

Bond University
Research Repository



Three Problems for the Evolutionary Debunking Argument

Davis, Oscar; Cox, Damian

Published in:
Ratio

DOI:
[10.1111/rati.12348](https://doi.org/10.1111/rati.12348)

Licence:
CC BY

[Link to output in Bond University research repository.](#)

Recommended citation(APA):
Davis, O., & Cox, D. (2022). Three Problems for the Evolutionary Debunking Argument. *Ratio*. Advance online publication. <https://doi.org/10.1111/rati.12348>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

For more information, or if you believe that this document breaches copyright, please contact the Bond University research repository coordinator.

Three problems for the evolutionary debunking argument

Oscar Davis  | Damian Cox

Faculty of Society and Design, Bond University, Gold Coast, Queensland, Australia

Correspondence

Oscar Davis, Faculty of Society and Design, Bond University, 14 University Drive, Gold Coast, QLD 4226, Australia.
Email: odavis@bond.edu.au

Funding information

Bond University

Abstract

In attempting to debunk moral realism through an appeal to evolutionary facts, debunkers face a series of problems, which we label the problems of scope, corrosiveness, and post-hoc justification. To overcome these problems, debunkers must assume certain metaphysical or epistemological positions, or otherwise pre-establish them. In doing so, they must assume or pre-establish the very conclusion they seek in advancing the argument. This means that such debunking arguments either beg the question against the moral realist or are undermined as standalone metaethical arguments.

KEYWORDS

evolutionary debunking, justification, metaethics, mind-independence, realism

1 | INTRODUCTION

Attempts to debunk evolutionary debunking arguments are not new. Nor is it new to demonstrate how they smuggle *old wine into new bottles* in the form of metaphysical and epistemological assumptions.¹ What we demonstrate here is the deeper implication of these claims, namely that evolutionary debunking is undermined as a standalone metaethical argument strategy.

We begin first by setting out the deep structure of evolutionary debunking argument strategies and explaining what we mean when we say that metaphysical and epistemological conclusions must be assumed or preestablished in order for them to succeed. We will then illustrate how these conclusions are assumed in the debunking literature by examining a particular debate in this literature, the demarcation challenge.

¹See for example, Clarke-Doane (2012), Das (2016), Klenk (2017), Schafer (2017), Vavova (2014, 2015), and Wielenberg (2010).

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Ratio* published by John Wiley & Sons Ltd.

2 | EVOLUTIONARY DEBUNKING ARGUMENTS

Evolutionary debunking arguments we discuss in this paper represent a metaethical challenge to moral realism. They seek to undermine the plausibility of the realist's claim that our moral sense could have evolved to track mind-independent moral truths. The argument is undermining in nature as opposed to rebutting. A rebutting argument against moral realism would furnish reasons to hold that key propositions of moral realism are false. An undermining argument against moral realism furnishes reasons to hold that beliefs in moral realism have no genuine truth-relevant ground. Moral realism is believed, when it is, only on a truth-irrelevant basis. If, for example, we were able to show that the belief-forming processes underlying a belief were off-track (i.e., non-truth tracking), then the belief would be undermined rather than rebutted.² Kahane (2011) structures this argument strategy in terms of justification:

1. S's moral belief that P is explained by φ
2. φ is an off-track process.
3. Therefore, S's belief is unjustified

Sophisticated versions of this argument differ in what we might call the epistemic premise—the premise which explains what is meant by *off-track* or *non-truth tracking* processes of belief formation. Consider the following three forms of evolutionary debunking arguments which differ in their formulation of the epistemic premise.

