On the Evolutionary Debunking of Morality*

Erik J. Wielenberg

I. INTRODUCTION

Many claim that the availability of evolutionary explanations for human moral beliefs threatens the view that humans have moral knowledge. Peter Singer suggests that evolutionary explanations can debunk moral claims.[1] Michael Ruse declares: "Morality is a collective illusion foisted upon us by our genes."[2] Sharon Street and Richard Joyce have recently offered sustained evolutionary debunkings of morality.[3] Proponents of such debunkings endorse the following thesis:

> The Evolutionary Debunking Thesis (EDT): If S's moral belief that P can be given an evolutionary explanation, then S's moral belief that P is not knowledge.[4]

* A shorter version of this article was presented at the second annual Rocky Mountain Ethics Congress at the University of Colorado in August 2009. I thank the audience and my commentator Brad Monton for helpful feedback on that occasion. Earlier versions of the manuscript were read and commented on by Matthew Braddock, Ben Bradley, David Enoch, Chris Heathwood, and assorted referees and editors for *Ethics*. I thank all of these people for their assistance. This article was written with the support of a DePauw University Faculty Fellowship.

1. Peter Singer, "Ethics and Sociobiology," *Philosophy and Public Affairs* 11 (1982): 40–64.

2. Michael Ruse, *Taking Darwin Seriously* (New York: Blackwell, 1986), 253.

3. Sharon Street, "A Darwinian Dilemma for Realist Theories of Value," *Philosophical Studies* 127 (2006): 109–66; Richard Joyce, *The Evolution of Morality* (Cambridge, MA: MIT Press, 2006).

4. A natural question to ask here is, why is the debunking limited to moral beliefs? The answer is that moral beliefs are seen as particularly susceptible to such debunkings because the evolutionary explanations proposed for them typically do not require that they be true. This aspect of evolutionary debunking arguments is examined in Sec. V.

Ethics 120 (April 2010): 441–464
 0014-1704/2010/12003-0003$10.00

The truth of EDT would obviously have dramatic implications for moral philosophy as well as the moral beliefs and practices of most human beings. If EDT is true, then human moral knowledge is in serious danger of being exposed as fantasy. And if there is no moral knowledge, then, as Jerry Fodor remarked in a different context, "practically everything I believe about anything is false and it's the end of the world."[5]

We may distinguish *metaphysical* debunking arguments and *epistemological* debunking arguments. The former seek to establish EDT by showing that no moral belief that can be given an evolutionary explanation is true. The latter make the case that the existence of an evolutionary explanation for a given moral belief implies that even if the belief is true, it is not knowledge.[6]

In this article, I debunk a variety of epistemological debunking arguments. To accomplish this I first sketch a possible evolutionary explanation for some human moral beliefs (Secs. II and III). Next, I explain how, given a reliabilist approach to warrant, my account implies that humans possess at least some moral knowledge (Sec. IV). In providing this explanation, I assume the truth of certain moral claims. Making such an assumption in this context may appear question-begging. To see that it is not, it is important to keep in mind the thesis that I seek to refute: the thesis that even if there are moral facts, humans lack moral knowledge. I seek to defeat this thesis by sketching an account of how we could have knowledge of certain moral truths, if they hold. After sketching such an account, I identify the basic structure shared by many epistemological evolutionary debunking arguments (Sec. V). Finally, I examine in detail the epistemological debunking arguments of Ruse (Sec. VI), Street (Sec. VII), and Joyce (Sec. VIII). I draw on the account of moral knowledge sketched in Section IV to illustrate how these arguments fail. I do more than debunk the arguments of the would-be debunkers; I also provide a plausible, if incomplete, model of human moral knowledge.

II. ON THE EVOLUTIONARY EXPLANATION OF MORALITY

A creature's evolutionary fitness relative to a given environment is directly proportional to the creature's capacity to transmit its genes on to the next generation.[7] Any trait of a creature that increases its capacity

5. Jerry Fodor, "Making Mind Matter More," in *A Theory of Content and Other Essays* (Cambridge, MA: MIT Press, 1992), 137–60, 156.

6. This distinction between metaphysical and epistemological debunkings is similar to Alvin Plantinga's distinction between de facto and de jure objections to Christian belief; see Alvin Plantinga, *Warranted Christian Belief* (New York: Oxford University Press, 2000), viii–ix.

7. For the point that fitness is always relative to a particular environment, see David Buller, *Adapting Minds: Evolutionary Psychology and the Persistent Quest for Human Nature* (Cambridge, MA: MIT Press, 2005), 28.

