

Only All Naturalists Should Worry About Only One Evolutionary Debunking Argument*

Tomas Bogardus

Evolutionary debunking arguments abound, but they're all murky where it counts most: exactly which epistemic principle combines with the facts of evolution to undermine moral realism? I'll identify some possible principles but show that most are false, spoiling the arguments built upon them. And every argument threatens only "representationalist" views of moral psychology, on which moral judgments rely on mental intermediaries, for example, sentiments. Only one argument remains a menace: a new "Argument from Symmetry." But it should worry only all naturalists, pressing a trilemma: abandon moral realism, accept a rationalism incongruous with naturalism, or reject naturalism. Nonnaturalists are free and clear.

INTRODUCTION

Do the facts of evolution undermine moral realism? The question has earned much recent attention.¹ Take incest taboos, for example. These

* I am grateful for helpful comments from Nathan Ballantyne, John Bengson, Eric Hagedorn, Clayton Littlejohn, Dustin Locke, Joel Mann, Chad Marxen, and Andrew Moon.

1. See, e.g., Richard Joyce, *The Evolution of Morality* (Cambridge, MA: MIT Press, 2006); Sharon Street, "A Darwinian Dilemma for Realist Theories of Value," *Philosophical Studies* 127 (2006): 109–66; Guy Kahane, "Evolutionary Debunking Arguments," *Noûs* 45 (2011): 103–25; Alex Rosenberg, *The Atheist's Guide to Reality: Enjoying Life without Illusions* (New York: Norton), 94–114; Justin Clarke-Doane, "Morality and Mathematics: The Evolutionary Challenge," *Ethics* 122 (2012): 313–40; and Katia Vavova, "Debunking Evolutionary Debunking" in *Oxford Studies in Metaethics*, vol. 9, ed. Russ Shafer-Landau (Oxford: Oxford University Press, 2014), 76–101, as well as "Evolutionary Debunking of Moral Realism," *Philosophy Compass* 10 (2015): 104–16.

These "evolutionary debunking arguments" are a subtype of a more general worry, which one might call "the problem of irrelevant causal factors," or "the problem of historical variability." See, e.g., Roger White, "You Just Believe That Because . . .," *Philosophical Perspectives* 24 (2010): 573–615; Nathan Ballantyne, "The Problem of Historical Variability," in *Disagreement and Skepticism*, ed. Diego Machuca, Routledge Studies in Contemporary

Ethics 126 (April 2016): 636–661

© 2016 by The University of Chicago. All rights reserved. 0014-1704/2016/12603-0004\$10.00

are often said to be explained—in full—by the reproductive advantage of inbreeding avoidance mechanisms, either in the form of hard-wired reulsions (for individuals) or cultural norms (for populations).² Does this evolutionary origin story debunk the common belief that incest really is wrong? Do these evolutionary considerations generate an epistemic challenge to moral realism, that is, the view that evaluative properties are mind-independent features of reality and we sometimes have knowledge of them?³

The Ayes think like this: the connection between us and any mind-independent evaluative properties is not good enough for beliefs about those properties (“moral beliefs” for short) to count as knowledge, even if those beliefs are true. And, when we are made aware of this frail connection, we acquire an undermining defeater for many if not all of those moral beliefs. In the face of that defeater, we can no longer sensibly claim to know that these moral beliefs are true. While this would not refute the metaphysical bits of moral realism—that evaluative properties are mind-independent—it would refute the epistemological component, that is, the claim that we sometimes have knowledge of these evaluative properties.⁴ In what follows, then, I will construe evolutionary debunking arguments as problems for moral knowledge, that is, as arguments for the conclusion that beliefs about mind-independent evaluative properties do not measure up to knowledge, even if they are true.

As for the premises, we will tour the field to see that different authors use different premises, and we will catalog that diversity. I will give evolutionary debunking arguments from the literature their most cred-

Philosophy (New York: Routledge, 2012), 239–59; Miriam Schoenfield, “Permission to Believe: Why Permissivism Is True and What It Tells Us about Irrelevant Causal Influences on Belief,” *Noûs* 48 (2014): 193–218; Adam Elga, “Lucky to Be Rational” (unpublished manuscript, Princeton University); and Joshua Schechter, “Luck, Rationality, and Explanation: A Reply to Elga’s ‘Lucky to Be Rational’” (unpublished manuscript, Brown University).

2. To dip your toe into the vast literature on this so-called Westermarck hypothesis, see Daniel Fessler and Carlos Navarrete, “Third-Party Attitudes toward Sibling Incest Evidence for Westermarck Hypothesis,” *Evolution and Human Behavior* 25 (2004): 277–94.

3. As opposed to nihilism—on which there are no moral truths—and subjectivist constructivism or expressivism, on which moral truths are functions of our evaluative attitudes themselves. Others, e.g., Russ Shafer-Landau in “Evolutionary Debunking, Moral Realism and Moral Knowledge,” *Journal of Ethics and Social Philosophy* 7 (2012): 1–37, take moral realism to include a semantic claim about the content of sincere moral judgments (those judgments express beliefs rather than conative attitudes, he says). But I agree with Kahane, “Evolutionary Debunking Arguments,” 103–4, that evolutionary debunking arguments do not (and are not meant to) directly challenge the realist’s “face-value” semantic claim, and so I don’t include it here for the purposes of this article.

4. And if moral realism expunged its epistemological component, these evolutionary debunking arguments would aim to push the realist into the awkward position of not knowing that his own view is true, and of believing and asserting a view (he knows) he doesn’t know.

ible interpretations, fortifying them with plausible and relevant epistemic principles. Some of these principles have been previously explored: sensitivity and nonaccidentality at some length,⁵ safety only briefly.⁶ I will offer novel objections to evolutionary debunking arguments resting on all these principles, and for the first time bring to light how evolutionary debunking arguments depend for their success on a controversial but widespread view of moral psychology that I call “Representationalism.” Finally, I will introduce to this literature another epistemic principle to support evolutionary debunking arguments, a principle concerning evidential symmetry. By then, I will have shown that the evolutionary debunking arguments found in the literature all fail, but I will conclude that a new version—the “Argument from Symmetry”—should upset those who hold certain combinations of widely popular views. In the end, we will see more clearly exactly who should worry about evolutionary debunking arguments, and who should not.

SENSITIVITY

Let’s start with Justin Clarke-Doane’s explanation of the evolutionary worry for moral realism: “The claim that we were selected to have true moral beliefs has counterfactual force. It implies that had the moral truths been very different, our moral beliefs would have been correspondingly different—that it would have benefited our ancestors to have correspondingly different moral beliefs. Accordingly, the key implication of the claim that we were not selected to have true moral beliefs is the negation of this counterfactual.”⁷ For Clarke-Doane, the skeptical worry is that facts of evolution show that this counterfactual is false: *had the moral truths been very different, our moral beliefs would have been correspondingly different.*⁸ And

5. See, e.g., Joel Pust, “On Explaining Knowledge of Necessity,” *dialectica* 58 (2004): 71–87; White, “You Just Believe That Because . . .”; Erik Wielenberg, “On the Evolutionary Debunking of Morality,” *Ethics* 120 (2010): 441–64; Shafer-Landau in “Evolutionary Debunking, Moral Realism and Moral Knowledge”; and Justin Clarke-Doane, “What Is the Benacerraf Problem?” in *New Perspectives on the Philosophy of Paul Benacerraf: Truth, Objects, Infinity*, ed. Fabrice Pataut (forthcoming), as well as “Justification and Explanation in Mathematics and Morality,” in *Oxford Studies in Metaethics*, vol. 10, ed. Russ Shafer-Landau (forthcoming).

6. See Clarke-Doane “What Is the Benacerraf Problem?” and also “Justification and Explanation in Mathematics and Morality.”

7. Clarke-Doane, “Morality and Mathematics,” 319.

8. Some readers have expressed concern that this counterfactual involves, on their view, a metaphysically impossible antecedent, and skepticism that we can evaluate such counterfactuals in any useful way. But let’s not sell ourselves short; we do have this power. Consider this counterfactual: *if you were the number two, you’d be odd.* And now this one: *if you were the number two, you’d be even.* Exactly one of those is true, and I bet you know which, despite the impossible antecedent. If your favorite theory of counterfactuals rules otherwise, that’s a problem for the theory, not the reality it means to model.

that counterfactual is so crucial for knowledge, the skeptic says, that once we realize it's false, we can no longer sensibly claim to have any knowledge of mind-independent moral facts. Hence the epistemic challenge to moral realism from evolution: our moral beliefs lack an important virtue, a virtue that epistemologists call *sensitivity*.