2.1 | The sensitivity formulation

1. The best explanation of the origins of our moral beliefs implies that the faculties responsible for the beliefs are insensitive to truth or falsity.
2. Thus, we would believe what we believe about right and wrong irrespective of the truth or falsity of our beliefs.
3. A belief cannot qualify as knowledge if we would believe it to be true even if it were false.
4. Therefore, our moral beliefs do not track truth well enough to count as knowledge.³

2.2 | The superfluity formulation

1. If S's moral belief that p is fully explained by φ , and φ neither entails nor presupposes the mind-independent truth of p, then p's being mind-independently true is explanatorily superfluous to explaining S's moral belief that p.
2. S's moral belief that p is fully explained by φ and φ neither entails nor presupposes the mind-independent truth of p.
3. Therefore, the mind-independent truth of moral belief is explanatorily superfluous.⁴

A third way of spelling out the epistemic premise of a debunking argument emphasises the independence between moral belief-forming processes and any putative moral truth.

²Moon argues that these types of arguments might be question-begging. See Moon (2021).

³This form of the debunking argument was first proposed by Michael Ruse. He argues that "...you would believe what you do about right and wrong, irrespective of whether or not a 'true' right or wrong existed." See, Ruse (1986, p. 254). The argument is based on an epistemological theory proposed by Robert Nozick (1981), though the idea resembles Fred Dretske's account of knowledge. See Dretske (1970).

⁴Gibbard touches on this form of argument in his claim that, "[i]f the [evolutionary] account is on the right track, then our normative capacities can be explained without supposing that there is a special kind of normative fact to which they typically respond." Gibbard (1998, p. 107). Richard Joyce employs this argument as an debunking strategy in his discussion of Occam's razor. Joyce (2006, pp. 188–189).

2.3 | The epistemic luck formulation

1. S's moral belief that p is explained by φ .
2. φ is an off-track process.
3. Given 2, moral realists can assert that:
 - a. there is nonetheless a relation between the fitness-tracking processes of evolution and mind-independent moral truths so that moral belief-forming processes turn out to track such truths after all, or
 - b. there is no such relation between the fitness tracking processes of evolution and mind-independent moral truth.
4. a is implausible because it is not empirically supported.
5. b is implausible because it commits us to the claim that our moral beliefs and mind-independent truths match up by coincidence, and we are not justified in believing in the product of such coincidental doxastic processes.
6. Therefore, S's belief is unjustified because both 3a and 3b are implausible.⁵

3 | THE PROBLEMS OF SCOPE, CORROSIVENESS, AND POST-HOC JUSTIFICATION

Aspiring debunkers must overcome at least three problems to construct a successful argument against moral realism, namely the problem of scope, the problem of corrosiveness, and most importantly, the problem of post-hoc justification. We aim to show that in overcoming these problems, aspiring debunkers must pre-establish or otherwise assume the very conclusion they seek to establish through their debunking arguments. To illustrate this, we follow an example from Schafer (2017) and Locke (2014, p. 231). Locke's Martian Case plays out an undermining argument against a belief about protons.

Jack has not received the training of an ordinary physics student. Rather, Martians brainwashed Jack to believe that certain kinds of streaks are caused by protons. Moreover, they brainwashed Jack to have this disposition, rather than a disposition to believe that something else causes such streaks, not because they themselves had done any physics, but simply because they liked the sound of the word "proton". You can even suppose, if you like, that there is some deep law of Martian psychology that makes them like the sound of the word "proton", and so it is in a certain sense inevitable (given the initial conditions) that Jack would come to have this disposition. Now Jack sees a streak and believes there goes a proton. However, just after coming to believe this, Jack learns that he was brainwashed in just the manner described above, and that he has not received the training of an ordinary physics student.

The argument begins with a belief about a certain state of affairs. The conclusion of the undermining argument calls for Jack to abandon that belief following his discovery that the belief is the product of a process that does not accurately track those states of affairs. The aspiring proton-belief debunker must take care to specify the target beliefs of this brainwashing, which here plays the role of off-track belief formation. It must be made clear just what beliefs are contained within the category of brainwashed beliefs (scope), what beliefs might be inferred from the brainwashed beliefs (corrosiveness), and what other non-brainwashed beliefs might allow us after the fact to infer that the brainwashed beliefs are after all justified (post-hoc justification). Facing these considerations allows us to understand the implications of the brainwashing as a distortive influence.

⁵This is a version of Street's highly influential Darwinian Dilemma. See Street (2006).