(relative to the competition) to pass its genetic material on to the next generation thereby increases that creature's evolutionary fitness. The most obvious way a trait can do this is by increasing the creature's capacity to survive and reproduce in the relevant environment, but this is not the only way. Another way a trait can increase a creature's level of evolutionary fitness is by disposing that creature to promote the survival and reproduction of creatures to which it is genetically related. For instance, the disposition to lay down one's life to save the lives of two creatures that are genetically identical to oneself may enhance a creature's evolutionary fitness even if it does not make the creature in question more likely to survive and reproduce.[8]

A necessary element of an evolutionary explanation for the presence of some trait T in a given population is an account of why, everything else being equal, members of the population who possess T would tend to have higher levels of evolutionary fitness (in the relevant environment) than members of the population who lack T. We may distinguish direct and indirect evolutionary explanations for a trait T. In providing a direct evolutionary explanation for T, we explain how T itself would contribute to the fitness of creatures who have T. In providing an indirect evolutionary explanation for T, we identify some other trait, T′, and explain how T′ would contribute to the fitness of creatures that have T′. We then make the further case that T and T′ are linked in such a way that creatures (in the relevant population) that have T′ are also likely to have T. The fact that T can be given an indirect evolutionary explanation does not imply that T itself enhances the fitness of creatures that have T; T may even decrease the fitness of creatures that have it, as long as T′ is sufficiently fitness-enhancing.

A simple example given by Michael Huemer illustrates this distinction. Let T = the capacity to see distant stars and let T′ = the capacity to see medium-sized objects in one's vicinity. All else being equal, sighted beings are more likely to survive and reproduce than unsighted creatures, so it seems likely that T′ can be given a direct evolutionary explanation. But it is not at all obvious that there are evolutionary benefits in being able to see distant stars. However, creatures with T′ are also likely to have T: "We can see the stars because we have vision, which is useful for seeing things on Earth, and once you

8. This is the central insight involved in Hamilton's concept of inclusive fitness; see W. D. Hamilton, "The Evolution of Altruistic Behavior," *American Naturalist* 97 (1963): 354–56. For useful discussions of this idea, see Elliot Sober and David Sloan Wilson, *Unto Others: The Evolution and Psychology of Unselfish Behavior* (Cambridge, MA: Harvard University Press, 1999), 64–67; and David Buss, *Evolutionary Psychology: The New Science of the Mind* (New York: Pearson Education, 2004), 13–15. When I speak of fitness in this article, I am speaking of inclusive fitness.

have vision, you wind up seeing whatever is there sending light in your direction, whether it is a useful thing to see or not."[9]

One debate regarding evolutionary explanations of morality concerns whether (i) evolutionary explanations can be given merely for human beings' capacity to form moral judgments of some sort or other or (ii) such explanations can be given for at least some of the specific moral judgments that human beings make as well.[10] Furthermore, among those who are sympathetic to evolutionary explanations for some of the specific moral judgments human beings make, there is disagreement about the exact contribution made by nature and nurture. For example, one view is that we are "hardwired" to form particular moral beliefs; another view is that our hardwiring merely provides us with built-in biases toward certain moral beliefs over others.[11]

In what follows I will sketch possible evolutionary explanations for some of our specific moral judgments. Although I will often speak of the evolutionary explanation of moral beliefs, it is probably dispositions to form certain moral beliefs (or, more weakly, to be biased toward certain moral beliefs) that are amenable to evolutionary explanation (if any phenomena in the vicinity are amenable to such explanations).[12] It is important to keep in mind the possibility that different moral beliefs may have different kinds of evolutionary explanations. It is also possible that some moral beliefs have multiple evolutionary explanations. I will not attempt to provide a complete evolutionary explanation for all human moral beliefs. Rather, I will sketch one possible explanation for at least some of our moral beliefs. Even if this account is plausible, it is certainly not a complete explanation for all human moral beliefs, and it may not even be the complete evolutionary story with respect to the moral beliefs that it does purport to explain.

III. THE BASIC ACCOUNT

Let us begin by considering moral beliefs about oneself. Human beings are disposed to view themselves as surrounded by a kind of moral barrier that it is wrong, unjust, evil, or somehow morally inappropriate or il-

9. Michael Huemer, *Ethical Intuitionism* (New York: Palgrave Macmillan, 2005), 216.

10. See, e.g., Giovanni Boniolo, "The Descent of Instinct and the Ascent of Ethics," in *Evolutionary Ethics and Contemporary Biology*, ed. Giovanni Boniolo and Gabrielle De Anna (Cambridge: Cambridge University Press, 2006), 27–42, 28.

