



# Why undermining evolutionary debunkers is not enough

Andrew del Rio<sup>1</sup>

Received: 4 January 2021 / Accepted: 11 March 2021 / Published online: 7 May 2021  
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

## Abstract

Denying the conclusion of a valid argument is not generally permissible if one suspends on one premise of the argument and believes the other premise(s). This can happen when one's only critique of an argument is to undermine one premise. There is incoherence there. Here I examine how this is relevant to the debate on evolutionary debunking of our moral knowledge. I argue that one significant line of response to the debunker is unsuccessful: merely undermining the debunker's empirical claim. It is not rational to respond this way and believe one has moral knowledge. First I present evidence that prominent critics of the debunking argument merely undermine the debunker's empirical claim. Then I argue for two premises: (1) merely undermining a premise can only justify a middling amount of doubt towards the premise and (2) we should have no more doubt about the conclusion of a valid argument than we do about the premises. Implications of the argument are explored.

**Keywords** Moral epistemology · Evolutionary debunking · Undermining defeaters · Undercutting defeaters · Suspension of judgment · Agnosticism

## 1 Introduction

Three characters—Theist, Atheist, and Agnostic—walk into a bar. The newspaper on the table reports chaos and tragic suffering around the world. Agnostic asks Theist a question. “Could God really co-exist with the amount of suffering in the world?” Upon reflection Theist responds “I really don't know. I guess I'm agnostic about whether God could co-exist with the amount of suffering in the world, but that doesn't shake my belief that God exists!”.

What's going on here? It seems that Theist is incoherent, but how? Theist doesn't go so far as to affirm a contradiction, nor to affirm propositions that entail a contradiction. Theist is in the following state:

---

✉ Andrew del Rio  
andrew.delrio@utexas.edu

<sup>1</sup> The University of Texas, Austin, TX, USA

<b>Believe</b>	(Suffering S obtains)
<b>Suspend</b>	(If S obtains then God does not exist)
<b>Believe</b>	(God does exist)

or perhaps:

<b>Very High Credence</b>	(Suffering S obtains)
<b>Middling Credence</b>	(If S obtains then God does not exist)
<b>Very High Credence</b>	(God does exist)

In the latter case the incoherence is probabilistic, in the former case it is harder to specify, but in both cases it looks like Theist is likely in an irrational state. This paper explores how a certain way of responding to the evolutionary debunking argument leaves us in the same position.

We can regiment the evolutionary debunking argument as follows<sup>1</sup>:

### Evolutionary Debunking argument

- (1) *Epistemological premise*: If our method for forming moral judgments is not truth-conducive or is accidentally truth-conducive, then none of our moral judgments amount to knowledge.
- (2) *Empirical premise*: Influenced as it is by evolution, our method for forming moral judgments is not truth-conducive or is accidentally truth-conducive.
- (3) None of our moral judgments amount to knowledge.

Section 2 differentiates three ways of responding to the debunker's argument: a constructivist way, a Moorean way, and an undermining way. The rest of the paper examines and critiques the undermining response. In Sect. 3 I defend the claim that our doxastic attitudes towards the premises of an argument rationally constrain our doxastic attitude towards the conclusion of the argument—specifically, we should have no more doubt about the conclusion of a valid argument than we do about the premises. In Sect. 4 I argue that merely undermining a premise can only justify an agnostic-like state or middling amount of doubt towards a premise. Section 5 applies the work of Sects 3, 4 to the evolutionary debunking argument showing that merely undermining the empirical premise does not rationally allow denying the debunker's conclusion. One who employs this mere undermining response ought to, by her own

<sup>1</sup> Morton (2018) gives a similar two premise—epistemological and empirical—regimentation. Here I use the non-accidentally truth-conducive condition, but as mentioned below (Sect. 5) my argument does not rest on formulating the debunker's argument with this particular condition for knowledge.

lights, be agnostic about whether she has moral knowledge. Section 6 explores the implications of my argument for our moral judgments.

## 2 Three strategies for saving moral knowledge

As just noted, evolutionary debunking arguments make an empirical claim and an epistemological claim. The empirical claim is that our evolutionary history has resulted in moral judgments that are accidentally correct if correct at all. The epistemological claim is that being accidentally correct is incompatible with knowledge. So we do not have moral knowledge. There are multiple strategies intended to rescue moral knowledge from the debunker. Here I will contrast three.

The first strategy is the *Constructivist Strategy*. The constructivist thinks the best response to the debunkers' challenge is to deny that moral facts are true independent of what we (would, in reflective equilibrium,) think or feel about them (Street, 2006, 2008). What makes it wrong to indiscriminately kill people is roughly that we (would, in reflective equilibrium,) think it is wrong and/or feel negatively toward that activity. The fact that our attitudes are correct is not accidental after all because the correctness of our attitudes is grounded in our having just those attitudes. So the debunker's empirical premise is false. It is false that our evolutionary history has resulted in moral judgments that are accidentally correct if correct at all. We construct the moral facts and since it is of their nature to be constructed by us, our evolutionary history is not a threat to our moral knowledge.<sup>2</sup>

