense "bound to be true", and if some evidence does not "threaten" your belief that your belief is bound to be true, then it hardly seems that that evidence can undermine your belief that p. In any case, we have no interest in challenging Basic Security. What we will argue is that Basic Security is too weak to block the moral evolutionary debunking argument.

Simplifying somewhat,[8] the debunker's contention is that Moral Disconnect—the thesis that moral beliefs are not explained by moral facts—undermines our moral beliefs. For this reason, we'll say that a security principle is *debunker-blocking* when accepting it is an obstacle to reasonably maintaining that Moral Disconnect undermines our moral beliefs. And we'll say that a security principle is *debunker-friendly* when it is not debunker-blocking.

---

[8] Strictly speaking, it is only first-order moral beliefs, and not metaethical beliefs like belief in Moral Disconnect itself, that are said to be undermined; it is only moral realists who are meant to be targeted by the argument; and it is not Moral Disconnect itself but rather justified belief in, or acquiring evidence of the truth of, Moral Disconnect that is meant to do the undermining. But, for ease of exposition, let's ignore those complications.

We contend that Basic Security is debunker-friendly. The only constraint that Basic Security places on the debunking argument is this: Moral Disconnect undermines our moral beliefs only if Moral Disconnect undermines belief in the security of our moral beliefs. Our contention is that Moral Disconnect does undermine belief in the security of our moral beliefs. To show that Moral Disconnect undermines belief in the security of our moral beliefs, we need to either show that it undermines belief in the safety of our moral beliefs or that it undermines belief in the sensitivity of our moral beliefs. We'll focus on safety, but a similar argument applies to sensitivity.

Let's begin by asking what, if anything, makes us prima facie justified in believing in the safety of our moral beliefs? The only plausible sources of justification that we are aware of ultimately depend for their justification on moral beliefs themselves. Take, for instance, the belief *that promises create pro tanto obligations*. A plausible account of why Dan is justified in believing this belief to be safe would go something like this:

1. Dan is prima facie justified in believing that necessarily (or at least in all nearby possible worlds), promises create pro tanto obligations. How? Perhaps on the basis of its self-evidence, or a moral intuition to this effect.[9]

2. Dan is prima facie justified in believing that it couldn't easily have been the case that he believed, via the same belief-forming method, that promises didn't create pro tanto obligations. After all, as the debunker herself would admit, we are evolutionarily hard-wired with dispositions to form fitness-enhancing beliefs like this, and the evolution of moral cognition doesn't unfold radically differently in *nearby* worlds.

3. Thus, Dan is prima facie justified in believing that it couldn't easily have been the case he formed a mistaken belief (via the same belief-forming method) about whether promises create pro tanto obligations. In other words, Dan is prima facie justified in believing that his belief that promises create pro tanto obligations is safe.[10]

---

[9] See Korman and Locke (2020: §4) on why drawing on one's moral intuitions in this context cannot be dismissed as "question begging".

[10] Something in the vicinity of this style of reason has been advanced by numerous authors: Nozick (1981: 346-348), Enoch (2010), Clarke-Doane (2015: 92-95, 2016a: §2.3, 2020: chs. 4.4 and 5.8), Huemer (2005: §8.6.4), Brosnan (2011), Talbott (2015), Wielenberg (2010), Skarsaune (2011), and Vavova (2015: §6). We say "in the vicinity", since the reasoning is typically advanced as part of an explanation of the reliability of our moral beliefs, as opposed to an explanation of how we are justified in believing them to be reliability.

Belief in the safety of beliefs about contingent moral propositions is justified in a similar way. Take, for instance, the belief *that Dan is pro tanto obligated to meet Dustin for lunch*:

1*. Dan is prima facie justified in believing that necessarily (or at least in all nearby possible worlds), whether he is pro tanto obligated to meet Dustin for lunch depends on various factors such as whether he promised to meet Dustin for lunch, whether Dustin is a friend in need of his help, whether some sufficiently greater good would come of his meeting Dustin for lunch, etc. How? Perhaps on the basis of moral intuitions about what such obligations depend on.

2*. Dan is prima facie justified in believing that it couldn't easily have been the case that he formed (via the same belief-forming method) a belief about [whether he is pro tanto obligated to meet Dustin for lunch] in a way that was not responsive to the factors described in (1*). After all, as the debunker herself would admit, we are evolutionarily hard-wired with dispositions to form fitness-enhancing beliefs about what sorts of factors are right-making, wrong-making, obligation-making, etc., and the evolution of moral cognition doesn't unfold radically differently in *nearby* worlds.