3.1 | The problem of scope

The scope of the proton-belief debunking argument needs prior specification. Just what exactly was Jack brainwashed to believe outside of the belief that certain kinds of streaks are caused by protons? Because beliefs do not exist as standalone propositions, Jack would probably require beliefs about what a proton is as a subatomic particle, that it forms part of an atom with electrons and neutrons and is positively charged. He would probably also require beliefs about what it means to create a streak in a cloud chamber, perhaps beliefs about fundamental laws of physics to assist in explaining that the streak had trajectory and velocity. If Jack did not possess beliefs necessary for possession of the central belief “certain kinds of streaks are caused by protons”, then that belief would not form. However, if Jack had not been brainwashed about a substantial portion of these beliefs, there may not be much of an argument in the first place. Jack could correctly deduce from his untainted beliefs that certain kinds of streaks are indeed caused by protons.

This point extends to moral beliefs in the specification of an evolutionary debunking argument against moral realism. Just which beliefs are contained within the scope of the argument? Levy and Levy (2020) argue that the scope of evolutionary debunking arguments seems to extend to mechanisms responsible for non-moral prudential and epistemic beliefs and therefore that such arguments have implications beyond the moral domain. Vavova (2014, p. 13) argues that an appropriate narrowing of the scope to purely moral beliefs leaves us knowing nothing about morality; it is not clear how we would go about knowing what we are mistaken about without some concrete notions of what morality *is*. Vavova argues that too narrow a scope either “accomplishes nothing new philosophically... [or] the argument collapses into the more ambitious and less promising one” (Vavova, 2014, pp. 17–18). This is the problem of corrosiveness to which we now turn.

3.2 | The problem of corrosiveness

The corrosiveness of the proton-belief debunking argument stems from the problem of its scope. What beliefs are vital to a functional understanding of physics which underpin the possibility of Jack's belief that certain kinds of streaks are caused by protons? Pretend for a moment that Jack only knows about protons being positively charged because of his observations of proton streaks. Suppose that Jack then infers from the proton's positive charge auxiliary beliefs that neutrons are neutrally charged, that electrons are negatively charged, and that this knowledge helps him to make sense of certain chemical interactions. Jack's discovery of his own brainwashing would in this case corrode all the beliefs which ground the content of his belief that certain kinds of streaks are caused by protons. From the category of beliefs undermined by Jack's discovery of the brainwashing, he can now no longer make sense of certain chemical reactions.⁶ This may not be a desirable conclusion for the proton-belief debunker, especially if the corrosiveness of the argument leaks into the very epistemic foundations of the proton-belief debunking argument itself.⁷ The aspiring proton-belief debunker has work to do here because the argument needs to undermine sufficient belief to warrant a significant challenge, but not so many beliefs that the “acid cannot be contained.”⁸

⁶Kyriacou argues that this occurs in the debunking of evaluative truths. See Kyriacou (2016, 2017).

⁷Vavova (2014, p. 20) makes a similar distinction and argues that it follows an inverse rule, “the potential strength of a debunking argument is inversely proportional to its ambition.”

⁸The association of debunking arguments with corrosiveness and acids may originally come from Dennett's book, *Darwin's Dangerous Idea*, where Dennett likens Darwinian evolution to a “universal acid” which “eats through just about every traditional concept, and leaves in its wake a revolutionized worldview, with most of the old landmarks still recognisable, but transformed in fundamental ways.” See Dennett (1995, p. 63).

3.3 | The problem of post-hoc justification

The final and we think most significant problem is that of post-hoc justification. Consider the following formulation of the brainwashed Jack example.

1. Jack believes certain kinds of streaks are caused by protons.
2. Jack's belief that certain kinds of streaks are caused by protons is itself caused by Martian brainwashing.
3. Martian brainwashing is an off-track process.
4. Therefore, Jack's belief is unjustified.