SENSITIVITY: S's belief that p, formed via method M, is **sensitive** if and only if, if p were false, S would not believe that p via M.⁹

To illustrate, suppose Smith believes truly that it is 70 degrees Fahrenheit in her office on the basis of checking her thermometer. To say that Smith's belief is sensitive is to say that, had it not been 70 degrees in there, Smith would not have believed via this method that it is (her thermometer would have "sensed" the truth and reported otherwise). Smith's belief is insensitive if and only if that counterfactual is false.

Guy Kahane also takes evolutionary debunking arguments to rely on the sensitivity condition. He says: "Evolution is not a truth-tracking process with respect to evaluative truth."¹⁰ Now, he doesn't offer an explicit definition of an "off-track" process, but he does give an example: "Think, for example, of the doting mother who believes that whatever her son does is admirable—and would think so whether the son did one thing or its very opposite."¹¹ The doting mother would continue to believe her son behaves admirably even if he didn't. That sounds like a violation of the sensitivity condition.¹²

9. On its face, Clarke-Doane's counterfactual is not equivalent to the sensitivity counterfactual—Clarke-Doane makes no mention of methods, for example. But at their core, both counterfactuals ask us to imagine the truth changing so that we might evaluate whether our beliefs change in a corresponding way. Clarke-Doane's rough-and-ready counterfactual bears a striking family resemblance to the sensitivity counterfactual, which is the product of decades of refinement in the philosophical furnace. Also, in more recent work ("What Is the Benacerraf Problem?" and "Justification and Explanation in Mathematics and Morality") Clarke-Doane is explicit in his intention to consider what epistemologists call sensitivity (though his statement of the counterfactual remains not maximally precise). So I count sensitivity, as I have put it here, as the charitable interpretation of Clarke-Doane's counterfactual. A more literal interpretive strategy is, I believe, less charitable: taking charcoal sketches of epistemic principles from Clarke-Doane (and other authors below) at face value results in arguments that are either not clear enough to critically evaluate or are not maximally rigorous and so vulnerable to more trivial objections. So, to my mind, it is better to follow the spirit rather than the letter of these arguments, as I have here. I am grateful to an anonymous referee of this journal for pressing me to defend this approach.

10. Kahane, "Evolutionary Debunking Arguments," 111.

11. *Ibid.*, 106.

12. In his presentation of evolutionary debunking arguments, William FitzPatrick, "Debunking Evolutionary Debunking of Ethical Realism," *Philosophical Studies* 172 (2015): 883–904, also uses this "tracking" talk, which suggests that he has sensitivity in mind. But, unlike Kahane, he gives no clearer indication that such talk is indeed shorthand for sensitivity. Shafer-Landau ("Evolutionary Debunking, Moral Realism and Moral Knowledge," 17) reads Joyce's (*The Evolution of Morality*, 179ff.) "belief pill" discussion as relying on the

So, for both Clarke-Doane and Kahane, evolutionary debunking arguments proceed like so:

- (1) Our moral faculty was naturally selected to produce adaptive moral beliefs, and not naturally selected to produce true moral beliefs.¹³
- (2) Therefore, it is false that: *had the moral truths been different, and had we formed our moral beliefs using the same method we actually used, our moral beliefs would have been different.*
- (3) Therefore, our moral beliefs are not sensitive.
- (4) Therefore, our moral beliefs do not count as knowledge.

This evolutionary debunking argument certainly deserves the name: it begins with a premise reporting evolutionary facts and ends in moral skepticism. And it lays bare the means by which we get from here to there: via a violation of the sensitivity condition. And it produces, in me at least, the foggy unease about moral realism that rolled in when Darwin first proposed his theory.

Yet, despite its virtues, this argument has two serious shortcomings; the first narrows its scope, but the second breaks it. First, the narrowing: the move from (1) to (2) is valid on only some views of moral psychology, for example, on what we might call “Representationalism.”¹⁴ On this view, all our moral judgments come by way of a mental intermediary, indication, report, or representation, which is delivered by our moral faculty and figures crucially into our formation of moral beliefs. The intermediaries might be sentiments like attraction and revulsion (or “affect-laden intuitions,” or “gut reactions”),¹⁵ as Sentimentalists maintain, but

sensitivity principle. But Joyce’s case is underdescribed and belief pills just as plausibly violate the safety condition (see below). Since some context indicates that Joyce relies on the safety condition, I classify him thusly below.

13. Perhaps you think the phrase “moral faculty” is more aptly applied to a complex set of mental and emotional capacities and dispositions, including very flexible reason-involving capacities, i.e., all those capacities involved in making moral judgments. I think that’s a legitimate use of the phrase “moral faculty.” However, for my purposes, I mean to cleave off just one of those capacities and call it our “moral faculty,” namely, whatever it is in us that produces our moral sentiments, gut reactions, affect-laden intuitions, etc. Your indulgence in my special use of “moral faculty” will be repaid: we will more easily be able to see how differing views of moral psychology determine to what extent we should worry about evolutionary debunking arguments.

14. By the move from *p* to *q* is valid on *r* I mean the conjunction of *p* and *r* (broadly) logically entails *q*. I am grateful to an anonymous referee of this journal for pressing me to clarify this locution.

15. For more on “affect-laden intuitions,” see Joshua Greene and Jonathan Haidt, “How (and Where) Does Moral Judgment Work?” *Trends in Cognitive Sciences* 6 (2002): 517–23. The “crucially” in the Representationalist claim that *mental intermediaries figure crucially into the formation of our moral beliefs* is to allow for possibility that the intermediaries don’t all by themselves determine which moral beliefs we form. The skeptical worry has fangs so

they might instead be calculations or thoughts.¹⁶ Either way, on this view, the inference from (1) to (2) looks valid. Suppose our moral faculty—the faculty which produces these intermediate mental states—figures crucially into all our moral judgments, and this faculty is the product of natural selection, as (1) says. If the moral facts had been different, then, as long as facts about which behaviors were adaptive had stayed the same, which moral faculty we ended up with easily might have stayed the same, and which moral beliefs we ended up with easily might have stayed the same.¹⁷ And so premise (2) seems to follow, on this view of moral psychology.

long as these intermediaries play an important role in however we arrive at our moral beliefs. If they do, then if the moral facts changed we easily might keep the same moral beliefs, and so the sensitivity condition will be violated.

16. Peter Singer and Joshua Greene think deontology relies on gut reactions but utilitarian reasoning doesn't and that is why utilitarians needn't worry about evolutionary debunking arguments but deontologists must. Yet unless they endorse what I call "Rationalism" (see below), they'll allow that utilitarian reasoning relies on some intermediary conscious states, on which moral judgments are based. And so long as our moral faculty is the product of natural selection (as they'd no doubt agree), Singer's and Greene's view is on the menu of this evolutionary debunking argument. For their views, see Singer, "Ethics and Intuitions," *Journal of Ethics* 9 (2005): 331–52; and Greene, "The Secret Joke of Kant's Soul," in *Moral Psychology: The Neuroscience of Morality: Emotion, Disease, and Development*, vol. 3, ed. Walter Sinnott-Armstrong (Cambridge, MA: MIT Press, 2007).

What unites this family of views I have dubbed "Representationalism" is that our moral judgments depend crucially on intermediate conscious states, delivered by our moral faculty. This view of moral psychology has much in common with the "Way of Ideas," a major foil for Thomas Reid—see, e.g., Thomas Reid, *An Inquiry into the Human Mind on the Principles of Common Sense*, ed. Derek R. Brookes (1764; repr., University Park: Pennsylvania State University Press, 2000). More recently this sort of view in the philosophy of perception has been called Representationalism or the Representative Theory; for seminal discussions, see, e.g., H. H. Price, *Thinking and Representation* (1946; repr., New York: Haskell House, 1977) as well as *Thinking and Experience* (1953; repr., Cambridge, MA: Harvard University Press, 2013); and also Frank Jackson, *Perception: A Representative Theory* (New York: Cambridge University Press, 1977). Sentimentalism is likely the most prominent version, but some, inspired by Kant, may hold that this intermediate conscious state is a pure, affect-less, and general thought, occasioned perhaps by characteristic perceptual experiences. And from this thought we infer our particular moral judgments. Still, so long as the faculty that produces these intermediary thoughts is the product of natural selection, even this nonsentimental variety of Representationalism will be committed to the first step of this evolutionary debunking argument. Thanks to John Bengson for suggesting (without necessarily endorsing) this sort of view to me, and for many other helpful comments besides.