3*. Thus, Dan is prima facie justified in believing that it couldn't easily have been the case that he formed a mistaken belief (via the same belief-forming method) about whether he was pro tanto obligated to meet Dustin for lunch. In other words, Dan is prima facie justified in believing that his belief that he is pro tanto obligated to meet Dustin for lunch is safe.

According to these answers, our prima facie justification for believing that our moral beliefs are safe rests on our prima facie justification for having certain moral beliefs, for instance the sort of moral beliefs described in (1) and (1*). Yet if answers like these give the correct explanations of how it is that we are prima facie justified in believing in the safety of our moral beliefs, then it is easy to see why Basic Security is no obstacle to the claim that Moral Disconnect undermines our moral beliefs. For if, as the debunker contends, Moral Disconnect undermines our moral beliefs, then it undermines *inter alia* the moral beliefs that are, according to (1) and (1*), prima facie justified. And if Moral Disconnect undermines the moral beliefs referenced by (1) and (1*), then it also undermines any belief that depends for its justification on our being justified in

having the beliefs described in (1) and (1*), including the belief that our moral beliefs are safe. Thus, Basic Security—according to which E undermines one's belief that p only if E undermines one's belief in the security of one's belief that p—is no obstacle to the claim that Moral Disconnect undermines our moral beliefs.

What we have just shown is that "defeat trickles up": whatever defeats our moral beliefs thereby also defeats any beliefs that depend for their justification on our moral beliefs, including the belief that our moral beliefs are safe.[11] By the debunker's lights, then, Moral Disconnect undermines our moral beliefs and *for that very reason* undermines belief in the safety and thus the security of those beliefs. Debunkers who accept Basic Security can therefore reasonably maintain that Moral Disconnect undermines our moral beliefs. In other words, Basic Security is debunker-friendly.

Those in search of a debunker-unfriendly security principle must find a formulation that blocks this sort of trickle-up defeat. One strategy would be to formulate the principle in such a way that evidence which undermines a belief is required to do so *by* rendering one unjustified in believing that that belief is secure. Clarke-Doane once embraced security principles to this effect (as we saw in §2), but he has since abandoned them, and rightly so: such formulations are open to powerful counterexamples which we have advanced elsewhere (and will not rehearse here).[12] Another strategy is to require evidence that undermines a belief to *rebut* the belief that that belief is secure. We'll explore that strategy in the following section.

Before assessing alternative security principles, however, it's worth noting that the search for a debunker-blocking security principle is, in an important sense, unmotivated. As far as we can tell, the two motivations for security principles offered by Clarke-Doane and Baras (canvassed in §2) do not motivate anything stronger than Basic Security.[13] Basic Security says that evidence

---

[11] Cf. Schechter (2018: 335-336) and Noonan (forthcoming: n.14). To escape the defeat-trickles-up argument, one would need to provide an account of our prima facie justification for believing that our moral beliefs are safe that does *not* depend on our prima facie justification for our moral beliefs. And we don't see how such an account would go. It strains credulity, for instance, that the belief *that one could not have easily had a false belief as to whether or not p, using the method that one actually used to determine whether or not p, where p is a moral proposition* is properly basic (e.g., on account of being self-evident).

[12] See Korman and Locke (2020); cf. Noonan (forthcoming).

[13] Cf. Schechter (2018: 335-336) and Woods (2018: 56) on more demanding security principles being unmotivated.

undermines the belief that p only if it renders one unjustified in believing that one's belief that p is secure. This by itself accommodates paradigm cases of undermining (like the pill case) and does justice to the idea that the evidence that undermines a belief must bear on an epistemically important feature of belief (granting, for the sake of argument, that at least safety is epistemically important). In short, we have already found a perfectly good, well-motivated, and (bonus!) ecumenical security principle. Its only shortcoming is that it doesn't do what anti-debunkers like Clarke-Doane and Baras hoped it would do. There is thus no reason to look any further, apart from a desire for a debunker-blocking principle. With that in mind, let us continue the search.

4. Rebutting Security

Basic Security is debunker-friendly because undermining our moral beliefs involves *ipso facto* undermining our belief that our moral beliefs are secure. What undermining our moral beliefs doesn't *ipso facto* do, however, is give us reason to believe that our moral beliefs are *not* secure. Thus, a natural security principle to consider is a principle we'll call

> **Rebutting Security**
> E undermines S's belief that p only if E rebuts S's belief that her belief that p is secure—that is, only if E gives S reason to believe that her belief that p is not secure.