The second strategy is the *Moorean Strategy*. Moore (1939) famously responded to the external world skeptic by raising his hands and saying "Here is one hand and here is another." Therefore "two human hands exist at this moment." The Moorean Strategy involves starting with some point that is putatively up for debate and using that to argue against one's opponent. So the external world skeptic thinks it is up for debate whether Moore's hand-wavy gesture is reason to believe hands exist at all. As the skeptic sees it, it is an open question whether there is a hand. Likewise, Moorean responses to the evolutionary debunker start their response with something that is putatively up for debate, for example: survival-promoting behavior is good (Enoch, 2011); society's basic needs are best served by adopting moral standards (Copp, 2007, 2008); life-affirming moral beliefs are reliable (Dogramaci, 2017). From there they argue that the debunker's empirical premise is false. It is false that our evolutionary history has resulted in moral judgments that are objectionably accidental. These responses rely on substantive moral claims, either evaluative or normative, to respond to an argument that we have no substantive moral knowledge.

---

<sup>2</sup> Berker (2014) argues that the constructivist is no better situated to respond to the argument than the moral realist.

A number of critiques are Moorean,<sup>3</sup> but as Kelly (2008) has noted, “many would hold that there is something deeply objectionable about dismissing skeptical theories on such Moorean grounds.” Indeed, in his estimation “perhaps most philosophers—or at least a significant plurality,” take issue with Moorean responses. This leads to another strategy.

The third strategy is the *Mere Undermining Strategy*. Call those who employ this strategy Mere Underminers. Mere Underminers focus on the weak support for the debunker’s empirical premise. Rather than arguing that the empirical premise is false as the Constructivist and Mooreans do, Mere Underminers argue that the premise is unsubstantiated.

On one interpretation of Shafer-Landau, FitzPatrick, and Vavova, they claim that we can undermine the debunker’s empirical premise and that undermining the premise is enough to reject the debunker’s conclusion. Following are some passages that support this interpretation. Keep in mind that each of these philosophers rejects the debunker’s conclusion that we lack moral knowledge.

Shafer-Landau (2012) seems to state the position explicitly: “Until such a time as we can mount a successful formal strategy for vindicating the reliability of our moral faculties, moral realists will have to focus directly on trying to undermine at least one premise of the Darwinian Argument.... All the realist needs is to show that [the premise] is inadequately supported.”<sup>4</sup>

FitzPatrick represents his own argument as one that “does *undermine* the more ambitious attacks on realism.” (emphasis added, 2014) He says “the key to *defusing* the debunking arguments lies in distinguishing between the real science associated with evolutionary theory and the debunker’s overreaching explanatory claims.” (emphasis added, 2014) And “it’s just not true that ‘the empirical data concerning human evolution’ support anything approaching [the] premise.... (i) the science *doesn’t tell us* how pervasive the evolutionary influence on our current beliefs is, and (ii) even where there is such influence we have *no reason* to suppose it’s vastly more incidentally distorting than it is incidentally supporting.” (emphasis added, 2015) Like Shafer-Landau, FitzPatrick’s attacks seem to merely assert that a premise of the debunker’s argument is inadequately supported.

Vavova’s overall strategy, as I understand it, creates a dilemma for the debunker. Either we can adopt the Moorean strategy above or we cannot rely on any of our moral judgments. On the first horn of the dilemma we can adopt the Moorean strategy and thereby have an adequate response to the debunker. But Vavova (2015) says that the Moorean Strategy seems “more or less directly question begging.” If we cannot adopt the Moorean Strategy, then we cannot rely on any of our moral judgments.

<sup>3</sup> See also White (2010), Wielenberg (2010) and Schafer (2010). Each Moorean response is much more sophisticated and nuanced than I have indicated here, but they share the idea that we can appeal to some substantive moral judgments in our response to the debunker.

<sup>4</sup> Shafer-Landau (2012) does say that moral beliefs are “strongly presumptively warranted.” This suggests that Shafer-Landau may embrace a Moorean strategy.

...morality could be about anything. And if morality could be about anything, then we have no idea what morality is about. So we *have no reason* to think that the attitude-independent truths and the adaptive beliefs don't overlap. But without that, we have no sense of what the chances are that we are mistaken. Therefore, we cannot get to the conclusion that we probably are mistaken. (emphasis added, 2014)

That is, on the second horn of the dilemma we lose reason to think the empirical premise is true—the premise is undermined.<sup>5</sup>

All three authors seem to claim that the crucial issue with the debunker's argument is that the empirical premise is unjustified rather than that it is false. They may indeed think it is false, but the reasons they give for not accepting it are merely undermining. Perhaps this is a wrong interpretation of these three authors. In any case I think the position is of independent interest and a tempting one. Rather than establish that the debunker's premise is false, the Mere Underminer is content to establish that we shouldn't believe the debunker's premise.

As we will see below, however, not believing is not enough to permit denying the debunker's conclusion and so the Mere Undermining Strategy does not save moral knowledge. The Mere Undermining Strategy partially deflects the debunker's blow—it shows that the debunker hasn't yet given us sufficient reason to think we don't have moral knowledge—but it conflicts with the position that we want to hold—that we have moral knowledge.