17. And this goes even if these intermediary states feed into "controlled processing," as Greene would have it in "The Secret Joke of Kant's Soul." Still: the output—an evaluative judgment—depends crucially on an intermediary, and therefore easily might have stayed the same had the intermediary stayed the same (even if the moral truths had differed). And it goes likewise for the similar response proposed by FitzPatrick in "Debunking Evolutionary Debunking of Ethical Realism": according to him, even if our moral faculty is the product of natural selection, no debunking conclusion follows since we might "develop and use" those faculties, with training, so as to be truth-sensitive. "Autonomous moral re-

Representationalism is popular enough, but it's not the only horse in the barn.¹⁸ Alternatives are available, and some do not license the inference from (1) to (2). For example, there's what we might call "Rational Insight Theory," or simply "Rationalism." On this view, moral beliefs may be formed partly or fully on the basis of mental intermediaries produced by our moral faculty, the faculty that Representationalists say kindles all our moral judgments. But Rationalists hold that moral beliefs also can be formed solely on the basis of what's been variously called rational insight, direct perception, direct apprehension, or simply presentation.¹⁹ No intermediary mental state is necessary for moral judgment; the truth (or falsity) of evaluative propositions may be presented to us in thought, not merely represented to us via a sentiment, a gut reaction, or a conscious seeming. We can "just see" that, for example, rape is wrong (emphasis on "just"). We need not indirectly see that rape is wrong, by feeling repulsed by it, or via any intermediary mental state at all.

Importantly, on Rationalism (1) to (2) looks invalid, or (more charitably) at least restricted to moral beliefs based, fully or partly, on those mental intermediaries (perhaps, for example, our belief that incest is wrong, when based on that characteristic feeling of disgust delivered by

flection," he calls it. But the question of what happens downstream of these mental intermediaries—the products of what I call our "moral faculty"—is independent of whether our moral faculty is sensitive to the truth. And how could it be, if it really was selected only for adaptiveness and not truth? Even if the moral facts had changed, our evolutionary history would have still produced the moral faculty we find ourselves with, with its characteristic outputs. So, given our moral faculty's insensitivity to the truth, FitzPatrick's hope for "Autonomous moral reflection" to set things aright sounds to me like the prospect of a brain in a vat digesting her seamless and systematically misleading sensory inputs and discovering via pure reflection—even fully developed and highly trained reflection—that she's envatted: completely hopeless. Garbage in, garbage out, as they say. Likewise with our moral faculty.

So, sadly for FitzPatrick, as long as our truth-insensitive moral faculty features crucially into our moral judgments, premise (2) in the above argument will follow. So neither Greene's "controlled processing" nor FitzPatrick's "autonomous moral reflection" will outrun this skeptical argument (unless they go full Rationalist; see below).

18. How many horses *are* in the barn? Well, either mental intermediaries from our moral faculty figure crucially into all our moral judgments, or they don't. If they do, Representationalism is true. If they don't, then something else figures crucially into at least some of our moral judgments. Perhaps that something else is our rational faculty—Rationalism. Or perhaps that something else is an input from divine testimony—Divine Revelation. (Or perhaps both divine testimony and rational intuition figure crucially into our moral judgments.) These seem to me to be the most widespread views in moral psychology. But there are other positions in logical space, and other views of moral psychology are limited only by the imagination and resolve of philosophers. For present purposes, it is enough to show that of three widespread views of moral psychology, evolutionary debunking arguments threaten only one.

19. For a nice recent articulation and defense of this type of view, see John Bengson, "Grasping the Third Realm," in *Oxford Studies in Epistemology*, vol. 5, ed. Tamar Szabo Gendler and John Hawthorne (New York: Oxford University Press, 2015), 1–38.

our moral faculty). In the other cases, our access to moral truths is unmediated: there is no third thing between us and the moral facts that could be sensitive to something other than truth. And so, on Rationalism, even if our moral faculty—the faculty that produces our moral sentiments, gut reactions, and so on—was selected for adaptiveness and not truth, (2) does not follow. In the good cases at least, our rational faculty figured crucially into our moral judgments, and we used an excellent method for forming moral beliefs: knowledge by direct insight. And so, given our ability to gain unmediated access to the truth of evaluative propositions, if the moral facts had been different, and had we used the same (five-star, top-drawer) method we actually used, we would have seen that the moral facts were different. Our moral judgments did not rely on any representation of the moral facts, so there was no risk of *misrepresentation* of those facts. Therefore (2) is false on Rationalism, even if (1) is true. And now you see that the fallout of this argument is limited: Rationalists need not fear this type of evolutionary debunking argument.

And there is a further restriction on the scope of this argument. On another view of moral psychology—let’s call it “Divine Revelation”—our moral and rational faculties may play determinative roles in the production of moral beliefs. But there is also a third option: moral beliefs can be formed solely on the basis of divine testimony, a supernaturally endowed conscience, the inward instigation of the Holy Spirit (as Aquinas might say), and so forth. On this view, as on Rationalism, (1) to (2) looks invalid, or (more charitably) at least restricted to moral beliefs based crucially on the deliverances of our evolved moral faculty. On Divine Revelation, there is indeed a third thing between us and the moral facts (namely, some vehicle of divine testimony), but it is not the product of our moral faculty, that is, the faculty that produces moral sentiments, gut reactions, and so on. So, even if our moral faculty was selected for adaptiveness and not truth, (2) does not follow. And that is because: if the moral facts had been different, the moral testimony of a loving God would have been different.²⁰ And if God’s moral testimony had been different, on this view, then our moral beliefs informed by that testimony would have been correspondingly different.²¹ So (2) is false on Divine Revelation, even if (1) is true. The scope of this argument is thereby further restricted: Divine Revelationists need not worry about this species of evolutionary debunking argument. This result is important if only because there have been and still are many Divine Revelationists—perhaps not lately in academia but certainly out there in the wild.

20. Even if, as some hold, God’s testimony (or will) determines the moral facts.

21. In “Evolutionary Debunking Arguments,” 109, Guy Kahane agrees (though he expresses some reservations later in his note 16): “If we are the designed products of God, then it does seem rational for us to rely on our natural doxastic dispositions given that these were implanted in us by an omniscient and omnibenevolent being.”

Aside from its narrow scope, there's a second, debilitating shortcoming of this argument: the move from (3) to (4) isn't valid. Recent work in epistemology has cast serious doubt on the claim that sensitivity is required for knowledge. For a useful summary of the case against sensitivity, the reader may consult Roger White's recent discussion.²² Here's one objection from White:

The problem arises . . . generally from cases of empirical underdetermination. As jury members we have overwhelming evidence that the defendant is guilty (finger prints, eyewitness reports, DNA matching blood on the murder weapon . . .). Although we are more than justified in believing that he is guilty, our evidence doesn't strictly entail it. There is always logical room for some wildly ad hoc conspiracy theory according to which the defendant was framed and all the evidence was planted. We will have to be justified in denying this conspiracy theory if we're justified in believing the defendant to be guilty (since they cannot both be true).²³ But if the conspiracy theory were true we would have all the same evidence and hence quite reasonably (but mistakenly) believe the defendant to be guilty and hence the conspiracy theory to be false. So again, by Truth Sensitivity we derive the absurd conclusion that we aren't justified in denying the conspiracy theory or believing that the defendant is guilty.

This objection and others have lately put sensitivity theorists on the back foot. If you find the objection persuasive, as I do, we may proceed on the assumption that sensitivity is not required for knowledge. And so, even if an evolutionary debunking argument could show that our moral beliefs were not formed sensitively, that by itself would not debunk our moral beliefs. I take this shortcoming to disable this evolutionary debunking argument, as construed by Clarke-Doane and Kahane. But there are other species of evolutionary debunking arguments, and the next was considered by Darwin himself.

22. White ("You Just Believe That Because . . .," sec. 4.1). Though for a few spirited defenses of sensitivity, the reader may consult the collection from Kelly Becker and Tim Black, *The Sensitivity Principle in Epistemology* (New York: Cambridge University Press, 2012).

23. To resist this argument, the sensitivity theorist may reject this sentence. But he will thereby open himself up to a common and powerful objection to sensitivity: that sensitivity theorists must deny that knowledge is closed under known (or, in White's case, justifiably believed) entailment. It is obvious (and knowable, and justifiably believable) that: if the defendant is guilty, then none of these ad hoc exculpatory conspiracy theories is true. To avoid rampant skepticism, we should allow that we can know (and justifiably believe) that the defendant is guilty. To deny that we can know (or justifiably believe) that none of these exculpatory theories is true would require giving up a highly plausible closure principle. And, in any event, denying we can know *none of the conspiracy theories is true* is implausible on its face, independent of a more general closure principle. So much the worse for sensitivity, and for any evolutionary debunking argument hitched on to it.

SAFETY

Our first attempt to state the epistemic challenge to moral realism from evolution centered on the notion of sensitivity. But other statements of the skeptical worry seem to focus on a different virtue of belief-forming methods, namely, safety.