Rebutting Security is debunker-blocking. Even those who think that Moral Disconnect undermines our moral beliefs should agree that Moral Disconnect doesn't give us reason to believe that those beliefs are *not* secure. Rather, they should insist that we adopt a skeptical attitude, withholding belief one way or the other about whether our moral beliefs are secure. In particular, they should concede that *if* our moral beliefs happen to be correct, then they may well be secure as well. (For instance, *if* your lunch obligation does in fact supervene on your promise in the way you think, then you couldn't easily have been mistaken about whether you have such obligations.[14]) But one who accepts Rebutting Security cannot think that Moral Disconnect undermines our moral beliefs in this way—that is, without rebutting belief in the security of those beliefs. Rebutting Security is debunker-blocking.

---

[14] Cf. Clarke-Doane (2015: §4-6, 2020: chs. 4.4 and 5.8) and Clarke-Doane and Baras (2021: §3).

However, Rebutting Security is subject to a range of counterexamples. Here is one that Clarke-Doane and Baras themselves introduce, to motivate the need for a more sophisticated security principle.[15]

> **WORM TESTIMONY**
> Wanda (the wonderer) wonders whether worms are insects, and Abe (the attester) tells her that they are. Not knowing whether to trust Abe, Wanda consults Ben (the booster), who vouches for Abe's reliability. Wanda then, on the basis of Abe's testimony, forms the belief that worms are insects (hereafter: her first-order worm belief). Additionally, on the basis of Ben's testimony, she forms the belief that her belief that worms are insects is secure (hereafter: her second-order worm belief). Finally, because she thinks that Ben is reliable, she thinks that beliefs formed on the basis of his testimony are secure, and so she forms the belief *that her belief that her belief that worms are insects is secure* is secure (hereafter: her third-order worm belief). Cas (the challenger), whom Wanda has overwhelming reason to trust, then tells Wanda that Ben cannot be trusted.

Cas's testimony plainly undermines Wanda's first-order worm belief. But Rebutting Security wrongly implies that it doesn't. Here's why. After hearing Cas's testimony and thereby losing her (Ben-sourced) reasons for believing that Abe is a reliable informant, Wanda is left without sufficient reason to believe that Abe is reliable and without sufficient reason to believe that Abe isn't reliable. Consequently, Cas's testimony renders her unjustified in believing that her first-order worm belief is secure. That is, it makes it the case that she ought to withdraw her second-order worm belief. But Cas's testimony does not give Wanda any reason to believe that her first-order worm belief is *not* secure—for all she knows, Abe is indeed reliable—and thus does not rebut her second-order worm belief. Thus, Rebutting Security wrongly implies that in WORM TESTIMONY, Cas's testimony doesn't undermine Wanda's first-order worm belief.

In retrospect, we should have expected that Rebutting Security would be subject to counterexamples like WORM TESTIMONY. According to Rebutting Security, E undermines your belief that p only if E indicates that your belief that p is *not* secure. Intuitively, such a requirement is simply too strong: insofar as one thinks that the epistemic status of one's higher-order belief about the security of one's belief that p has bearing on the epistemic status of one's belief that p,

---

[15] See Clarke-Doane (2020: 116-117) and Clarke Doane and Baras (2021: 177). We have taken the liberty of filling in some details of the case.

one ought to think that evidence that merely leaves one unjustified in having the higher-order belief is enough to do the trick.[16] Cases like WORM TESTIMONY make this intuition concrete.

5. Iterated Security

Basic Security is plausible but debunker-friendly. Rebutting Security is debunker-blocking but implausible. Anti-debunkers must look elsewhere for a security principle that is both debunker-blocking and plausible. Where to look?

Here is a clue. Notice that, while Cas's testimony does not rebut Wanda's second-order worm belief, it does rebut her third-order worm belief—that is, her belief that her second-order worm belief is itself secure. By Wanda's lights, her second-order belief is secure (if it is) because Ben is a reliable source. Cas's testimony justifies her in believing that Ben is not reliable and thus justifies her in believing that her second-order worm belief is not secure. In other words, Cas's testimony rebuts her third-order worm belief. A security principle that permits this sort of *even higher*-order challenge to undermine Wanda's first-order worm belief would be poised to accommodate WORM TESTIMONY.

To formulate a security principle that can do the job, it will be helpful to introduce some jargon. Let *the p-security propositions* refer to all and only the following propositions:

(1) the proposition that S's belief *that p* is secure

(2) the proposition that S's belief *that S's belief that p is secure* is secure

(3) the proposition that S's belief *that S's belief that S's belief that p is secure is secure* is secure

(4) …

---

[16] Why? Because if the epistemic status of one's (higher-order) belief that one's belief that p is secure has bearing on the epistemic status of one's (first-order) belief that p, then, presumably, if one is not justified in believing that one's belief that p is secure, then one is not justified in believing that p. (Cf. Woods 2018: §3.2.) Some may object that such a view makes undermining too easy, since it might seem to imply that children, who have no beliefs about the security of their beliefs, have no justified beliefs at all. But our claim here implies no such thing. The fact that children lack the *concepts* of sensitivity, safety, security, etc., is no objection to the claim that children are (propositionally) justified in believing that their beliefs are secure.