### 3 Doxastic constraints in valid arguments

To show that the Mere Undermining Strategy conflicts with a claim to moral knowledge I will first argue that the collective amount of doubt we have in the premises of a valid argument sets a limit on the amount of doubt we can rationally have in the conclusion. I will work this idea out first in terms of traditional doxastic states (belief, suspension, disbelief) and then in terms of credences. This will be relevant because, as I will discuss in Sect. 4, undermining justifies some doubt in the premises but not enough for a coherent denial of the conclusion.

When evaluating an argument, it is not uncommon to take a two step procedure: check for validity and check for true premises. If the argument fails at either step, then it is unsound and can be rejected. This method is fine as far as it goes, but if the second step is practiced without sensitivity to which specific doxastic attitudes we take towards the premises, we may be lead to ignore an argument unjustifiably.

Say an argument is valid, or can be made so, and we want to check for the truth of the premises. Here traditional doxastic attitudes are helpful. You can believe a premise, disbelieve it (logically equivalent to believing the negation), or suspend judgment. Following Friedman (2013), I am thinking of suspension of judgment as “an attitude that expresses or represents or just is one's neutrality or indecision about

<sup>5</sup> See also Vavova ([forthcoming](#)).

which of  $p$ ,  $\neg p$  is true.” Accordingly, suspension is different from having no attitude towards a proposition. Barring a lack of the relevant concepts, once we have considered a premise to check for its truth, we would (almost without exception<sup>6</sup>) have some doxastic attitude towards the premise. If you rationally disbelieve a premise, then you are rational to ignore the argument—that argument does not provide reason for you to revise your belief about the conclusion.<sup>7</sup> But what if a premise is one that you should neither believe nor disbelieve?

We might think that if we do not believe every premise of the argument, then the argument has failed the truth test (as far as we are concerned) and therefore allows us to reject the conclusion. This would be a mistake. We need to be sensitive to the fact that an argument can fail to pass the truth-test in two ways: when we believe it has a false premise (disbelieve a premise) or when we take an attitude of suspension toward a premise. When it fails in this latter way, we cannot just reject the conclusion.

Consider the following example:

### Coin game argument

- (1) If the coin flip is heads, then I win.
- (2) The coin flip is heads.
- (3) I win.

Premise 1 lays out the rules of the game we are playing. What should I think of this argument before the coin is flipped? Clearly I should not believe (2) but I also should not disbelieve it. Rather I should suspend judgment about (2). What should I believe about the conclusion? I should neither believe nor disbelieve (3). Rather, in the absence of any other evidence that I win, I should suspend judgment about (3).<sup>8</sup> Is affirming (3) rationally compatible with suspending on (2)? It could be. There could be other relevant evidence that supports the conclusion, for example, if it is true in this (admittedly boring) game, that on tails you also win. But, believing (1), is denying (3) rationally compatible with suspending on (2)? It is not and could not be. Any evidence against (3) must also lower your confidence in some premise. Say I have justification for disbelieving (3)—the oracle told me I will not win the game—then I would know that the coin will not come up heads. The rules of the game entail this.<sup>9</sup> Failing to believe some premise does not mean that I can reject the conclusion. Why is this? We might say there is a general principle that can be stated as follows:

<sup>6</sup> Friedman (2013) Sect. 3.1 provides some exceptions.

<sup>7</sup> I am assuming here that the argument is minimal in that all of its premises are used to deduce the conclusion. If it has extra premises and you disbelieve one of those while accepting the rest, then that argument still provides reason for you to accept the conclusion.

<sup>8</sup> And the same would be true if I suspended judgment about (1) instead of (2) (imagine I see the coin flip is heads, but the rule of the game about which of heads/tails is the winner is sealed in an envelope).

<sup>9</sup> I am assuming *closure* of propositional justification across known entailment but I am not assuming *transmission* of propositional justification across known entailment. Though closure of justification is not without its critics (see Avnur (2012)), it is widely endorsed.

**Argument Rule**

For a valid argument, it is irrational to suspend judgment about one premise, believe the rest of the premises (if any), and disbelieve the conclusion.<sup>10</sup>

The Argument Rule seems plausible when we are talking about sufficiently short arguments. We could generate many confirming instances of this rule by taking a short argument and ascribing the relevant doxastic states to a subject. It will be intuitive over and over again that the subject is somehow irrational or incoherent if they suspend on one premise, believe the other (if any), and disbelieve the conclusion. In the Coin Game the intuition is generated when we consider a person who believes they will lose the game (disbelieves the conclusion) but suspends on whether the coin flip will be heads while believing that if it is heads then they win. Something is wrong with that. Likewise in my opening example Theist believes God exists while believing suffering S obtains but suspending on the conditional if S obtains, then God does not exist. This violates the Argument Rule for the following argument:

**Atheist argument**

- (1) If suffering S obtains, then God does not exist.
- (2) S obtains.
- (3) God does not exist.