Here 2 is a causal premise (it describes the causal conditions of Jack's belief) and 3 is an epistemic one. 2 has it that the cause of Jack's belief is Martian brainwashing. 3 has it that belief produced in this way does not reliably track the truth. In addition to the causal and epistemic premises, Kahane (2011, p. 106) stresses the importance of a further premise.⁹ He states, "our understanding of the causal premise... needs to rule out what we might call post-hoc justification." Say that Jack seeks to reinforce his belief in proton streaks by looking for evidence of them. It is vital to understand whether any such evidence would be a post-hoc justification or a post-hoc rationalisation. Consider this argument.

1. Jack believes certain kinds of streaks are caused by protons.
2. Jack's belief that certain kinds of streaks are caused by protons is explained by Martian brainwashing.
3. Martian brainwashing is an off-track process.
4. Martian brainwashing leads Jack to look for evidence that certain kinds of streaks are caused by protons and he finds genuine and compelling evidence that this is true.
5. Therefore, Jack's belief is justified. (It is justified post-hoc.)

Jack's belief is caused by an off-track process, but it is subsequently supported by evidence. This supplementary evidence really does justify his belief. Jack is epistemically lucky that the off-track process hit the truth in that it motivated him to find independent evidence for his belief. In this case, premises 2 and 3 no longer provide an undermining defeater of Jack's belief. For the debunking strategy to succeed, the debunker must show that 3 rules out 4 in that a post-hoc justification either is not possible or cannot suffice for knowledge. This is the problem of post-hoc justification.

Now consider this argument.

1. Jack believes certain kinds of streaks are caused by protons.
2. Jack's belief that certain kinds of streaks are caused by protons is explained by Martian brainwashing.
3. Martian brainwashing is an off-track process.
4. Martian brainwashing leads Jack to look for evidence that certain kinds of streaks are caused by protons. He finds what he takes to be evidence in a physics textbook. In fact, however, the textbook has been written by people who also acquired their beliefs about streaks and protons through Martian brainwashing.
5. Jack's belief remains unjustified. (His putative justification is in fact a post-hoc rationalisation. He has found company, not epistemic support.)

In this case Jack is not justified in his belief at all. Jack's arguments for his beliefs are post-hoc rationalisations due to the nature of the evidence he adduces. Kahane shows that the debunking argument cannot clearly decide

⁹Joyce (2006, p. 180) also acknowledges this but does not appear to credit the gravity of this consideration in the moral case. He concludes his belief pill analogy (which is similar to the Martian case) with the claim that, "...unless you can find some concrete evidence either in favour or against the belief you should cease to believe this thing..."

the nature of Jack's belief without a premise to the effect that all post-hoc reflection on the belief is a rationalisation, not a justification.

The problem of post-hoc justification arises from the many possible alternative grounds we may have for our beliefs. What is required of the debunker is to show, without begging the question, that other grounds one has, or may discover, for a belief are insufficient for knowledge or justification. One must rule out such grounds for the debunking attempt to be fully successful. Kahane highlights the importance of the hidden premise that all available grounds for a debunked belief are post-hoc rationalisations rather than justifications. However, the significance of this observation for evolutionary debunking arguments is more significant still. To establish that available grounds of belief are not justifications and are instead rationalisations, we will argue, is not to debunk at all. Simply put, the problem of post-hoc justification shifts the character of the argument from an undermining argument to a rebutting one.

This is problematic for the aspiring debunker. To effectively eliminate the possibility of a post-hoc justification for moral realism, they must rule out the availability of alternative good reasons to believe moral realism. This becomes problematic on two fronts. Since such a task would require demonstrating that there are no available good reasons to believe moral realism, the debunker can only proceed by either assuming or first establishing that there are such reasons. The former means that the debunker begs the question against the realist. The latter means that the debunker has already established that there are no available good reasons to believe moral realism, so they have rendered their debunking otiose. This shows that evolutionary debunking strategies are weak as standalone arguments and instead rely on underlying metaethical assumptions and arguments.