With the notion of p-security propositions in hand, we can now formulate a new security principle that is poised to handle WORM TESTIMONY:

Iterated Security

E undermines S's belief that p only if E rebuts S's belief in one of the p-security propositions.[17]

Here is how Iterated Security permits Cas's testimony to undermine Wanda's first-order worm belief. According to Iterated Security, the conditions for undermining her first-order worm belief are satisfied so long as a belief of hers with any of the following contents is rebutted:

(1') the proposition that her belief that worms are insects is secure

(2') the proposition that her belief that her belief that worms are insects is secure is secure

(3') the proposition that her belief that her belief that her belief that worms are insects is secure is secure is secure

(4') …

As we saw two paragraphs back, Cas's testimony (that Ben-sourced beliefs are not secure) does rebut Wanda's belief in proposition (2'), aka her third-order worm belief. Thus, Cas's testimony does satisfy the constraint that Iterated Security places on undermining; Iterated Security rightly permits Cas's testimony to undermine Wanda's first-order worm belief.

Is Iterated Security debunker-blocking? To assure oneself that it is debunker-blocking, one would have to confirm that Moral Disconnect does not rebut belief in any of the following (infinitely many) propositions:

(1") the proposition that our belief that we're obligated to keep our promises is secure

(2") the proposition that our belief that our belief that we're obligated to keep our promises is secure is secure

---

[17] More in §8 about how Iterated Security relates to Clarke-Doane and Baras' preferred security principle.

(3″) the proposition that our belief that our belief that our belief that we're obligated to keep our promises is secure is secure is secure

(4″)…

We saw in §4 that Moral Disconnect doesn't rebut belief in proposition (1″). To check whether Iterated Security is truly debunker-blocking, we would need to check all the rest in order to be sure that it doesn't rebut belief in any of them either. (We leave that as an exercise for the reader.) However, as we shall see shortly, it hardly matters whether Iterated Security is debunker-blocking, since—for *any* belief—there is readily available evidence that meets Iterated Security's conditions for undermining it.

6. Against Iterated Security

We will raise two objections to Iterated Security. The first turns on a simple observation: people do not normally form beliefs about the security of their beliefs, let alone beliefs about the security of their beliefs about the security of their beliefs. Iterated Security is able to handle WORM TESTIMONY only because we artificially built into the case that Wanda (the weirdo) forms these esoteric higher-order beliefs. But consider a variation on WORM TESTIMONY, whose protagonist forms beliefs like a normal person:

> WORM REDUX
> Norma (the normal person) wonders whether worms are insects, and Abe (the attester) tells her that they are. Not knowing whether to trust Abe, Norma consults Ben (the booster), who vouches for Abe's reliability. Norma then, on the basis of Abe's testimony, forms the belief that worms are insects. But it does not occur to Norma to wonder whether her belief that worms are insects is secure, and she does not form a belief one way or the other about whether it is. Cas (the challenger), whom Norma has overwhelming reason to trust, then tells Norma that Ben cannot be trusted.

Like Wanda, Norma is, just after consulting Ben, justified in believing that worms are insects, and Cas's testimony subsequently undermines her belief that worms are insects. However, by the lights of Iterated Security, Cas's testimony undermines Norma's worm belief only if it rebuts belief in one of the following propositions—that is, only if it gives her reason to believe that one of them is false.

(1') the proposition that her belief that worms are insects is secure

(2') the proposition that her belief that her belief that worms are insects is secure is secure

(3') the proposition that her belief that her belief that her belief that worms are insects is secure is secure is secure

(4') …

Cas's testimony does not give Norma reason to believe that (1') is false. At most, belief in (1') is undermined.

Does Cas's testimony give Norma reason to believe that any of the other security propositions are false? Apart from (1'), each of the security propositions makes reference to a non-existent belief, a belief that Norma does not have. We are prepared to agree that Norma *does* have reason to believe those propositions are false and, indeed, is (propositionally) justified in believing that they are false. But what gives Norma reason to believe that those propositions are false isn't Cas's testimony but rather the simple fact that there is no such thing as "her belief that her belief that worms are insects is secure." At least on a Russellian account of reference failure, (2') is the proposition that *there is exactly one x such that (i) x is her belief that her belief that worms are insects is secure & (ii) x is secure*. So understood, Norma is (propositionally) justified in believing that (2') is false because she is justified in believing that she has no beliefs about the security of her worm beliefs, and thus that the first conjunct of the Russellian analysis is not satisfied. So, while she does arguably have reason to believe that (2') is false, it's for reasons having nothing to do with Cas's testimony.