The Argument Rule, however, seems to have some exceptions. Reflection on Kyburg's (1961) Lottery Paradox and Makinson's (1965) Preface Paradox should show that the Argument Rule is not universal. In those cases it would seem you can believe all the premises of a very long argument but also disbelieve the conclusion without transgressing any rational constraints. And of course if you can believe all the premises and disbelieve the conclusion, then it is no surprise that it is rational to believe all the premises, *except for one* which you suspend on, and disbelieve the conclusion. Nevertheless, we need not rely on a principle and infer that something is wrong in these short arguments. The reason the principle looked appealing to begin with is because we can see that something is wrong in the specific cases cited.<sup>11</sup>

So far we have been thinking through these arguments with the traditional doxastic attitudes in mind: belief, suspension of judgment, and disbelief. And we have seen that we need to be sensitive to how a premise is not believed. The force of an argument will be different when a premise is disbelieved than when one suspends

<sup>10</sup> Rosa (2020) explores and defends rational requirements like the Argument Rule for arguments where one suspends on a premise. McGrath (forthcoming) explicitly endorses an instance of the rule: "if I am agnostic about whether supernatural beings exist and I know that if God exists there are supernatural beings, then I cannot, without irrationality, believe that God exists."

<sup>11</sup> Rosa (2019) argues that while we cannot affirm a general principle that it is irrational to suspend judgment about one premise, believe the rest of the premises (if any), and disbelieve the conclusion, we can affirm a general principle that one has a reason not to be in that state. I think one has a defeasible defeater for the set (or at least one) of those doxastic attitudes. A defeater is different than a reason, but I cannot develop this theory here.

on the premise. The Argument Rule helps us increase our doxastic sensitivity when evaluating arguments, but it is not universally applicable. Formal epistemology provides more resources that can explain why the Argument Rule coheres with our intuitions about short arguments but not to our intuitions about very long arguments.

We can model our doxastic states with a credence function ascribing a number from 0 to 1 to each proposition we have taken an attitude towards. Most of our beliefs will not be modeled with a 0 or a 1.<sup>12</sup> We are usually less than maximally confident. In these cases even though one may have a high credence in a premise, one will still have some doubt about it. So even though I am highly confident that I will go home tonight, I am aware that circumstances could arise where I do not end up going home tonight. I ascribe the lot of these scenarios very little credence, but insofar as I give them any credence then I have some degree of uncertainty about whether I will go home tonight. Call the degree to which your credence falls short of certainty, your uncertainty. With these notions of credence and uncertainty in mind, we can now see why the Argument Rule looks good in the case of short arguments but looks bad in the case of long arguments like the paradoxes cited above. The sum total of your uncertainties for all the premises of an argument are like a maximum allocation of uncertainty that you can spend in the conclusion. You cannot have more uncertainty about the conclusion than is allotted by the sum of the uncertainties you have in the premises. In the case of short arguments, a small bit of uncertainty in each of the few premises does not allow for a lot of uncertainty to accumulate. But in the case of much longer arguments like the Lottery and Preface cases, a small bit of uncertainty in each of the many premises adds up to allow for a lot of uncertainty in the conclusion.

Formally it looks like this. Following Adams and Levine (1975), the *uncertainty* of a premise is equal to one's credence in a premise subtracted from 1, or  $U(p) = 1 - C(p)$ , where  $U(p)$  is your uncertainty in  $p$  and  $C(p)$  is your credence in  $p$ . So if you have a 0.7 credence in a premise, your uncertainty in that premise is 0.3. A constraint on the relationship between one's uncertainties in the premises and one's uncertainty in the conclusion of a valid argument can be stated with the Uncertainty Rule:

**Uncertainty Rule** Where  $p_1$ – $p_n$  are an arguments premises,  $c$  is the conclusion, and the argument is deductively valid,  $U(c) \leq U(p_1) + \dots + U(p_n)$ .<sup>13</sup>

In other words, your uncertainty in the conclusion of a valid argument must be less than or equal to the sum of your uncertainties in the premises of that argument. This Uncertainty Rule gives us a precise lower bound for the probability of the conclusion

<sup>12</sup> There are hard questions here about the relationship between belief/suspension and credence. For the purposes of this paper we can assume that belief, suspension, and disbelief are associated with high, middling, and low credence. Ultimately, however, I find that view unsatisfactory. I argue that there are no good credal accounts of belief, suspension, and disbelief in del Rio ([manuscript](#)).

<sup>13</sup> See also Adams (1998), 31–53. These premises need not be probabilistically independent.



of the argument, given the probabilities of the premises.<sup>14</sup> For example, according to the Uncertainty Rule, someone who places a credence of 1.0 in the first premise of Coin Game Argument and a credence of 0.5 in the second premise is probabilistically incoherent if they have a credence less than 0.5 in the conclusion.

We have seen in this section that we must be sensitive to the doxastic attitudes we have towards the premises of arguments. If we want to reject the conclusion of an argument, then we need to pay attention to more than just whether we believe the premises. Rather we need to pay attention to our degree of confidence in the premises. Even if we do not believe all the premises, ascribing some credence to the premises can place a constraint on how low our confidence can be in the conclusion. In a short two-premise argument, if we have a high degree of confidence in one premise and a middle credence in the other premise, we will be probabilistically incoherent if we take a low credence in the conclusion.<sup>15</sup> This is precisely the state which Theist was in and in which we find ourselves if we merely undermine the debunker.