We might consider that alternative reasons for believing that mind-independent moral truths exist can be found in the work of Kant, in the arguments of naturalist reductionist realists, or even in writings of theological ethicists like Aquinas (to barely scratch the surface). What is required of the aspiring debunker is to show that all such alternative attempts at justification are mere post-hoc rationalisations. After such an exhaustive appraisal, if the aspiring debunker has found no justified alternative grounds, they can reasonably assert that *morality is a myth*. However, having made their point with purely rebutting facts (by discrediting Kant and the like), evolutionary debunking is superfluous. Such strategies are attempted argumentative shortcuts that fail.

4 | OTIOSITY OF THE EPISTEMIC LUCK ARGUMENT

According to the Epistemic Luck formulation of the evolutionary debunking strategy, there is a blanket reason to reject all forms of post-hoc justification. The argument, as developed by Street (2006), has it that moral realists are caught in a dilemma between accepting an implausible story about the connection between mind-independent moral truth and fitness-tracking processes of evolution or an equally implausible story about the coincidence between mind-independent moral truth and fitness. If this argument succeeds, then a defender of evolutionary debunking can argue on its basis that all post-hoc considerations in support of moral realism are in fact post-hoc rationalisations.

The mind-independent truth of moral beliefs does not appear to add to their fitness. A propensity to form moral beliefs and act upon them would facilitate species cooperation just as well if these beliefs were false. This is the claim of superfluity behind the epistemic luck argument. But this superfluity does not preclude a perfectly reasonable explanation of why a grasp of mind-independent moral truths, if such things exist, would constitute a successful strategy for advancing human cooperation. Say that there are mind-independent moral truths and that grasping some of them turns out to be an effective way of learning to cooperate. For example, we might say that moral truths are fundamentally about respecting persons, and respecting persons facilitates cooperation between persons. Of course, it is not the only way of establishing intra-species cooperation. Bees have found another way. But it is one way. So why is it a surprise when a vulnerable, intelligent, and profoundly interdependent species hits

upon morality as a survival strategy? They need to cooperate. The moral way is one path to cooperation. It suits rational beings like us.¹⁰

This is an unexceptional explanatory story. The fact that the *mind-independent truth* of moral beliefs plays no role in their success as a cooperative strategy is not to the point. If there are mind-independent moral truths, and belief in them facilitates cooperation, then it is no surprise that it would be extremely useful and evolutionarily advantageous for us to develop a propensity to believe them. To reject the consequent of this conditional, one must either reject the first part of the antecedent or the second. Rejecting the first (there are mind-independent moral truths) begs the question against the realist. Rejecting the second part of the antecedent (that mind-independent moral truth is such that belief in it facilitates cooperation) requires a rebuttal of substantive moral theorising. It requires that we show that true moral beliefs are not, could not be, or are very unlikely to be, things that facilitate cooperation. This initiates a discussion about the substantive nature of morality. The argument between the moral realist and anti-realist is thus not settled by appeal to alternative explanations for our propensity to believe in objective moral truths. There is much more work to be done.

Tersman (2008, p. 403) refines the evolutionary debunking strategy. "If one wants to avoid ... global scepticism [about the justification of moral intuitions], one must show that there are intuitions for which no debunking explanation can be given or where the debunking explanations are inferior to non-debunking ones." In this case, the argument focuses upon the explanatory credentials of moral intuitions. The realist and the aspiring debunker face each other on rebutting grounds to provide either (i) an epistemological explanation as to whether it is possible that our moral beliefs (or intuitions) might be ratified externally to any etiological explanation (the problem of post-hoc justification), (ii) whether the etiological explanation succeeds in fully explaining only our moral beliefs (the problem of scope), and (iii) whether that etiological explanation can be offered without becoming self-undermining (the problem of corrosiveness). Here each explanation involves shifting the argument from an undermining argument to a substantive metaethical argument.

5 | CONTAINING THE UNIVERSAL ACID

To overcome the problem of corrosiveness, debunkers must meet a demarcation challenge. Recall in the cloud streak belief case that the central issue was how to determine the scope of the debunking argument so that it would allow a clean and contained target for the Martian brainwashing. We raised the question of how we could set out the scope of Martian brainwashing so that that "protons cause streaks" is debunked without this debunking carrying over into (and thus corroding) fundamental beliefs about nuclear physics, beliefs required to give sense to the belief that protons cause streaks.