Since Cas's testimony does not (itself) give Norma reason to believe that either (1') or any of the other security proposition is false, Iterated Security wrongly entails that Cas's testimony doesn't undermine Norma's worm belief.

Our second objection is that Iterated Security is in direct tension with one of Clarke-Doane and Baras' stated motivations for security principles.[18] The objection begins by observing that beliefs about the sensitivity of one's beliefs—when one *does* form them—are invariably going to be insensitive.[19] For example, suppose you have not just the first-order belief that you have hands,

---

[18] The argument that follows is in the spirit of Huemer (2001: 186), White (2010: 581), and Vogel's (2012: 130-131) argument that sensitivity constraints mishandle higher-order beliefs.

[19] There may be rare exceptions, e.g., Cogito beliefs. Let's ignore that complication.

but also the second-order belief that your first-order belief that you have hands is sensitive. Is your second-order belief sensitive? To answer this question, we need to consider the nearest world—call it 'W'—in which your first-order belief that you have hands is not sensitive, and ask whether, in that world, you still would have believed that it was (using the same methods you actually used to form this second-order belief). As we are about to see, you still would have.

By definition, in W, your belief that you have hands is not sensitive. However, you wouldn't believe that you had hands unless it *seemed to you* that your perceptual faculties were in good working order. So, in W, you believe that your belief that you have hands is sensitive, and you form this second-order belief by the same methods that you form it in the actual world. In W, then, your belief that you have hands is not sensitive, but you believe that it is (using the relevant methods). In other words, your second-order belief that your first-order belief that you have hands is sensitive is not itself sensitive. A similar line of reasoning applies to virtually all of our beliefs, and since security requires sensitivity, we can conclude that virtually all of our beliefs about the security of our first-order beliefs are not themselves secure.

What we've just shown is that for virtually any *p* that you believe, your second-order belief that your first-order belief that *p* is secure—if you have such a second-order belief—is not itself secure. Since we've just shown this, you've just learned it (if you didn't already know it). Does it follow that virtually all of your beliefs are undermined? Of course not. What best explains why not? The best explanation, we think, is this: your first-order beliefs have not been undermined because the security (and, in particular, the sensitivity) of your *second-order* belief that your first-order belief that p is secure is not an epistemically important feature of your *first-order* belief that p. This means that Iterated Security is in direct tension with one of Clarke-Doane and Baras' stated motivations for security principles (discussed in §2), which claimed that evidence that undermines must be evidence that bears on some "important epistemic feature" of belief.[20]

7. Recursive Security

We have postponed discussion of Clarke-Doane and Baras' currently preferred security principle, which we mentioned way back in §1:

---

[20] A natural revision for the security theorist at this point is to drop reference to sensitivity and formulate Iterated Security purely in terms of safety. However, we believe that our Cloud Chamber case (2020: §7, 2023: §5) serves as a counterexample to an "Iterated Safety" principle.

"If evidence, E, undermines our belief that P, then E gives us direct reason to doubt that our belief is sensitive or safe *or* E undermines the belief that <the belief that P is safe and sensitive>" (2021: 178)

Adjusting the formulation to match that of the principles discussed above:

Recursive Security

Evidence E undermines S's belief that p only if (i) E gives S direct reason to doubt that her belief that p is secure **or** (ii) E undermines S's belief that her belief that p is secure.[21]

As the label suggests, Recursive Security is meant to be read as a recursive account of what it takes for evidence to undermine a belief. The second disjunct in the consequent mentions undermining, indicating that one way for some evidence to undermine a belief is for it to undermine the belief that that belief is secure in the way specified by the first disjunct—namely, by giving one direct reason to doubt that *the belief that that belief is secure* is secure.

Notice that our security principles and Clarke-Doane and Baras' are "packaged" somewhat differently. In particular, we have been deliberately formulating security principles without using their preferred locution, "direct reason to doubt". This is for a couple of reasons.