SAFETY: S's belief that *p*, formed via method *M*, is **safe** if and only if, if S were to believe that *p* via *M*, *p* would be true.

Take, for example, the argument that E. O. Wilson and Michael Ruse offer in favor of their moral skepticism:

Suppose that, instead of evolving from savannah-dwelling primates, we had evolved in a very different way. If, like the termites, we needed to dwell in darkness, eat each other's feces and cannibalise the dead, our epigenetic rules would be very different from what they are now. Our minds would be strongly prone to extol such acts as beautiful and moral. And we would find it morally disgusting to live in the open air, dispose of body waste and bury the dead. Termite ayatollahs would surely declare such things to be against the will of God. . . . Ethics does not have the objective foundation our biology leads us to think it has.²⁴

The idea seems to be that, had things gone differently in our evolutionary history—and they easily might have—the method by which we arrived at our current moral beliefs would have produced quite different moral beliefs. And so, imagining that we “rewind the tape,” so to speak, and rely on the same method of moral belief formation that we actually used, we see that we easily might have arrived at moral beliefs that, by our own lights, would have been false. And that, one might think, proves that our moral beliefs were not formed safely, and so do not count as genuine knowledge.²⁵

Charles Darwin considers a similar thought: “If, for instance, to take an extreme case, men were reared under precisely the same conditions as hive-bees, there can hardly be a doubt that our unmarried females

24. E. O. Wilson and Michael Ruse, “The Evolution of Ethics,” *New Scientist* 108 (October 17, 1985): 52

25. Here, it is quite clear that Ruse and Wilson mean to arrive at moral skepticism via a violation of the safety condition. But in a slightly later work, “Moral Philosophy as Applied Science,” *Philosophy* 61 (1986): 173–92, 187, Ruse and Wilson seem to rely instead on a violation of the *sensitivity* condition: “even if external ethical premises did not exist, we would go on thinking about right and wrong in the way that we do.” There they imagine the moral facts changing and conclude that our moral beliefs would not keep pace: insensitivity. We have already seen, however, that such an argument resting on sensitivity is ruined by a limited scope and invalid inferences.

would, like the worker-bees, think it a sacred duty to kill their brothers, and mothers would strive to kill their fertile daughters; and no one would think of interfering.”²⁶ Though he doesn’t commit himself to the same skeptical conclusion as Wilson and Ruse, Darwin does fret over the evolutionary threat to moral realism, and he seems to believe that the danger lurks in the neighborhood of the counterfactual that Wilson and Ruse highlight: had our evolutionary history gone differently, we may well have had moral beliefs that, by our own lights, would have been false.

Richard Joyce also seems convinced by this variety of evolutionary debunking argument: “Were it not for a certain social ancestry affecting our biology, the argument goes, we wouldn’t have concepts like obligation, virtue, property, desert, and fairness at all. If the analogy is reasonable, therefore, it would appear that once we become aware of this genealogy of morals we should (epistemically) . . . cultivate agnosticism regarding all positive beliefs involving these concepts until we find some solid evidence either for or against them.”²⁷ Joyce asks us to imagine our social ancestry having gone differently and concludes that we would have ended up with radically different evaluative beliefs, beliefs in tension with our current moral systems.²⁸ That, he thinks, provides enough fuel for an evolutionary debunking argument, and the argument persuades him.²⁹

26. Charles Darwin, *The Descent of Man* (1871; repr., Amherst, NY: Prometheus Books, 1998), 102.

27. Joyce, *The Evolution of Morality*, 181.

28. In *The Atheist’s Guide to Reality*, 112–13, Alex Rosenberg endorses an evolutionary debunking argument that relies on the idea that natural selection “can’t be a process that’s reliable for providing us with what we consider correct moral beliefs.” The context makes clear that he thinks of reliability as a modal notion (“natural selection was relentlessly moving through the design space of alternative ways of treating other people, animals, and the human environment,” he says on 107–8). He also comes close to an explicit statement of a safety violation: “If the environment had been very different, another moral core would have been selected for. . . . But it wouldn’t have been made right, correct, or true by its fitness in that environment.” And so I believe his argument should be classified as a variant of the safety argument in this section.

29. Shaun Nichols, “Process Debunking and Ethics,” *Ethics* 124 (2014): 727–49, advances an evolutionary debunking argument that relies on our moral faculties being “epistemically defective processes.” He wishes to sidestep the question of what makes a process epistemically defective, thinking it unnecessary for the debunker’s purposes since (*ibid.*, 733) “we can often register that a process is defective without taking sides on what makes the process defective.” Unless he expects the moral realist to just admit, without argument, that our moral faculties are defective, this simply will not do for a non-question-begging evolutionary debunking argument. However, Nichols approvingly cites Alvin Plantinga’s proper-function account of warrant as a gesture toward explaining “epistemically defective process,” an account found in Plantinga’s *Warranted Christian Belief* (New York: Oxford University Press, 2000). Presumably Nichols’s thought is that a naturally evolved faculty (selected for adaptiveness and not truth) violates this condition of Plantinga’s account: *the process is functioning according to a design plan successfully aimed at truth*. And, one might think, if a process violates that condition, then it also violates the safety condition. So, Nichols is

And so, inspired by Wilson and Ruse, Darwin, and Joyce, we might run an evolutionary debunking argument through the safety condition on knowledge like so:

- (5) Our moral faculty was naturally selected to produce adaptive moral beliefs, and not naturally selected to produce true moral beliefs.
- (6) Therefore, it is false that: *had our species evolved elsewhere, elsewhen and we later formed moral beliefs using the same method we actually used, our beliefs would be true.*
- (7) Therefore, our moral beliefs are not safe.
- (8) Therefore, our moral beliefs do not count as knowledge.

This argument, as with the sensitivity argument of the last section, does produce the rippling unease toward moral realism that one looks for in an evolutionary debunking argument. And this argument relies not on sensitivity but on a widely accepted and more plausible condition on knowledge: safety. Two cheers for this argument.

But the argument has three shortcomings: the first narrows its scope while the second and third sink it. First, of the three widespread views of moral psychology we have discussed, the move from (5) to (6) relies on Representationalism. If the process of moral judgment relies on intermediary mental states (sentiments, gut reactions, “affect-laden intuitions,” etc.), and the faculty that produces these intermediaries was selected for adaptiveness but not selected for truth as (5) says, then (6) is guaranteed: had our evolutionary history differed so as to change what behaviors were adaptive, then that moral faculty delivering our moral sentiments would have produced different intermediaries, and—if Representationalism is true—we therefore easily might have ended up with moral beliefs that, by our own lights, would have been false.

But, as with the sensitivity argument of the last section, things look different on Rationalism and on Divine Revelation. On Rationalism, even if (5) is true, we actually used our ability to “just see” the truth value of evaluative propositions in order to form (at least some of our) moral beliefs, and this method would not easily lead us astray even if our social ancestry had been different. Nothing stands between us and the truth that might have misled us had we evolved elsewhere, elsewhen. So this argument does not threaten those moral beliefs formed in the light of our rational faculty and not merely our moral faculty. And on Divine Reve-

cagey about the details of debunking arguments. But if, like me, you long to understand the epistemic principles beneath the evolutionary challenge to moral realism and you share my willingness to “read between the lines” to this end, you’ll join me in placing his argument among the safety arguments of this section. (Otherwise, we might just admit Nichols’s argument is too indeterminate to be usefully classified at this level of detail.)

lation, even if (5) is true, we actually used divine testimony to form (at least some of) our moral beliefs. And the testimony of a loving God would not have led to radically false moral beliefs had our social ancestry been different.³⁰ So those moral beliefs are not threatened by this argument. Again, then, of the three widespread views of moral psychology we have discussed, this argument is a danger only to Representationalism.

Yet there are two additional, more serious objections to this argument. The first is that the move from (6) to (7) isn't valid. Premise (6) tells us that, had something gone differently in our past, then, using the same method we actually used to acquire moral beliefs, we easily might have believed falsely. But it does not follow from this that we actually believed unsafely, as (7) says. To see this, consider a nonepistemic analogy. On your drive home from work, you always use the Homeward Bridge. Today it is as sturdy as ever, and you safely cross it. Earlier that day, however, a pack of unprincipled teenagers tried to detonate some small-scale explosives under Homeward Bridge, which would have rendered it unsound and hazardous by the time of your approach. Fortunately for you, their matches were wet and they shuffled home in low spirits. But had they succeeded, as they easily might have, the bridge would have been unsafe. Nevertheless, Homeward Bridge was a safe way for you to get home today. In this case, an analogue of premise (6) does not entail an analogue of premise (7): the fact that something nearly happened in the past that would have made your way unsafe doesn't entail that the way was unsafe when you actually used it.