11. For part of this debate, see Chandra Sekhar Sripada, "Nativism and Moral Psychology: Three Models of the Innate Structure that Shapes the Contents of Moral Norms," in *Moral Psychology*, vol. 1, *The Evolution of Morality*, ed. Walter Sinnott-Armstrong (Cambridge, MA: MIT Press, 2007), 319–43, together with the ensuing comments and reply (345–65).

12. See Buller, *Adapting Minds*, 53.

legitimate for others to cross.[13] This barrier is conceptualized differently in different cultures. In the West, the barrier is presently conceptualized in terms of rights. Under Islam it is conceptualized in terms of duties.[14] In Neo-Confucianism, it is conceptualized in terms of legitimate desires.[15] Despite various cultural differences, human beings normally believe that there are certain things that others simply ought not do to them, for example, rape them, enslave them, steal from them, or kill them for entertainment.[16] It is not hard to see how the disposition to form such beliefs might be fitness-enhancing. At least part of the explanation for such moral beliefs' ability to do this presumably lies in the fact that they can function as conversation (and deliberation) stoppers.[17] Viewing ourselves as possessing boundaries that may not be transgressed no matter what provides a distinctive kind of motivation to resist such transgressions by others. Holding such beliefs disposes one to resist behavior on the part of others that typically dramatically decreases one's prospects for survival and reproduction.

It is also not hard to see how the disposition to believe that one's kin possess such barriers might be favored by evolutionary forces. Such beliefs will dispose one to act to prevent one's kin from being oppressed or exploited in various ways, and this in turn may enhance one's evolutionary fitness by making it more likely that one's genetic relatives will survive and reproduce.

It is harder to see why the disposition to believe that nonkin are subject to similar moral barriers would be fitness-enhancing. Without taking a position on whether such a disposition is fitness-enhancing, I suggest that this disposition can be given an indirect evolutionary explanation. Consider the principle that things that are alike with respect to their known properties are alike with respect to their unknown properties (call this the *Likeness Principle*). Recognizing this principle, or being disposed to reason in accordance with it, provides many advantages.[18] For example, it helps us to determine which of the many things we might eat will nourish and which will poison us.

13. See Donald Brown, *Human Universals* (Philadelphia: Temple University Press, 1991), 138.

14. Jack Donnelly, *Universal Human Rights in Theory and Practice*, 2nd ed. (Ithaca, NY: Cornell University Press, 2003), 72–73.

15. See Stephen Angle, *Human Rights and Chinese Thought* (New York: Cambridge University Press, 2002), 98–100.

16. See Ellen Messer, "Pluralist Approaches to Human Rights," *Journal of Anthropological Research* 53 (1997): 293–317.

17. See Daniel Dennett, *Darwin's Dangerous Idea: Evolution and the Meanings of Life* (New York: Simon & Schuster, 1995), 506–7; and Joyce, *Evolution of Morality*, 111.

18. Peter Carruthers makes a similar argument with respect to our tendency to employ the inference to the best explanation; see Peter Carruthers, *Human Knowledge and Human Nature* (Oxford: Oxford University Press, 1992), 109–10, 183–87.

When applied to moral properties, the Likeness Principle can generate beliefs that are not necessarily fitness-enhancing. If I believe that I am a bearer of certain moral barriers and that others are similar to me with respect to their known properties, I am disposed to form the belief that those others possess similar moral barriers. An example given by Kwame Appiah illustrates this idea:

> Think of the icon, devised by a Quaker abolition society in the eighteenth century, depicting a shackled slave, kneeling and with hands raised in supplication, with the legend "Am I not a man and a brother?" It was not a moral treatise, but it had an argument to make: see me as someone like yourself.[19]

Peter Singer has likened reason to "an escalator" that can lead us "to places that are not of any direct advantage to us, in evolutionary terms."[20] One way reason might do this is by using the Likeness Principle to lead us to conclude that nonkin possess the same moral barriers as ourselves. Because reasoning in accordance with the Likeness Principle benefits us across a variety of contexts, we are disposed to reason in accordance with it in moral contexts as well, where it may not serve our interests so well.[21] Thus, Steven Pinker may be on the right track when he claims that "no creature equipped with the circuitry to understand that it is immoral for you to hurt me could discover anything but that it is immoral for me to hurt you."[22]

Many note the "in-group, out-group" aspect of human morality.[23] As Jack Donnelly observes, "in their past, *all* major regional civilizations have at times been dominated by views that treated some significant portion of human beings as 'outsiders' . . . not entitled to guarantees that could be taken for granted by 'insiders'."[24] An important element of this aspect of morality is that those belonging to the out-group are seen as being different from those in the in-group not just with respect to their moral properties but with respect to their nonmoral properties as well. In fact, those in the out-group are often viewed as less than fully

19. Kwame Anthony Appiah, *Experiments in Ethics* (Cambridge, MA: Harvard University Press, 2008), 160.

20. Peter Singer, "Morality, Reason, and the Rights of Animals," in *Primates and Philosophers: How Morality Evolved*, ed. Frans de Waal (Princeton, NJ: Princeton University Press, 2006), 140–60, 146.