First, Clarke-Doane and Baras clearly think that some evidence can give one reason to doubt that p without giving one reason to believe not-p: "It is not obvious that every reason to doubt P is a reason to believe ~P. Some, for instance, are reasons to suspend judgment about P, which arguably is something else." (2021: 176 n.23). But (we think) they are just wrong about this. When a source tells you that there is sulfur on Mars, and you then discover that your source is unreliable, you do not in virtue of that have reason to doubt that there is sulfur on Mars. Perhaps you have reason to "doubt your belief" that there is sulfur on Mars, or reason to "have your doubts about whether" there is sulfur on Mars, or reason to be "in a state of doubt about whether" there is

---

[21] Clarke-Doane and Baras are explicit that "reason to doubt that the belief that p is safe or sensitive" is to be read as "reason to doubt that the belief that p is safe or reason to doubt that the belief that p is sensitive" (2021: 166), which is equivalent to "reason to doubt that the belief that p is secure".

sulfur on Mars, but you do not have "reason to doubt" that there is sulfur on Mars.[22] On our view, a reason to doubt that p is always a reason (often a very small and outweighed reason) to believe not-p. This divergence in our understanding of the notion of a "reason to doubt" makes it awkward for us to work with.

Second, the notion of a "direct", as opposed to an "indirect", reason to doubt is elusive, and Clarke-Doane and Baras are reluctant to provide a precise account of what it means (2021: 176-177). We will have more to say about possible interpretations of "direct reason to doubt" in the following section.

Nonetheless, despite our concerns about Clarke-Doane and Baras' notion of a "direct reason to doubt", we can offer a relatively straightforward critique of Recursive Security—a critique that goes through regardless of what they mean by "direct reason to doubt". The critique is just that Recursive Security is debunker-friendly, and it is debunker-friendly for exactly the same reason that Basic Security is debunker-friendly. That is, accepting Recursive Security is no impediment to maintaining that Moral Disconnect undermines moral beliefs. To see this, suppose you think that Moral Disconnect does undermine our moral beliefs and for that very reason undermines belief in the security of our moral beliefs (see §3 above). By your lights, then, Moral Disconnect does satisfy condition (ii) in the consequent of Recursive Security. One who accepts Recursive Security can therefore reasonably maintain that Moral Disconnect undermines our moral beliefs. Recursive Security is thus debunker-friendly.

8. Direct Iterated Security

We noted just above that Clarke-Doane and Baras' security principles are packaged somewhat differently from ours, invoking the notion of a "direct reason to doubt." Our last point of business in this paper is to revisit Iterated Security and assess whether repackaging it in terms of "direct reasons to doubt" affects our case against it. Consider, then, the following variation on Iterated Security:

Direct Iterated Security

---

[22] See Schoenfield (2022: 272) for an explanation of the distinction between a reason to doubt that p and a reason to be in a state of doubt with respect to p.

E undermines S's belief that p only if E gives S direct reason to doubt one of the p-security propositions.

We suspect it is this principle that Clarke-Doane and Baras (2021) meant to defend, despite their official endorsement of Recursive Security. Just after introducing Recursive Security, Clarke-Doane and Baras characterize Recursive Security as saying that "in order for evidence, E, to undermine our belief that P, there must be a chain of undermining evidence which bottoms out in a direct reason to doubt safety or sensitivity" (2021: 178). This characterization fits Direct Iterated Security better than Recursive Security. As we saw at the end of the previous section, Recursive Security turned out to be debunker-friendly because it does not actually require evidence that undermines our moral beliefs to give us direct reason to doubt the security of our moral beliefs (on any understanding of "direct reason to doubt"), or the security of our belief that our moral beliefs are secure, or…. and so on. Hence, Recursive Security does not actually require that there "be a chain of undermining evidence which bottoms out in a direct reason to doubt safety or sensitivity". Directed Iterated Security, by contrast, does imply that E cannot undermine our moral beliefs without being a direct reason to doubt the security of one of these higher-order beliefs. So, again, we suspect that it is Direct Iterated Security, and not Recursive Security, that Clarke-Doane and Baras actually meant to defend.

As noted above, Clarke-Doane and Baras are reluctant to provide a precise account of the notion of a "direct reason to doubt". This makes Direct Iterated Security exceedingly difficult to assess. Needless to say, we cannot address or anticipate every possible understanding of what a "direct reason to doubt" might amount to. But we can at least address one natural understanding of the notion.[23] On this understanding, the direct–indirect distinction has to do with whether the reason to doubt proceeds, so to speak, via some intermediary proposition. More precisely:

---

[23] Clarke-Doane and Baras plainly see a tight connection between getting direct reason to doubt that p and getting reason to believe ~p (see Clarke-Doane 2020: 99 and Clarke-Doane and Baras 2021: 176-177), suggesting that perhaps a "direct reason to doubt" that p *just is* a (possibly very small or heavily outweighed) reason to believe ~p. We set this interpretation aside, since Direct Iterated Security on this interpretation just is Iterated Security.