Here's an epistemic case to cement the point. Suppose that the infamous Evil Epistemologist has poisoned the world's water supplies with a drug that radically impairs human cognitive faculties. Once exposed to the drug, all of one's faculties become completely unreliable. However, a benevolent nurse used his only dose of antidote to immunize you in the maternity ward. Your faculties are therefore safe from this poison, while everyone else's faculties are in serious danger. Yet it is true that, had you been born in a different time or a different place, and used the same faculties and methods you actually used, you easily might have believed things that, by your own lights, would be false. So we have here a counterexample to the general form of the inference that is meant to carry us from (6) to (7). The fact that something might have happened (to my ancestors, say) which would have rendered my faculties unsafe does not

30. The testimony of a loving God may have differed had our social ancestry differed, if other behaviors had conduced to our flourishing as a result (just as a virtuous teacher's advice to an aspiring sculptor may differ from her advice to an aspiring lyre player, and a farmer's advice to her chickens may differ from her advice to her bees). But this testimony would not have been false, and so (6) would not follow from (5). Thanks to Eric Hagedorn and Joel Mann for raising this objection.

entail that my faculties are actually unsafe. The second inference of the argument is, therefore, invalid.³¹

And so is the inference from (7) to (8), for it assumes that safety is required for knowledge. That is, the move from (7) to (8) assumes that a subject can know that *p* via some method only if that method would not easily have delivered error. Only if, that is, not easily would the subject believe that *p* via this method without it being the case that *p*. But, as I will now argue, safety is not required for knowledge.

Here's a general recipe for whipping up a counterexample to the safety condition: first, pick the most virtuous belief-forming method you can imagine, and have a subject form a belief via that method. Second, add a twist of fate: put the method in danger of malfunctioning, but let the danger remain purely counterfactual. Now, since things could have gone less well epistemically but didn't, it is quite tempting to admit that the virtuous method produced knowledge. And yet the belief was not formed safely, for there are many nearby possible worlds in which the method goes awry. Sometimes our actual performance is enough for knowledge, even under threat of various externalities.

For example, suppose that the world's most accurate clock hangs in Smith's office, and Smith knows this. This is an atomic clock: its accuracy is due to a clever radiation sensor, which keeps time by detecting the transition between two energy levels in cesium-133 atoms. This radiation sensor is very delicate, however, and could easily malfunction if a radioactive isotope were to decay in the vicinity. This morning, against the odds, someone did in fact leave a small amount of a radioactive isotope near the world's most accurate clock in Smith's office. This alien isotope has a relatively short half-life, but—quite improbably—it has not yet decayed at all. It is 8:20 a.m. The alien isotope will decay at any moment, but it is indeterminate when exactly it will decay. Whenever it does, it will

31. It is worth noting that if the inference from (6) to (7) were valid, then we would plausibly have counterexamples to the inference from (7) to (8), i.e., we could plausibly construct cases of unsafe knowledge. See Juan Comesaña, "Unsafe Knowledge," *Synthese* 146 (2005): 395–404; Ram Neta and Guy Rohrbaugh, "Luminosity and the Safety of Knowledge," *Pacific Philosophical Quarterly* 85 (2004): 396–406; and Christoph Kelp, "Knowledge and Safety," *Journal of Philosophical Research* 34 (2009): 21–31. These authors all develop alleged cases of unsafe knowledge that depend on the move in question: something easily might have happened in the past which would have rendered my beliefs now unsafe, therefore they are now unsafe. If you think that is a good inference—i.e., if you think the move from (6) to (7) is valid—then these philosophers provide a major obstacle to the move from (7) to (8). However, I deny the inference from (6) to (7), and I believe this is a fatal shortcoming for the proposed cases of unsafe knowledge from Comesaña, Neta and Rohrbaugh, and Kelp. (For a fuller defense of this claim, see Tomas Bogardus, "Knowledge under Threat," *Philosophy and Phenomenological Research* 88 [2014]: 289–313.) But there are clear examples of unsafe knowledge that do not rely on this hasty inference, as we will see below.

disrupt the clock's sensor, and—for complicated sciencey reasons that I cannot explain here—it will freeze the clock on the reading “8:22.”

The clock is running normally at 8:22 a.m. when Smith enters her office. Smith takes a good hard look at the world's most accurate clock—what she knows is an extremely well-designed clock that has never been tampered with—and forms the true belief that it is 8:22 a.m.

Many classic and widely accepted theories of knowledge rule that Smith knows it is 8:22 a.m.³² Further, it strikes me as obvious that Smith knows it is 8:22 a.m., and perhaps the same goes for you. The threat to her clock remains purely counterfactual, and everything is functioning properly when she forms her belief. Yet, since the isotope could easily have decayed and frozen the clock at “8:22,” and since Smith may easily have checked the clock a moment earlier or later, Smith might easily have believed it is 8:22 a.m. without its being 8:22 a.m. Smith formed her belief in a way that could easily have delivered error.

Therefore, Smith knows and yet her belief was not formed safely. Safety, then, is not required for knowledge. And so the inference from (7) to (8) above is invalid, as was its friend, the inference from (6) to (7). And therefore we're knee deep in reasons to think that the safety argument does not threaten the rationality of moral belief, even if Representationalism is true. Let's turn elsewhere, then, to see if we might find a more intimidating evolutionary debunking argument.

ACCIDENTALITY

Many epistemologists share an anti-luck intuition—knowledge excludes a certain kind of luck—and they typically offer Gettier cases as paradigmatic examples of this type of toxic luck. Gettier himself seems to have the intuition, when he pointed out that it is only by the “sheerest coincidence” that Smith's belief turned out to be true, in support of his claim that Smith lacked knowledge in one of the cases in his famous 1963 paper. As conditions on knowledge, safety and sensitivity are often offered as analyses of this sort of knowledge-precluding luck.

But suppose that, as we have argued above, neither sensitivity nor safety is able to sharpen an evolutionary debunking argument into a stinging attack on moral realism. It still may seem as though evolutionary debunking arguments concern how lucky one would have to have been to get her moral beliefs right, if moral realism is true. Think on all the ways we might have evolved, all the moral beliefs we might easily have formed via our evolved moral faculty: if one happened to get those mind-independent moral facts right, it could be due (the skeptic says) only to

32. For a defense of this claim, see Bogardus, “Knowledge under Threat.”

that sort of knowledge-precluding luck. Though safety and sensitivity look unpromising, there is still hope of developing the evolutionary challenge to moral realism along these lines. For we may pursue another strand of thought within anti-luck epistemology: an analysis of luck in terms of accidentality.

Peter Unger, for example, offers this analysis of knowledge: S knows that p just in case it is not at all accidental that S is right about its being the case that p.³³ Unger is careful to point out that some types of accidentality are compatible with knowledge. For example, suppose that a disgruntled student tries to run me down with his car, but at the last moment loses control and rolls harmlessly into a ditch. “I can’t believe you crashed your car trying to kill me!” I exclaim, though I really can and do. Indeed, I know it. Now, that the car crashed was an accident, and so the content of my belief that it crashed is true by accident. Yet I know that it crashed. So that sort of accidentality is compatible with knowledge. And it is an accident that I am alive and able to believe that the car crashed, since the driver failed to kill me only by accident. And so this kind of accidentality too is compatible with knowledge. What’s incompatible with knowledge, for Unger, is that there should be any accidentality “between the man and the fact,” as he says. That is, given my evidence—which may exist or have been acquired by accident—and given that I am there to consider the evidence—this too may be the result of an accident—it cannot be an accident that I form a true belief on my evidence. Or, as Unger says, in believing that the car crashed, it cannot be an accident that I am right about its being the case that the car crashed.

Casting the evolutionary challenge to moral realism as an attack on the accidentality of moral belief rather than the safety or sensitivity of moral belief will avoid the debilitating objections raised in the previous sections. In addition, this strategy has the virtue of being faithful to other statements of evolutionary debunking arguments in the literature.

Sharon Street, for example, puts the evolutionary challenge to moral realism this way: “The realist must hold that an astonishing coincidence took place—claiming that as a matter of sheer luck, evolutionary pressures affected our evaluative attitudes in such a way that they just happened to land on or near the true normative views among all the conceptually possible ones.”³⁴ And Matthew Bedke has it that “cosmic coincidence” is a defeater for intuitive nonnaturalism; he too seems to gesture toward

33. Peter Unger, “An Analysis of Factual Knowledge,” *Journal of Philosophy* 65 (1968): 157–70, 159. See also William Harper, “Knowledge and Luck,” *Southern Journal of Philosophy* 34 (1996): 273–83; and Masahiro Yamada, “Getting It Right by Accident,” *Philosophy and Phenomenological Research* 83 (2011): 72–105.