#### 4 Doxastic results of mere undermining

Mere Underminers find themselves in this state because mere undermining only justifies an agnostic-like attitude. An *undermining defeater* shows that some evidence E that (purportedly) justified believing some proposition  $p$  does not in fact (sufficiently) support believing  $p$ . Merely undermining a premise does not give you reason to believe its negation, but it does give you reason not to believe that premise. As Sturgeon (2014) argues, undermining defeaters “do not make for reasons to believe.” Bergmann (2006) says this is the differentiating feature between rebutting and undercutting defeaters—undercutting (or undermining) defeaters do not provide reason for thinking a defeated claim is false while rebutting defeaters do. Therefore, on its own, undermining supports taking a neutral doxastic attitude. “The under[ining] defeater gives me reason to be an agnostic.” (Bergmann, 2005).

Consider the following proposition:

**Even Stars**      There are an even number of stars.

<sup>14</sup> Another way of getting a similar result is to employ the Entailment Rule:  $(A \models B) \rightarrow P(A) \leq P(B)$ . Since the premises of a valid argument entail the conclusion, the conclusion cannot have a lower probability than the conjunction of the premises. This method can create a higher lower bound than the Uncertainty Rule, and so is more constraining. However it is less user friendly since the probability of the conjunction of the premises must be calculated, taking into account whether or not the premises are probabilistically independent. The Uncertainty Rule will always provide a lower bound that is equal to or less than the lower bound provided by the Entailment Rule. My argument is given using the weaker constraint.

<sup>15</sup> For the purposes of this paper I will assume that identified probabilistic incoherence is irrational. Details for a view like this are worked out in Dogramaci (2018).

Say I believe this because a scientist told me that each star has one and only one twin star somewhere in the universe. This belief of mine can be undermined if the scientist tells me that she was just pulling my leg about the twin stars. This undermines my belief in Even Stars—Even Stars is not sufficiently supported by my total evidence. I can now see that the scientist’s original testimony is not sufficient support for Even Stars. But notice that undermining a proposition does not in any way suggest that it is false. I should not believe that there are an odd number of stars! On their own, undermining defeaters push us to suspension of judgment or a middle credence, not to disbelief or a low credence.

One might object that Even Stars stacks the deck against the Mere Underminer. Of course we ought to suspend judgment about Even Stars. There are either an even or an odd number and we have no evidence one way or the other. That is just a paradigmatic case for suspension. It is true that there are just an even or odd number. Two options. But the same two options exist for every proposition; it’s true or its negation is true (Odd Stars is just the negation of Even Stars).

The reason Even Stars is a fitting example is because it is clearly a case where *no rebutting evidence is available*. We want that to be the case because we are trying to isolate the force of mere undermining. Mere undermining of  $p$  only removes reason for believing  $p$ . It does not provide reason to disbelieve  $p$ . It is true that removing epistemic reason for a subject  $S$  to believe  $p$  might create a situation where disbelieving  $p$  is epistemically rational for  $S$ . But that is because of antecedent epistemic reason for disbelieving  $p$ —that is rebutting evidence. None of these cases will be cases of *mere* undermining.

For example, consider the case of Magic Mike.

**Magic Mike** a magician, Mike, selects from four pieces of paper each designating a suit—hearts, diamonds, spades, clubs. He places one piece of paper in an envelope, concealing which one it is. He then shuffles a deck of cards and asks me to pick a card. Mike says the envelope contains the suit of the card I will pick. I draw the six of hearts. But before I look in the envelope, I find out by some trustworthy means that Mike is not a magician at all and is only pretending to perform a trick.

For the sake of argument, let’s assume that trusting Mike’s testimony while I thought he was a magician yields justified beliefs. Then this is a case where I gained a justified belief in Hearts.

**Hearts** The content of the envelope says “hearts.”

My belief was then undermined. But Hearts isn’t *merely* undermined. I have some evidence that Hearts is false, for now I know that Mike picked a suit randomly and the chance of randomly getting the matching suit is  $1/4$ . I have rebutting evidence that is crucially relevant in rejecting Hearts. Perhaps disbelief that Hearts is not in order. Nevertheless, this case is not fitting for illustrating the consequences of undermining by itself. The only fitting cases for understanding the force of mere

undermining are cases where no rebutting evidence is available. All such cases are like Even Stars. Hence Sturgeon's and Bergmann's claims above that undermining only gives reason for agnosticism.

In this section I have argued that mere undermining of  $p$  requires an agnostic attitude towards  $p$ . In the previous section I argued that taking an agnostic attitude towards one premise of a valid two-premise argument (while endorsing the other premise) is not rationally compatible with denying the conclusion. We are now ready to apply these points to the Mere Underminer of the evolutionary debunking argument.

## 5 Doxastic constraints on mere underminers of the debunking argument

Again the evolutionary debunking argument can be stated as follows:

### Evolutionary Debunking argument

- (1) *Epistemological premise*: If our method for forming moral judgments is not truth-conducive or is accidentally truth-conducive, then none of our moral judgments amount to knowledge.
- (2) *Empirical premise*: Our method for forming moral judgments is not truth-conducive or is accidentally truth-conducive.
- (3) None of our moral judgments amount to knowledge.