E gives S a direct reason to doubt that p iff$_{\text{def}}$ (i) E gives S reason to doubt that p and (ii) there is no further proposition q such that E gives S reason to doubt that p only in virtue of giving S reason to doubt that q.

Direct Iterated Security, on the assumption that "direct" means what the above says it means, amounts to

Evidence E undermines S's belief that p only if, for some p-security proposition q, (i) E gives S reason to doubt that q, and (ii) there is no proposition r such that E gives S reason to doubt that q only in virtue of giving S reason to doubt that r.

A wrinkle. Direct Iterated Security, so understood, involves the bare notion of a "reason to doubt" a proposition, and as indicated in §7 we and Clarke-Doane and Baras understand this notion differently. We think that a reason to doubt that p has to be a reason to believe ~p. If that's right, then Direct Iterated Security mishandles WORM REDUX in exactly the same way as Iterated Security: although Cas's testimony does give Norma reason to not believe the p-security proposition *that her belief that worms are insects is secure*, it does not give her reason to believe that this (or any other) p-security proposition is false. Accordingly, on our construal of "reason to doubt", she does not have a reason to doubt any p-security propositions, and Direct Iterated Security (so construed) wrongly entails that Cas's testimony doesn't undermine Norma's belief that worms are insects.

Suppose, however, that we understand "reason to doubt" as Clarke-Doane and Baras do, and count a mere reason to not believe p as a reason to doubt that p. In that case, Cas's testimony does give Norma a reason to doubt (i.e., a reason to not believe) that her belief that worms are insects is secure, and Direct Iterated Security delivers the correct result Cas's testimony undermines Norma's belief that worms are insects in WORM REDUX.[24] Charitably, then, that is how we ought to understand "reason to doubt" in Direct Iterated Security. Making it explicit:

Direct Iterated Security*

---

[24] That said, our second objection to Iterated Security in §6—that it is in tension with the stated motivations for security principles—applies equally to Direct Iterated Security so understood.

Evidence E undermines S's belief that p only if, for some p-security proposition q, (i) E gives S reason to not believe q, and (ii) there is no proposition r such that E gives S reason to not believe q only in virtue of giving S reason to not believe r.

It's easy to see, however, that this principle is overly demanding. Consider the following, final variation on WORM TESTIMONY:

WORM BRIBE

Abe tells Norma that worms are insects. Cas, whom Norma has overwhelming reason to trust, then tells Norma that Abe was bribed (by someone who neither knew nor cared whether this was true) to tell her this.

Cas's testimony plainly undermines Norma's first-order worm belief. According to Direct Iterated Security*, in order for Cas's testimony to undermine Norma's first-order worm belief, some proposition from the following list must meet certain conditions:

(1') the proposition that her belief that worms are insects is secure

(2') the proposition that her belief that her belief that worms are insects is secure is secure

(3') the proposition that her belief that her belief that her belief that worms are insects is secure is secure is secure

(4') …

Specifically, for some proposition on that list, Cas's testimony must give Norma reason to not believe that proposition, and it must do so directly, that is, not by way of giving her reason to not believe some other proposition.

Cas's testimony does arguably give Norma reason to not believe (1'). However, she gets reason to not believe (1') only in virtue of getting reason to not believe that Abe is faithfully reporting his actual beliefs about worms. Since Cas's testimony therefore does not give Norma a reason to not believe (1')—and since as far as we can tell it does not give her a reason to not believe any of the others either—Direct Iterated Security* wrongly entails that Cas's testimony does not undermine Norma's beliefs.

One last epicycle. Direct Iterated Security* mishandled WORM BRIBE because it prohibits Cas's testimony from undermining belief in p-security proposition (1') in a mediated way. But we know (from §3) that a security principle, if it is to be debunker-blocking, must somehow disallow evidence from undermining belief in p-security propositions by way of undermining belief in p itself. A tempting fix, then, would be to permit undermining to be mediated any which way with precisely one exception: evidence cannot undermine belief in p-security propositions by way of giving one reason to not believe p. More precisely:

Direct Iterated Security**

Evidence E undermines S's belief that p only if, for some p-security proposition q, (i) E gives S reason to not believe q, and (ii) it's not the case that E gives S reason to not believe q only in virtue of giving S reason to not believe p.[25]