In terms of morality, a further risk is that a debunking argument in metaethics might extend to beliefs outside of the moral domain. If we take all our belief-forming capacities to be the product of evolution, and we accept the general epistemic claim that evolution selected for fitness and not truth, then we threaten to undermine the very faculties responsible for producing our beliefs about evolution and thus undermine the evolutionary debunking argument itself. Kyriacou (2017, 2019) calls our attention to the fact that the debunker must show a clear way of distinguishing between domains which can be epistemically justified in the light of natural selection and domains which cannot. This is known as the demarcation challenge. Debunkers rely on something like the following argument in facing this challenge:

1. Domains which are epistemically justified are those which pertain to cognitive faculties which would have reliably facilitated survival and reproduction.

¹⁰This would be a third-factor response, or a pre-established harmony argument. See Copp (2008), Enoch (2011), and FitzPatrick (2015).

2. Faculties which were responsible for producing beliefs about environmental states of affairs (e.g., there is a tiger in the bushes) needed to be reliable.
3. Scientific beliefs are generated by faculties which are responsible for generating beliefs about environmental states of affairs.
4. Therefore, scientific beliefs are reliable.

The question can be posed in this way. What fundamental beliefs involved in constructing the evolutionary debunking argument must be protected from an undermining argument in order to allow for the argument to go through?

Griffiths and Wilkins (2015) face the demarcation challenge by proposing something they call "The Milvian Bridge Principle"—named after the battle between Constantinus and Maxentius who fought for the Roman Empire over the Tiber in Rome in 312. As a result of Constantinus' belief in divine intervention, he won the battle. One could posit that either he won because God supported him and thus made the victory occur, or that he won because he thought that God was supporting him and so he (and his army) fought all the harder and longer.¹¹ The historical example can be reduced to the thesis that one cannot argue from the successful outcome of believing in x that x is therefore true. The analogy then drawn is that the success of moral values in our society means only that members who hold certain moral beliefs and act on them will tend to flourish in conditions which favour those actions. Moral beliefs which are deemed fitness-tracking are by no means truth-tracking. Belief-forming processes thus divide into those that track fitness only because they track the truth and those that track fitness on other grounds. These second, unreliable belief-forming processes are susceptible to debunking. This distinction allows things like sensory faculties and the corresponding empirical beliefs drawn from them to survive the debunking argument.

Griffith and Wilkins (Griffiths & Wilkins, 2015, p. 202) put the point in the following terms.

Milvian Bridge: The X facts are related to the evolutionary success of X beliefs in such a way that it is reasonable to accept and act on X beliefs produced by our evolved cognitive faculties.

Certain beliefs, such as perceptual or memory beliefs, are reasonable to accept because their reliability corresponds to evolutionary success. These beliefs include, they write (Griffiths & Wilkins, 2015, p. 213),

...those which guide mundane action...the existence of ... [a person's] body, and of other human bodies and inanimate bodies, all arranged in space and time, as well as the fact that those other human bodies knew similar things.

If such commonsense beliefs were false, we would have been misled about essential features of our environment and this would have been fatal to us.

It seems quite plausible to argue this, as one could posit that, from an evolutionary perspective, navigating the complexities of Pleistocene life required doxastic processes grounded in observations which were generally or roughly true. Griffiths and Wilkins (2015, p. 214) extend commonsense beliefs to form the foundation to scientific beliefs by arguing that "the world of the commonsense must be seen as...the world described using the perceptual and conceptual categories available." Scientific beliefs are separate to commonsense beliefs but serve as a kind of extension from the commonsense realm to the scientific realm. In a way, these commonsense beliefs serve as the shoulders on which scientific beliefs stand.

¹¹Or even that Constantinus could pretend to believe that they were supported by their God, and this would have had the same influence on the troop's morale.