What makes the argument *evolutionary* is the evolutionary considerations offered in support of premise 2.<sup>16</sup> What makes it a *debunking* argument is that it asserts that some condition for knowledge is not satisfied. I think the argument is best built on a non-accidentally truth-conducive condition as defended by Yamada (2011), but my charge against the Mere Underminer is consistent with debunking arguments that use other conditions for knowledge, such as a reliability condition or a proper function condition.<sup>17</sup>

The argument is valid. So however much we want to doubt the conclusion, we need at least that much doubt to arise in the premises. In other words, if we want to disbelieve or have a very low credence in the conclusion, we need sufficient uncertainty in the premises. Merely failing to believe a premise does not get us off the hook of this argument.

Now we might just disbelieve the empirical premise because we believe moral claims that entail the conclusion is false and we believe the epistemological premise is true. That would be to employ a Moorean strategy. Everything I've said here is compatible with the Moorean strategy succeeding. I am evaluating a kind of strategy that is supposed to satisfy those with anti-Moorean sentiments. The Mere

<sup>16</sup> My purposes do not require me to go into these consideration here. See Joyce (2001, 2005), Mackie (1977), Street (2006) and Vavova (2015).

<sup>17</sup> See also Schafer (2014) and Setiya (2012) for defense of a non-accidentality condition.

Undermining Strategy is one such strategy. There are others. For example, one might argue on theistic grounds that the empirical premise is false. If we know God exists and foreordained the evolutionary process to produce in us moral knowledge, then the empirical evidence from evolution is undermined and the premise is rebutted. That is, we might take their to be independent and sufficiently strong reasons for thinking God guarantees that our moral beliefs are formed in a non-accidentally truth-conducive way. Such anti-Moorean strategies may succeed, though Street (2014) argues they do not. My thesis is that the Mere Undermining strategy cannot succeed.

Grant that the Mere Underminer does undermine the premise of the debunker's argument. Undermining a premise does not give you reason to believe its negation. Therefore, if one's only critique of the argument is to undermine a premise, then one should suspend judgment or take a middle credence on that premise. Taking either of these attitudes towards the empirical premise (while granting the epistemological premise) requires not also believing we have moral knowledge. So the Mere Underminer is irrational if she believes she has moral knowledge. Put another way, Mere Underminers cannot rationally deny the debunker's conclusion—they cannot believe they have moral knowledge and be rational.

## 6 Further implications

One might object: "What I really care about are my moral judgments, not whether I know I have moral knowledge." This objection accepts the conclusion that the Mere Underminer can't rationally believe she has moral knowledge, but asks why we should care about rationally believing that we have moral knowledge. After all, my argument is not that particular moral judgments, like "slavery is wrong," are unknown, much less unjustified. Yet the argument may well have this consequence.

The first thing to be said is that the Mere Underminer will have to undergo some drastic belief revision about what she knows—presuming that she takes herself to know all kinds of moral facts. For it isn't just the general claim of moral knowledge that is rationally incompatible with mere undermining. Say the Mere Underminer takes a credence of 0.5 in the empirical premise and accordingly 0.5 in the conclusion that none of our moral judgments amount to knowledge, as I argue is required. Can they still think they know particular moral judgments, like slavery is wrong? No. Consider the following argument:

### Particular moral judgment argument

- (1) If none of my moral judgments amount to knowledge, then my moral judgment that slavery is wrong doesn't amount to knowledge.
- (2) None of my moral judgments amount to knowledge.
- (3) My moral judgment that slavery is wrong doesn't amount to knowledge.

The same line of reasoning that led us to take a middle credence about the conclusion of the evolutionary debunking argument will lead us to a middle credence in

the conclusion of this argument. (1) is necessarily true. Say we assign it a credence of 1. (2) is the conclusion of the debunking argument which we have now assigned a credence of 0.5. Therefore there is a total sum of 0.5 for the uncertainty of the premises. So we can have less than or equal to 0.5 uncertainty in the conclusion of this argument. In terms of credences, we can have a credence of 0.5 or greater in this conclusion. Anything less is probabilistically incoherent. Agnosticism about whether one knows slavery is wrong is presumably undesirable.

The second thing to be said is that even the Mere Underminer's first order moral judgments may be in jeopardy. There are at least two ways of making things worse for the Mere Underminer. The first way employs a principle that says suspension of judgment towards whether you know  $p$ , makes it irrational to believe  $p$ . The second way employs a principle that says suspension of judgment about whether your belief that  $p$  was formed in a non-accidentally truth-conducive way, makes it irrational to believe  $p$ . Either way the Mere Underminer's moral judgments will be irrational.

Huemer (2007, 2011) defends what he calls the Metacoherence Requirement:

**Metacoherence requirement (MR)**      Categorically believing that  $p$  commits one, on reflection, to the view that one knows that  $p$ .