34. Sharon Street, “Reply to Copp: Naturalism, Normativity, and the Varieties of Realism Worth Worrying About,” *Philosophical Issues* 18 (2008): 208–9.

this nonaccidentality, anti-luck condition on knowledge.³⁵ We have good textual reason, then, to pursue accidentality as the “active ingredient” in evolutionary debunking arguments.³⁶

35. Matthew Bedke, “Intuitive Non-Naturalism Meets Cosmic Coincidence,” *Pacific Philosophical Quarterly* 90 (2009): 188–209.

36. Though one often finds, nestled in with this talk of coincidence, suggestions that moral realism is explanatorily superfluous given the facts of evolution. See, e.g., Philip Kitcher, “Biology and Ethics,” in *The Oxford Handbook of Ethical Theory*, ed. David Copp (Oxford: Oxford University Press, 2006), 163–85, 176: “When we see children setting light to a cat, do we perceive the wrongness of what they are doing? Is it necessary to invoke the ‘fact that it is wrong’ to explain the judgment? Apparently not.” See also Joyce (*The Evolution of Morality*, 209–10, 219) for talk of explanatory superfluity: the problem for justification of any moral belief on moral realism is the existence of “an empirically supported theory that provides a complete explanation of why we have that belief while nowhere presupposing its truth.” This thought may trace back to Gilbert Harman, *The Nature of Morality: An Introduction to Ethics* (Oxford: Oxford University Press, 1977), 3–26. And, for further development, see Dustin Locke, “Darwinian Moral Skepticism,” in *Challenges to Moral and Religious Belief: Disagreement and Evolution*, ed. Michael Bergmann and Patrick Kain (New York: Oxford University Press, 2014), 220–36. For a nice response to “explanationist” skepticism about intuitions generally, see Joel Pust, “Against Explanationist Skepticism Regarding Philosophical Intuitions,” *Philosophical Studies* 106 (2001): 227–58.

One might wonder about an evolutionary debunking argument that turned on a principle we might call “Ockham”: if we know of facts that completely explain some phenomenon, and these facts do not appeal to or presuppose X (or even make X likely), and we have no reason to believe in X other than that it might explain this phenomenon, then we should not believe in X. Combined with the claims that a purely natural moral genealogy can completely explain our moral judgments, and that this genealogy does not appeal to, presuppose, or make likely mind-independent evaluative properties, and that we have no other reason to believe in such properties, this “Ockham” principle would yield a skeptical conclusion.

But moral realists can reasonably deny the premise that there is no reason to believe in mind-independent evaluative properties but for their ability to explain our moral judgments. Even if pain best explains my pain judgments, my primary reason to believe in pain is that I perceive it—that pain seems obvious—not how neatly it explains my judgments. Similarly, a moral realist might hold that our primary reason to believe in mind-independent evaluative properties is that we perceive them—that they seem obvious—not that they nicely explain our moral judgments. Moral realism is not best seen as a theoretical posit, least of all to explain our moral judgments; it is a report of apparent data, of how things strike us as obvious. And, without that no-other-reason-to-believe-in-X premise, a revised Ockham principle would be vulnerable to counterexamples, like so: I learn that a wizard is giving me visual hallucinations that are occasionally veridical, i.e., visual hallucinations that sometimes, by chance, “match” the world. I have a complete wizard-laden explanation of why I am visually appeared to apple-ly right now that does not appeal to or presuppose or even make likely that there is an apple before me. But I also feel the apple with my hands and taste it with my mouth, two senses that (let us stipulate) I know are wizard-proof. This case would meet the revised “Ockham” principle, and that principle would therefore rule that I should not believe there’s an apple before me. That is the wrong result, and so this (admittedly bizarre) apple-wizard case takes down that revised “Ockham” principle, along with any evolutionary debunking argument that turns on it.

Rationalists and Divine Revelationists may reasonably deny that we have a complete explanation of our moral judgments, an explanation that makes no appeal to mind-

And so perhaps the argument is meant to go like this: when we reflect on certain counterfactuals featuring us descending from different social ancestries, we realize that we easily might have ended up with moral beliefs that we would consider false. As Street says, we “just happened to land on or near the true normative views among all the conceptually possible ones.” And these counterfactuals should convince us that our moral beliefs, if true at all, are true by sheer accident. And by this route we should conclude that our moral beliefs do not amount to knowledge. A bit more carefully, we begin as we did before with the sensitivity argument and the safety argument:

- (9) Our moral faculty was naturally selected to produce adaptive moral beliefs, and not naturally selected to produce true moral beliefs.

From there we might proceed as we did with the safety argument:

- (10) Therefore, it is false that: *had our species evolved elsewhere, elsewhere and we later formed moral beliefs using the same method we actually used, our beliefs would be true.*

But now we cut a new path and conclude:

- (11) Therefore, it is accidental that we are right about our moral beliefs (or we are not right at all).

And, hence,

- (12) Our moral beliefs do not count as knowledge.

Whatever its virtues, the argument has problems. As with the safety argument above, the first inference of this argument should concern only

independent evaluative properties, since, on their view, not all of our moral judgments come by way of our naturally evolved moral faculty. On their views, we have alternative means of ascertaining mind-independent evaluative properties: rational intuition or divine revelation. And so, according to them, we’re a bit like the character in the apple-wizard case who can check for the apple with his additional senses. Representationalists, on the other hand, think that our naturally evolved moral faculty is crucially involved in all our moral judgments, and so they cannot take this same escape route.

But, in their defense, Representationalists might reasonably point out that, whatever force there is to considerations of explanatory superfluity, upon closer inspection that force resolves itself into insensitivity, unsafety, or accidentality. Why worry about a complete explanation of our moral judgments that makes no appeal to mind-independent moral facts, if not because such an explanation shows that our moral judgments don’t have the right sort of modal connection to the truth, or that they’re accidentally right if right at all? So even the Representationalist need not worry about this “explanatory superfluity” evolutionary debunking argument, since it’s merely one of those previous arguments dressed up in a new guise, arguments which we’ve dismissed. I’m grateful to Dustin Locke for helpful comments on this type of argument, and for his dissenting opinion, see Locke, “Darwinian Moral Skepticism.”

Representationalists. Rationalists and Divine Revelationists need not fear this argument.

And neither should Representationalists, in fact, since the inference from (10) to (11) is invalid. Premise (10) tells us that, had something accidental and out of our control gone differently in the past, then the method we used to acquire moral beliefs could easily have led us astray. But this premise is blind to Unger's distinction between benign and toxic kinds of accidentality discussed above. For it may be that, had something accidental and out of our control gone differently in the past, then we would have lacked good evidence that we actually had, and that is why the method we used could easily have led us astray. But that sort of accidentality is not the toxic, knowledge-precluding luck Unger tried to capture with the phrase featured in (11). And so it is possible to put together a counterexample to the inference from (10) to (11).

Here's one. Suppose it's 1515 and you are an elderly friend of Copernicus in Fromburk. Copernicus has collected enough data to justify belief in heliocentrism, but, for fear of ecclesiastical reprisals, he has not and will not share these data with the world for over a decade, long after your death. He has, however, shared his findings with you. You can come to know on the basis of Copernicus's testimony that heliocentrism is true, and it is no accident that you are right. Yet had the past been different—had you been born and raised elsewhere, elsewhere, for example, far away from your friend Copernicus—and formed astronomical beliefs using the same method you actually used (namely, expert testimony), you easily could have been led astray into thinking that heliocentrism is false. Only by a mere accident of birth did you acquire Copernicus's solid gold testimony. But this species of accidentality is not the sort described in premise (11); given the evidence you have acquired, it is not an accident that you are right about heliocentrism's being true. So we have here a counterexample to the general form of the inference that is meant to shuttle us from (10) to (11). The accidentality captured in (10) may not infect the relationship "between the man and the fact" required by (11), and so (10) doesn't entail (11).³⁷

37. As we have just seen, Rationalists and Divine Revelationists may grant (9) while denying the inferences from (9) to (10) and from (10) to (11). And, as argued in the last section, Rationalists and Divine Revelationists have good reason to reject (10). But they also have good reason to reject (11): if, as Rationalism holds, we can form moral judgments on the basis of unmediated access to the truth of some moral facts, then there would be no accidentality, as Unger puts it, "between the man and the fact" that could prevent knowledge—there's nothing at all "between the man and the fact," on this view, and so it won't be an accident that these moral judgments are correct. And if Divine Revelationism is true, we might form moral judgments on the basis of divine testimony: here too it will be no accident that we are correct about such moral judgments, since it will be no accident that the divine testimony is true.