Direct Iterated Security** does avoid the WORM BRIBE counterexample (exercise for the reader). The problem with Direct Iterated Security** is not that it is open to counterexamples, but that (like the other counterexample-free security principles) it is debunker-friendly. To see this, suppose you think (as the debunker does) that Moral Disconnect undermines your belief that you are obligated to keep your promises. (Call this your *promise-obligation belief*.) Direct Iterated Security** does not prevent you from saying so. After all, Moral Disconnect (by your lights) also gives you reason to not believe that you have at least some obligations. (Call this your *some-obligations belief*.) Since defeat trickles up (see §3), it also gives you reason to not believe that your some-obligations belief is secure, and *a fortiori* reason to not believe that your promise-obligation belief is secure. Condition (i) is satisfied because Moral Disconnect does give you reason to not believe that your promise-obligation belief is secure. Condition (ii) is satisfied because it does so not only in virtue of giving you reason to not believe that you are obligated to keep your promises, but also in virtue of giving you reason to not believe that you have any obligations at all. Accepting Direct Iterated Security** is therefore no obstacle to accepting that Moral Disconnect undermines moral beliefs. Direct Iterated Security** is debunker-friendly.

---

[25] Thanks to Chris Noonan for suggesting that we consider this security principle.

9. Conclusion

We have considered numerous security principles, but have found none that is well-motivated, permissive enough to avoid counterexamples, and yet strong enough to cause trouble for the debunker. Basic Security, which says that evidence undermines a belief only if it renders one unjustified in believing in the security of one's belief, was shown to be debunker-friendly, and thus unsuitable for anti-debunking purposes. Rebutting Security, which says that evidence undermines a belief only if it rebuts belief in that belief's security, was shown to be subject to counterexamples involving somewhat indirect challenges. Iterated Security, which says that evidence undermines a belief only if it rebuts *some* relevant security belief, likewise succumbs to counterexamples and was shown to be in tension with the stated motivations for security principles. Clarke-Doane and Baras' (officially) preferred security principle, Recursive Security, which says that evidence undermines a belief only if it gives one direct reason to doubt that the belief is secure or else undermines belief in that belief's security, was shown to be debunker-friendly. Finally, Direct Iterated Security, which we suspect is the principle that Clarke-Doane and Baras' actually prefer, and which says that E undermines S's belief that p only if E gives S direct reason to doubt one of the p-security propositions, was shown to either face counterexamples, be inconsistent with Clarke-Doane and Baras' stated motivation for endorsing a security principle, or both.

Some may feel that the repeated failure to find a security principle that is both debunker-blocking and counterexample-free just means we need to keep Chisholming away until we find the right security principle. As we observed in §3, however, we have already found "the right" security principle: Basic Security. Clarke-Doane and Baras offer a variety of reasons for embracing a security principle, and Basic Security checks off every reason on the list. The search for a security principle that is both counterexample-free *and* debunker-unfriendly is unmotivated, and even if one found it, it is hard to see what reason there would be to prefer it to the debunker-friendly Basic Security.[26]

References

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

Clarke-Doane, Justin (2012), 'Morality and Mathematics: The Evolutionary Challenge', *Ethics* 122: 313-340.

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

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

Clarke-Doane, Justin (2016b), 'What is the Benacerraf Problem?', in Fabrice Pataut (ed.), *Truth, Objects, Infinity* (Dordrecht: Springer), pp.17-44.

Clarke-Doane, Justin (2017), 'Debunking Arguments: Mathematics, Logic, and Modal Security', in Michael Ruse and Robert J. Richards (eds.), *The Cambridge Handbook of Evolutionary Ethics* (Cambridge: Cambridge University Press), pp. 202-209.

Clarke-Doane, Justin (2020), *Morality and Mathematics* (Oxford: Oxford University Press).

Clarke-Doane, Justin and Baras, Dan (2021), 'Modal Security', *Philosophy and Phenomenological Research* 102: 162-183.

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

Huemer, Michael (2001), *Skepticism and the Veil of Perception* (Lanham: Rowman & Littlefield).

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

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

Korman, Daniel Z. and Locke, Dustin (2020), 'Against Minimalist Responses to Moral Debunking Arguments', *Oxford Studies in Metaethics* 15: 309-332.

Korman, Daniel Z. and Locke, Dustin (2023), 'An Explanationist Account of Genealogical Defeat', *Philosophy and Phenomenological Research* 106: 176-195.

Noonan, Christopher (forthcoming), 'Evolutionary Debunking Arguments, Explanationism and Counterexamples to Modal Security', *Erkentnnis*.

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

Schechter, Joshua (2018), 'Is There a Reliability Challenge for Logic?', *Philosophical Issues 28, Philosophy of Logic and Inferential Reasoning*: 325-347.

Schoenfield, Miriam (2022), 'Deferring to Doubt', *Proceedings of the Aristotelian Society* 122: 269-290.

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

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

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.

Vavova, Katia (2015), 'Evolutionary Debunking of Moral Realism', *Philosophy Compass* 10: 104-116.

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 (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.

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