It would seem then, that our cognitive faculties were not selected for their ability to do calculus, statistics or even to take a series of known premises, say about a historical event, and deduce from those premises new information about a time in the distant past. But Griffiths and Wilkins (2015, p. 219) argue that such cognitive processes are justified because,

...if evolution does not undermine our trust in our cognitive faculties, neither should it undermine our trust in our ability to use those faculties to debug themselves—to identify their own limitations, as in perceptual illusions or common errors in intuitive reasoning.

They argue that new concepts and methods which have not been directly shaped by the forces of natural selection might still be introduced and justified through the use of evolved faculties. On these grounds, Griffiths and Wilkins believe they meet the demarcation challenge.

Griffiths and Wilkins' success in meeting the challenge is questionable. Kyriacou (2018) responds that the number of unsalvageable commonsense beliefs is large enough to topple the structure which Griffiths and Wilkins sought to build, namely that of grounding scientific beliefs. Kyriacou concludes that debunkers are required to clarify their meta-epistemological commitments, as their argument faces many challenges.¹² Our focus is on how attempts to meet the demarcation challenge play out in terms of the critique of evolutionary debunking strategies we have developed here.

In order to both distinguish between belief-forming processes that are truth-tracking and those that are not and to assign commonsense beliefs to the first category and moral beliefs to the second, we must assume or establish a particular account of the nature of both commonsense and moral truth. Beliefs based on an immediate perception of one's environment must be true (if they are true) in virtue of a causal connection between environmental facts and one's perceptions of them such that one's perceptions reliably generate true beliefs. Applying this model to moral beliefs: moral beliefs must be true (if they are true) in virtue of a causal connection between moral facts and one's moral perceptions ("S sees another's action as wrong") such that one's moral perceptions reliably generate true beliefs. While we may be confident in the first sort of causal process, we have no reason to accept the second sort. The distinction drawn by Griffiths and Wilkins is a distinction between two sorts of belief fixing processes: the causally legitimate and the causally suspect. This begs the question against the moral realist. More than this, it assumes that moral beliefs are made true by causal connection with moral facts. The latter claim does not so much beg the question against the moral realist as burden them with a highly uncongenial account of moral truth. But now the question becomes, what *kind* of thing would mind-independent moral truth be if it exists? And this brings us right back to substantive metaethical argument about the nature of morality, the nature of truth, and the possibility of moral truth. Defending evolutionary debunking arguments against the demarcation challenge demonstrates another way in which debunkers must assume or pre-establish the very conclusion they seek to establish through a debunking argument.

6 | CONCLUSION

Aspiring debunkers must overcome at least three problems to construct a successful debunking of moral realism, namely the problem of scope, the problem of corrosiveness, and most importantly, the problem of post-hoc justification. We have argued that in overcoming these problems, aspiring debunkers must assume or pre-establish the very conclusion they seek to establish. The success of debunking arguments ultimately relies on the outcome of underlying substantive metaethical debates, and this undermines their status as standalone arguments.

¹²Kyriacou also develops a normative defence of Cuneo and Schafer-Landau's claim that there exists a host of substantive claims that are non-natural conceptual truths called moral fixed points. Kyriacou argues that failing to acknowledge and act in accordance with these moral fixed points exemplifies a failure of our rational agency, which he identifies as a meta-conceptual deficiency. Kyriacou thus identifies another "way out" of the challenge raised here. See Kyriacou (2018, 2021).

ACKNOWLEDGEMENT

Open access publishing facilitated by Bond University, as part of the Wiley - Bond University agreement via the Council of Australian University Librarians.