This principle states that rationality demands a coherence between our first order beliefs and our higher order beliefs about whether the first order beliefs amount to knowledge. If we believe that it is raining outside and we reflect on whether we know it, we must conclude that we do or else stop believing that it is raining. Huemer (2011) says "If one believes that  $p$ , and one either denies or withholds that one knows that  $p$ , then, according to MR, one exhibits a sort of irrationality in virtue of the clash between one's two attitudes." The primary reason for this comes from what Huemer calls Moore-paradoxical sentences—sentences like 'It is true that it's raining, but I do not believe it is true that it's raining' and 'It is true that it's raining, but I do not know that it is true that it's raining.' There is, according to Huemer, something irrational about even *thinking* these things. Linguistic conventions and rules of assertion cannot explain this fact, but the Metacoherence Requirement does. We can restate the principle like this:

**MR\***      If, upon reflection, one should not believe that one's belief  $p$  amounts to knowledge, then one should not believe  $p$ .

But recall that we have concluded above that the Mere Underminer should not, upon reflection, believe that her moral judgment—call it  $m$ —amounts to knowledge. Therefore, according to MR\*, the Mere Underminer should not believe  $m$ . The Mere Underminer should neither believe that she knows slavery is wrong, nor that slavery is wrong.<sup>18</sup>

<sup>18</sup> The same result looks like it comes from the knowledge norm of belief: one ought only believe  $p$  if one's belief that  $p$  would amount to knowledge (Williamson, 2000). Since the Mere Underminer must believe  $m$  does not amount to knowledge, accepting the knowledge norm of belief means that she should not, by her own lights, believe  $m$ .

Bergmann (2005) gets to this conclusion directly from the Mere Underminer's suspension on the empirical premise. Bergmann writes:

If you are considering whether the actual basis of your belief that  $p$  is indicative of  $p$ 's truth and you find yourself resisting the belief that it is (because you have considered the matter and you have no idea whether it supports  $p$  or not), that seems to undercut your justification for believing  $p$  in the same way as if you believed outright that the actual basis for your belief that  $p$  did not indicate  $p$ 's truth.<sup>19</sup>

He motivates the point with a case. Say you find yourself in a factory looking through a window with a narrow field of view at some widgets traveling down a conveyor belt. They look red to you and you form the belief <The widgets are red>. Then another spectator asks you if the widgets are red or if there is a red light shining on them. You consider the higher-order question of whether you formed your belief in a reliable way. Having no idea whether the higher-order proposition is true, you suspend judgment on the matter. This also gives you, according to Bergmann, a defeater for your belief that the widgets are red. You should suspend judgment about whether the widgets are red.

But this seems to be just the sort of case we have with the evolutionary explanation of our moral judgments. We are asked to reflect on whether the actual basis for a moral belief  $m$  is indicative of  $m$ 's truth. The Mere Underminer must suspend judgment on the matter. But that defeats her justification for believing  $m$  at all. Again, the Mere Underminer should neither believe that she knows  $m$ , nor believe  $m$ .<sup>20</sup>

So if the argument of this paper is correct, Mere Undermining requires a substantive skepticism about whether we know particular moral judgments, and if the extended arguments of Huemer or Bergmann are correct, then skepticism about each of our particular moral judgments is required as well.

## 7 Concluding remarks

In closing, I want to emphasize the limited scope of this argument but also its broader implications. I am not arguing that the debunker is successful, nor that the debunker is unsuccessful. I am not endorsing any particular way of responding to the argument, or critiquing the Constructivist or the Moorean. I am also not arguing that Shafer-Landau, FitzPatrick, and Vavova are in fact Mere Underminers. They may well employ some rebutting evidence in their ultimate evaluation of the debunker, in addition to the undermining arguments in their papers. I am, however, ruling out Mere Undermining. Undermining is still valuable, but it must be supplemented with some other evidence, empirical or a priori, for the reliability of our moral judgments.

<sup>19</sup> Bergmann (2005, 426).

<sup>20</sup> See Hazlett (2012) and White (2010) for critique of these positions.

The argument also has wider implications. What I have said here applies across the board. A general *undermining fallacy* is committed when someone denies (disbelieves) the conclusion of an argument while their only critique is that one premise of the argument is unsubstantiated or inadequately supported. I suspect this situation is not uncommon in many philosophical disputes. If we find ourselves in that mere undermining position with respect to the premises of an argument, we might just have to be more skeptical than we would like or grant that we are doing something not easily distinguished from question begging. Undermining, by itself, is not enough.

**Acknowledgements** Thanks to the audience at the World Congress of Philosophy in Beijing, China for discussion of my argument. I'm grateful for many conversations about evolutionary debunking with Daniel Eaton, Steven Gubka, Derek Haderlie, Amelia Kahn, and Justin Morton. Thanks to Nevin Climenhaga, Jonathan Dancy, Josh Dever, Liz Jackson, Cory Juhl, Miriam Schoenfield, Katia Vavova and two anonymous referees for critical discussions. Special thanks are due to Sinan Dogramaci and David Sosa for feedback on every stage of this paper's development.