I have argued elsewhere against the general form of the inference from (11) to (12).³⁸ A brief version of the argument might borrow John Hawthorne's swampwatch, a particle-for-particle duplicate of the world's most accurate wristwatch "created by a fortuitous coming together of molecules."³⁹ Swampwatch reports the time but, given its birth from chaos, its reports are not aimed at the truth or anything else. And so, if a subject were to believe truly that *p* on the basis of swampwatch's reports, it would be an accident—in a familiar and legitimate sense of the word "accident"—that the subject is right about *p*'s being the case. And yet, Hawthorne says, when a subject "uses his swampwatch to inform him about the time . . . we are intuitively ready to say that [he] knows what time it is." If you share Hawthorne's intuition, as I do, then we have here a case in which someone knows that *p* and yet it is accidental that she's right about *p*'s being the case. In ascribing knowledge, we prize regular, reliable success so highly that we will tolerate even some accidentality "between the man and the fact." This may well go for our evolved moral faculty; for all the argument shows, our moral faculty is regularly, reliably successful, like a swampwatch. We thereby see that (11) does not entail (12). With such a narrow scope and two invalid inferences, this evolutionary debunking argument is about as dangerous as a paper tiger. Can the skeptic do better?

THE ARGUMENT FROM SYMMETRY

We have seen that all of the most prominent statements of evolutionary debunking arguments are limited in scope and invalid. But, for me at least, the conviction persists that there's something to them. What could it be? I suggest that we take a page from a closely related skeptical argument, an argument for the conclusion that religious beliefs are undermined by their historical contingency. Here's how Philip Kitcher puts the argument: "Had the Christians been born among the aboriginal Australians, they would believe, in just the same ways, on just the same bases, and with just the same convictions, doctrines about Dreamtime instead of about the Resurrection. The symmetry is complete. . . . Given that they are all on a par, we should trust none of them."⁴⁰ Obviously, this argument targets religious belief, not moral realism. But perhaps the moral skeptic could adapt it to her purposes. As before, the facts of evolution

38. See Tomas Bogardus, "The Problem of Contingency for Religious Belief," *Faith and Philosophy* 30 (2013): 371–92.

39. John Hawthorne, "Deeply Contingent A Priori Knowledge," *Philosophy and Phenomenological Research* 65 (2002): 247–69.

40. Philip Kitcher, "Challenges for Secularism," in *The Joy of Secularism: 11 Essays for How We Live Now*, ed. George Levine (Princeton, NJ: Princeton University Press, 2010), 24–56, 26.

should convince us of a certain kind of “modal fragility” of our moral beliefs. In particular, it ought to convince us that had our species evolved elsewhere, elsewhen—as easily might have happened—and we later formed moral beliefs using the same method we actually used, our beliefs may easily have been incompatible with our actual beliefs, false by our own lights. Now, the novelty: that modal fragility reveals a troubling evidential symmetry between us-as-we-actually-evolved and us-as-we-easily-might-have-evolved. The facts of evolution teach us, on this line of reasoning, that we have no independent, non-question-begging reason to believe that we are in a better epistemic position on moral questions than our many nearby counterfactual selves who disagree with us on moral questions.

This argument maps out terrain that has already been extensively explored (and colonized) by those working on the epistemic significance of peer disagreement. Analogies one finds in that literature suggest themselves here: if my thermometer says it’s 70 degrees in here while your thermometer says it’s 80 degrees, and I have no reason independent of this disagreement to favor mine over yours, it would be unacceptably question-begging or arbitrary to nevertheless stick with the report of my thermometer over yours. The fact that this thermometer is mine is not an epistemically respectable symmetry breaker.⁴¹ And the same goes even if you and your thermometer never enter the scene, so long as I learn that—without changing the temperature of the room—my thermometer easily might have said that it is 80 degrees in here. Here I actually am with the thermometer’s report of 70 degrees, but there I easily might have been with equally good evidence that it is 80 degrees. When I learn of this evidential symmetry between my actual self and my nearby counterfactual self, why trust my actual evidence? That this report is actual is not an epistemically respectable tie breaker, if the actual world is an island in a sea of nearby worlds where the thermometer delivers incompatible reports.

So too with the reports of our moral faculties, one might think, especially if our moral faculty was selected for adaptiveness and not truth. Suppose I accept a particular moral code on the basis of my evolved moral faculty but I come to learn that—even holding the moral facts constant—had my social ancestry gone differently, I easily might have accepted a different moral code on that same basis. Wouldn’t it be unacceptably arbitrary or question-begging to nevertheless stick with the actual deliverances of my moral faculty? The fact that this moral faculty of mine is actual cannot properly break the epistemic symmetry, and so—one might think—I ought to give up my moral beliefs.

41. For more on epistemically respectable symmetry breakers, see David Christensen, “Disagreement, Question-Begging, and Epistemic Self-Criticism,” *Philosophers’ Imprint* 11 (2011): 1–22.

This evolutionary debunking argument would be on rather firmer ground than the previous versions we have discussed. For it relies not on sensitivity, safety, or nonaccidentality, but rather on something resembling a conciliatory, “conformist” view of peer disagreement, a view that has become prevalent among those epistemologists working on peer disagreement. And the challenge to moral realism could get its grip without the burden of the dubious claim that adherents of various religions really are epistemic peers. Skepticism looms large if we notice merely that it is unclear whether we are in a better epistemic position than we would have been had we evolved elsewhere, elsewhen.⁴² That far-reaching unclarity should, I believe, be understood to constitute the core of an evolutionary debunking argument at its most threatening.

And so I’d recommend that the argument begin as before:

- (13) Our moral faculty was naturally selected to produce adaptive moral beliefs, and not naturally selected to produce true moral beliefs.
- (14) Therefore, it is false that: *had our species evolved elsewhere, elsewhen and we later formed moral beliefs using the same method we actually used, our beliefs would be true.*

Our beliefs easily might have been incompatible with our actual beliefs, false by our own lights; there’s counterfactual disagreement, near enough to cause trouble. So far, the argument proceeds as before. But our new evolutionary debunking “Argument from Symmetry” might continue not via a safety condition, nor a nonaccidentality condition, but rather as follows:

42. For nice developments of this suggestion for those looking toward disagreement as fuel for skepticism, see Sarah McGrath, “Moral Disagreement and Moral Expertise,” *Oxford Studies in Metaethics* 3 (2008): 87–108; Ballantyne, “The Problem of Historical Variability”; Nathan King, “Disagreement: What’s the Problem? or A Good Peer Is Hard to Find,” *Philosophy and Phenomenological Research* 85 (2012): 249–72; and Sanford Goldberg, “Disagreement, Defeat, and Assertion,” in *Disagreement: New Essays*, ed. David Christensen and Jennifer Lackey (New York: Oxford University Press, 2013), 167–89.

It is reasonable to doubt that mere counterfactual disagreement has the same skeptical force as actual disagreement. But counterfactual disagreement becomes more troubling the closer this disagreement is to actuality, the more easily such disagreement might have happened. (Wouldn’t it bother you to learn that, for example, your calculator easily might have delivered a different answer to the same question?) Indeed, one might think that what’s so troubling about actual disagreement is that it shows just how easily disagreement might happen: it is maximally close disagreement; it so easily could have happened that it did. Reflection on the merely possible vicissitudes of our species’ evolutionary adventure cannot deliver that sort of maximally close disagreement, but it does reveal counterfactual disagreement within shouting distance, close enough to make salient a very troubling kind of epistemic symmetry. This is enough to fuel a debunking argument for moral realism. I am grateful to an anonymous referee of this journal for helpful comments on this section.

- (15) It is unclear whether we are in a better epistemic position than our many nearby counterfactual selves who disagree with us.
 (16) Therefore, our moral beliefs do not count as knowledge.

Or, at least, it would be unacceptably arbitrary or question-begging to insist that they do, in the face of this evidential symmetry. The conclusion makes trouble for moral knowledge, or the propriety of our claim to it, and it plausibly follows, at least for those who are aware of and accept (13)–(15).

To my mind, this is the most plausible evolutionary debunking argument we have yet examined. And it should worry Representationalists. Recall from above that the inference from (13) to (14) is valid on Representationalism. And (15) looks true on Representationalism. If those mental indications (attraction, disgust, etc.) issued by our moral faculty are our primary guides to moral truth, then how could it be clear to us that we are better positioned to answer moral questions than we would have been had our faculties issued different reports, as they easily might have? On Representationalism, we and our nearby counterfactual selves are like those two disagreeing thermometers: we have no independent means by which to see that the reports of our faculties are more reliable than the incompatible reports our faculties would have delivered had we evolved differently. The only means we have to answer moral questions is the faculty that is now called into question. (Worse, while our thermometers are at least aimed at the truth, our moral faculty was selected only for its adaptiveness—even more reason to be conciliatory!) So, on Representationalism, it is unclear that we are now better positioned than we would have been had we evolved differently, just as (15) says. The skeptical considerations that make conciliation attractive are in force here, urging us toward (16): to insist that our moral beliefs do measure up to knowledge, in spite of this evidential symmetry, would be unacceptably arbitrary or question-begging. So, to resist this argument, the Representationalist must defuse these motivations toward a relatively modest form of that prominent conciliatory position in the epistemology of disagreement literature.