ORCID

Oscar Davis  <https://orcid.org/0000-0002-6637-4465>

REFERENCES

- Clarke-Doane, J. (2012). Morality and mathematics: The evolutionary challenge. *Ethics*, 122(2), 313–340.
- Copp, D. (2008). Darwinian skepticism about moral realism. *Philosophical Issues*, 18(1), 186–206.
- Das, R. (2016). Evolutionary debunking of morality: Epistemological and metaphysical. *Philosophical Studies*, 173(2), 417–435.
- Dennett, D. (1995). *Darwin's dangerous idea: Evolution and the meanings of life*. Simon and Schuster.
- Dretske, F. (1970). Epistemic operators. *Journal of Philosophy*, 67(24), 1007–1023.
- Enoch, D. (2011). *Taking morality seriously: A defense of robust realism*. Oxford University Press.
- FitzPatrick, W. (2015). Debunking evolutionary debunking of ethical realism. *Philosophical Studies*, 172, 883–904.
- Gibbard, A. (1998). *Wise choices, apt feelings*. Clarendon Press.
- Griffiths, P., & Wilkins, J. (2015). Crossing the Milvian Bridge: When do evolutionary explanations debunk belief? In P. R. Sloan, G. McKenny, & K. Eggleston (Eds.), *Darwin in the 21st century: Nature, humanity and god* (pp. 201–231). Notre Dame University Press.
- Joyce, R. (2006). *The evolution of morality*. The MIT Press.
- Kahane, G. (2011). Evolutionary debunking arguments. *Nous*, 45(1), 103–125 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3175808/pdf/nous0045-0103.pdf>
- Klenk, M. (2017). Old wine in new bottles: Evolutionary debunking arguments and the benacerraf-field challenge. *Ethical Theory and Moral Practice*, 20(4), 781–795. <https://doi.org/10.1007/s10677-017-9797-y>
- Kyriacou, C. (2016). Are evolutionary debunking arguments self-debunking? *Philosophia*, 44(4), 1351–1366.
- Kyriacou, C. (2017). Evolutionary debunking: The Milvian Bridge destabilized. *Synthese*, 196, 1–19. <https://doi.org/10.1007/s11229-017-1555-0>
- Kyriacou, C. (2018). From moral fixed points to epistemic fixed points. In C. Kyriacou & R. McKenna (Eds.), *Metaepistemology* (pp. 71–95). Palgrave Macmillan.
- Kyriacou, C. (2019). Evolutionary debunking: The demarcation problem. *Logos and Episteme*, 10, 175–182.
- Kyriacou, C. (2021). Moral fixed points, rationality and the “why be moral” question. *Erkenntnis*, 86, 647–664.
- Levy, A., & Levy, Y. (2020). Evolutionary debunking arguments meet evolutionary science. *Philosophy and Phenomenological Research*, 100, 491–509. <https://doi.org/10.1111/phpr.12554>
- Locke, D. (2014). Darwinian normative skepticism. In M. Bergmann & P. Kain (Eds.), *Challenges to moral and religious belief: Disagreement and evolution* (pp. 220–236). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199669776.001.0001>
- Moon, A. (2021). Circular and question-begging responses to religious disagreement and debunking arguments. *Philosophical Studies*, 178, 785–809.
- Nozick, R. (1981). *Philosophical explanations*. Harvard University Press.
- Ruse, M. (1986). *Taking darwin seriously*. Blackwell.
- Schafer, K. (2017). Evolutionary debunking arguments, explanatory structure, and anti-realism. In R. Debes & K. R. Stueber (Eds.), *Ethical sentimentalism new perspectives* (pp. 52–65). Cambridge University Press.
- Street, S. (2006). A darwinian dilemma for realist theories of value. *Philosophical Studies*, 127, 109–166.
- Tersman, F. (2008). The reliability of moral intuitions: A challenge from neuroscience. *Australasian Journal of Philosophy*, 86(3), 389–405. <https://doi.org/10.1080/00048400802002010>
- Vavova, K. (2014). Debunking evolutionary debunking. In R. Shafer-Landau (Ed.), *Oxford studies in metaethics* (Vol. 9, pp. 76–101). Oxford University Press.
- Vavova, K. (2015). Evolutionary debunking of moral realism. *Philosophy Compass*, 10(2), 104–116. <https://doi.org/10.1111/phc3.12194>
- Wielenberg, E. (2010). On the evolutionary debunking of morality. *Ethics*, 120(1), 441–464.

How to cite this article: Davis, O., & Cox, D. (2022). Three problems for the evolutionary debunking argument. *Ratio*, 00, 1–10. <https://doi.org/10.1111/rati.12348>