## References

- Adams, E. (1998). *A primer of probability logic*. . Csl Publications.
- Adams, E., & Levine, H. (1975). On the uncertainties transmitted from premises to conclusions in deductive inferences. *Synthese*, 30, 429–460.
- Avnir, Y. (2012). Closure reconsidered. *Philosophers' Imprint*, 12, 1–16.
- Bergmann, M. (2006). *Justification without awareness: A defense of epistemic externalism*. Oxford University Press.
- Bergmann, M. (2005). Defeaters and higher-level requirements. *Philosophical Quarterly*, 55, 419–436.
- Berker, S. (2014). Does evolutionary psychology show that normativity is mind-dependent? In J. D'Arms & D. Jacobson (Eds.), *Moral psychology and human agency: Philosophical essays on the science of ethics*. Oxford University Press.
- Copp, D. (2007). *Morality in a natural world: Selected essays in metaethics*. . Cambridge: Cambridge University Press.
- Copp, D. (2008). Darwinian skepticism about moral realism. *Philosophical Issues*, 18, 186–206.
- del Rio, A. (manuscript). Suspending belief in credal accounts.
- Dogramaci, S. (2017). Explaining our moral reliability. *Pacific Philosophical Quarterly*, 98, 71–86.
- Dogramaci, S. (2018). Rational credence through reasoning. *Philosophers' Imprint*, 18, 1–25.
- Enoch, D. (2011). *Taking morality seriously: A defense of robust realism*. . Oxford University Press.
- FitzPatrick, W. (2014). Why there is no Darwinian dilemma for ethical realism. In Michael Bergmann & Patrick Kain (Eds.), *Challenges to moral and religious belief: disagreement and evolution*. Oxford University Press.
- FitzPatrick, W. (2015). Debunking evolutionary debunking of ethical realism. *Philosophical Studies*, 172, 883–904.
- Friedman, J. (2013). Suspended judgment. *Philosophical Studies*, 162, 165–181.
- Hazlett, A. (2012). Higher-order epistemic attitudes and intellectual humility. *Episteme*, 9, 205–223.
- Huemer, M. (2007). Moore's paradox and the norm of belief. In S. Nuccetelli & G. Seay (Eds.), *Themes from G.E. Moore: New essays in epistemology and ethics*. Clarendon Press.
- Huemer, M. (2011). The puzzle of metacoherence. *Philosophy and Phenomenological Research*, 82, 1–21.
- Joyce, R. (2001). *The myth of morality*. . Cambridge University Press.
- Joyce, R. (2005). *The Evolution of Morality*. Bradford.
- Kelly, T. (2008). Common sense as evidence: Against revisionary ontology and skepticism. *Midwest Studies in Philosophy*, 32, 53–78.
- Kyburg, H. (1961). *Probability and the logic of rational belief*. Wesleyan University Press.

- Mackie, J. L. (1977). *Ethics: Inventing right and wrong*. Penguin Books.
- Makinson, D. C. (1965). The paradox of the preface. *Analysis*, 25, 205–207.
- McGrath, M. (forthcoming). Being neutral: suspension of judgement, agnosticism and inquiry. *Noûs*.
- Moore, G. E. (1939). Proof of an external world. *Proceedings of the British Academy*, 25, 273–300.
- Morton, J. (2018). When do replies to the evolutionary debunking argument beg the question? *Australasian Journal of Philosophy*, 97, 265–280.
- Rosa, L. (2019). Logical principles of agnosticism. *Erkenntnis*, 84, 1263–1283.
- Rosa, L. (2020). Rational requirements for suspended judgment. *Philosophical Studies*, pp. 1–22.
- Schafer, K. (2010). Evolution and normative scepticism. *Australasian Journal of Philosophy*, 88, 471–488.
- Schafer, K. (2014). Knowledge and two forms of non-accidental truth. *Philosophy and Phenomenological Research*, 89, 373–393.
- Setiya, K. (2012). *Knowing right from wrong*. Oxford University Press.
- Shafer-Landau, R. (2012). Evolutionary debunking, moral realism and moral knowledge. *Journal of Ethics and Social Philosophy*, 7, 1–38.
- Street, S. (2006). A Darwinian dilemma for realist theories of value. *Philosophical Studies*, 127, 109–166.
- Street, S. (2008). Constructivism about reasons. *Oxford Studies in Metaethics*, 3, 207–246.
- Street, S. (2014). If everything happens for a reason, then we don't know what reasons are: Why the price of theism is normative skepticism. In M. Bergmann, & P. Kain, (eds.), *Challenges to moral and religious belief: Disagreement and evolution*. Oxford University Press.
- Sturgeon, S. (2014). Pollock on defeasible reasons. *Philosophical Studies*, 169, 105–118.
- Vavova, K. (2014). Debunking evolutionary Debunking. *Oxford Studies in Metaethics*, 9, 76–101.
- Vavova, K. (2015). Evolutionary Debunking of moral realism. *Philosophy Compass*, 10, 104–116.
- Vavova, K. (forthcoming). The limits of rational belief revision: A dilemma for the darwinian debunker. *Noûs*.
- White, R. (2010). You just believe that because. *Philosophical Perspectives*, 24, 573–615.
- Wielenberg, E. (2010). On the evolutionary debunking of morality. *Ethics*, 120, 441–464.
- Williamson, T. (2000). *Knowledge and its Limits*. Oxford University Press.
- Yamada, M. (2011). Getting It Right By Accident. *Philosophy and Phenomenological Research*, 83, 72–105.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.