Rationalists and Divine Revelationists, on the other hand, need not worry about this argument. For reasons given above, (13) will not yield (14) on either view. And (15) is plausibly false on Rationalism and Divine Revelation, or at least restricted to moral beliefs fueled by sentiment, gut reactions, and so on. For in those other cases—the good cases, on these views—it will be clear that we are in a better position than our disagreeing counterfactual selves and those actual people who disagree with us. On Rationalism, when we form a moral belief via rational insight, we “just see” the truth of our own view, and therefore our superior position, as we might “just see” that we are better positioned than a friend who insists that

$5 + 5 = 55$.⁴³ And for those Divine Revelationists who take their moral judgments to be divinely certified, it is clear to them that they're better positioned than those who disagree, just as it is clear to me that the expert testimony I have received positions me better on some question than my friend who relies on nonexpert testimony.⁴⁴

WHO SHOULD FEAR EVOLUTIONARY DEBUNKING ARGUMENTS?

We completed our tour of evolutionary debunking arguments that dominate the literature and noted that every one is invalid and relies on a dubious epistemic principle like sensitivity, safety, or nonaccidentality. And we have just seen an evolutionary debunking Argument from Symmetry, which relies on a more plausible conciliatory epistemic principle. None of these arguments, however, should worry Rationalists or Divine Revelationists; at best, they concern only Representationalists.

So, one way to rationally maintain moral realism in the face of evolutionary debunking arguments is to accept Divine Revelation. Some skepticism will still be called for, though: the facts of evolution serve Divine Revelationists a defeater for those moral beliefs they take to have been based on the deliverances of their evolved moral faculty (e.g., incest is wrong, rape is wrong, etc.). But Divine Revelationists can in good conscience—and in full knowledge of the facts of evolution—maintain their moral beliefs they take to have been formed via divine revelation (e.g., incest is wrong, rape is wrong, etc.).

Of course, many philosophers are naturalists for whom Divine Revelation cannot give relief from evolutionary debunking arguments.⁴⁵

43. We can safely reject any conciliatory principle that calls for us to reject the deliverances of rational insight. But more likely conciliatory views of disagreement will not call for conciliation in such cases. For more on why, see Tomas Bogardus, "A Vindication of the Equal-Weight View," *Episteme: A Journal of Social Epistemology* 6 (2009): 324–35.

44. Perhaps you suspect that this Argument from Symmetry (and others like it) may be retooled to undermine religious belief, given the highly contingent nature of our religious beliefs. If we'd been born and raised elsewhere, elsewhen, after all, our religious beliefs easily might have been different. But I believe that such arguments pose no serious threat to religious belief. For the reasons, see Bogardus, "The Problem of Contingency for Religious Belief."

45. According to David Bourget and David Chalmers, 49.8 percent of target faculty accept or lean toward naturalism, "What Do Philosophers Believe?" *Philosophical Studies* 170 (2013): 465–500, though one should keep in mind the survey's low response rate and selective sample. And just what's a naturalist? That survey doesn't tell us, but I think Michael Rea had it right when he argued that naturalism is best understood not as an ontology but as a research program, and a naturalist is someone disposed to treat "the methods of science and those alone as basic sources of evidence." See his *World without Design* (New York: Oxford University Press, 2002), 67. These days, those methods are commonly taken to imply or otherwise support the austere ontology of, e.g., Quine: the fundamental physical

What's a naturalist to do? She might turn to Rationalism; we have seen in this article how Rationalism can evade evolutionary debunking arguments. However, while they're logically compatible, Rationalism and naturalism nevertheless have a tense relationship. Philosophers who self-describe as rationalists tend to be nonnaturalists.⁴⁶ Why is that? Perhaps it's the incongruity of intuition pressed into a naturalistic worldview. After all, intuition is meant to involve direct, unmediated access to the truth-values of whole classes of propositions, a *prima facie* mysterious intellectual power in crucial ways unlike any of our sense modalities. Such a power would be very surprising if naturalism were true and much less surprising if naturalism were false.⁴⁷ And so Rationalism counts against naturalism; if true, it disconfirms naturalism. Taking this escape hatch from evolutionary debunking arguments would, therefore, be a considerable cost for the naturalist, a dash from the smoke into the flame.

One final alternative for the naturalist is to maintain Representationalism but to sacrifice moral realism on the altar of Darwinism just as Wilson, Ruse, Joyce, and others have done. Moral antirealism is a minority view among philosophers—27.7 percent of target faculty accept or lean toward it, according to Bourget and Chalmers—and one suspects it fares even worse with the folk.⁴⁸ After all, there is strong empirical evidence that humans come equipped with an unconscious, quick, automatic, noninferential, evolutionarily old faculty for evaluating the world along a moral dimension,⁴⁹ and that these evaluations presuppose moral

entities and forces that the physicists tell us about (fermions, bosons, etc.), no immaterial souls, Platonic forms, and certainly no God, sets are tolerable, numbers are suspect unless they're just sets, etc. While there are self-described naturalists with less austere ontologies, they're widely regarded (even by themselves) as bearing an explanatory burden for their deviance from purer forms of naturalism.

46. See Bourget and Chalmers, "What Do Philosophers Believe?" for data supporting a high correlation between nonnaturalism and rationalism on a recent PhilPapers survey of target philosophy faculty. See also the high correlations between moral antirealism and naturalism, empiricism, physicalism, and atheism. While Rationalism is distinct from Divine Revelationism, neither pairs well with naturalism.

47. For a recent articulation of this line of thought, see Thomas Nagel, *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature Is Almost Certainly False* (New York: Oxford University Press, 2012), who thinks that our powers of cognition comprise one of the main reasons why, as his subtitle proclaims, "the materialist Neo-Darwinian conception of nature is almost certainly false."

48. As Richard Joyce, *The Evolution of Morality*, 209, admits, "I think it is safe to say that [moral non-naturalism and moral supernaturalism] come the closest to capturing what ordinary speakers believe."

49. See Antonio Damasio, *Descartes' Error* (Boston: Norton, 1994); R. A. Shweder and Jonathan Haidt, "The Future of Moral Psychology: Truth, Intuition, and the Pluralist Way," in *Reaching Out: Caring, Altruism, and Prosocial Behavior*, ed. B. Puka (New York: Garland, 1994) 336–41; Greene and Haidt, "How (and Where) Does Moral Judgment Work?"; and John Mikhail, *Elements of Moral Cognition: Rawls' Linguistic Analogy and the Cognitive Science of Moral and Legal Judgment* (New York: Cambridge University Press, 2011).

realism.⁵⁰ This is not to say that our hard-wired inclinations are infallible. It is only to say that moral realism, as we have defined it, has the status of common sense, or at least an attractive starting point. Moral antirealists are pushed there by arguments, not appearances. To give up moral realism, then, is a theoretical cost, a cost proportional to how obvious one initially finds the existence of mind-independent evaluative properties. And many people find that quite obvious, agreeing with Thomas Nagel here and counting it difficult to give up a view on which “pain is really bad, and not just something we hate, and . . . pleasure is really good, and not just something we like.”⁵¹ Moral realism dies hard.

To sum up, we have surveyed several species of evolutionary debunking argument. Nonnaturalists needn’t worry about any: they may (and likely already do) adopt Rationalism or Divine Revelation and sidestep the skeptical threat. But, while most evolutionary debunking arguments fail, the Argument from Symmetry catches naturalists between the devil and the deep sea: they must give up moral realism or accept Rationalism, a heavy price for the naturalist in either direction. And that is why only all naturalists should worry about only one evolutionary debunking argument.

50. See Nichols, “Process Debunking and Ethics”; Shaun Nichols and Trisha Folds-Bennett, “Are Children Moral Objectivists?” *Cognition* 90 (2003): B23–B32; Cecilia Wainryb, Leigh A. Shaw, Marcie Langley, Kim Cottam, and Renee Lewis, “Children’s Thinking about Diversity of Belief in the Early School Years,” *Child Development* 75 (2004): 687–703; as well as Geoffrey Goodwin and John Darley, “The Psychology of Meta-ethics,” *Cognition* 106 (2008): 1339–66.

51. Nagel, *Mind and Cosmos*, 110.