21. The Likeness Principle could similarly "expand outward" various moral claims—e.g., my pain is bad; therefore, other people's pain is bad.

22. Steven Pinker, "Evolution and Ethics," in *Intelligent Thought: Science versus the Intelligent Design Movement*, ed. John Brockman (New York: Vintage, 2006), 142–52, 151.

23. See, e.g., Appiah, *Experiments in Ethics*, 142–45; Brown, *Human Universals*, 138–39; de Waal, *Primates and Philosophers*, 53; and Michael J. Perry, *The Idea of Human Rights: Four Inquiries* (Oxford: Oxford University Press, 1998), 58–60.

24. Donnelly, *Human Rights*, 41.

human, as nonpersons. Large-scale exploitation and genocide typically include dehumanization. The oppressors have an ideology according to which their victims are not fully human.[25] Perhaps this ideology is required in order to prevent the oppressors from reaching the conclusion that their victims possess the same moral barriers as themselves. Apparently it is extremely psychologically difficult for human beings to believe that others have different moral properties from themselves while also believing that those others are similar to themselves with respect to their nonmoral properties.[26] This constitutes some (admittedly weak) empirical evidence in support of the proposal I have sketched here. It suggests that we are inclined to extend the moral barrier to those we know to be similar to ourselves in nonmoral ways whether we want to do so or not. In cases where we are strongly motivated to exploit other human beings, we need an ideology according to which those we wish to exploit are not fully human in order to make the exploitation psychologically possible.

IV. MORAL KNOWLEDGE

In this section I will make the case that if we have moral barriers or rights of the sort sketched in the preceding section, then there is a plausible explanation of how we could have knowledge of these rights. That we have such rights is an assumption; I offer no argument for the existence of such rights here. The reason I take this approach is that I am concerned here exclusively with rebutting epistemological debunking arguments. Such arguments are not aimed at showing that there are no moral truths. Rather, such arguments are aimed at showing that even if there are moral truths, human beings lack knowledge of such truths. In arguing against this conditional claim, it is not question-begging to assume the truth of its antecedent (that there are moral truths). That is the strategy I pursue in this section.

I take it that being produced by reliable processes confers warrant on beliefs and that sufficiently reliable processes can confer a degree of warrant sufficient for knowledge.[27] I will make the case that the model

25. See, e.g., Jonathan Glover, *Humanity: A Moral History of the Twentieth Century* (New Haven, CT: Yale University Press, 2000); and Philip Zimbardo, *The Lucifer Effect: Understanding How Good People Turn Evil* (New York: Random House, 2007).

26. See Martha Nussbaum, "Human Functioning and Social Justice: In Defense of Aristotelian Essentialism," *Political Theory* 20 (1992): 202–46.

27. I mean 'warrant' to indicate the feature of beliefs, whatever it is, that makes the difference between mere true belief and knowledge. The claim that reliability is sufficient for warrant is, of course, controversial. However, David Enoch has argued that the heart of the challenge to moral knowledge posed by evolutionary debunkers is that of explaining the correlation between normative truths and our normative judgments; see David Enoch, "The Epistemological Challenge to Metanormative Realism: How Best to Understand It, and How to Cope with It," *Philosophical Studies* 148 (2010): 413–38. The model I propose

sketched in the previous section implies (given the assumption of the reality of rights) that at least some human moral beliefs are formed by significantly reliable processes. An important feature of this model is that at least some of the beliefs produced by way of reliable processes need not be inferred from other things the agent believes. Such beliefs are justified basic beliefs, and the brand of reliabilism I am employing here is therefore a kind of foundationalism.[28]

Consider a typical Westerner's belief that she herself has the right not to be tortured for entertainment. The model sketched above says little about the precise nature of the proximate processes by which this belief is produced; the model is compatible with a wide range of possible proximate processes. To see this, consider two simple possibilities. One possibility is that the process is intuition-driven. A number of thinkers suggest that at least some moral truths are self-evident and can simply be "seen" to be true in a direct way.[29] C. S. Lewis colorfully puts it this way: "The primary moral principles on which all others depend are rationally perceived. . . . Their intrinsic reasonableness shines by its own light."[30]

This proposal is certainly true to the phenomenology of the formation of some moral beliefs. In this respect there is a similarity between some moral beliefs and some mathematical beliefs. Consider Kurt Gödel's remarks about mathematical knowledge: "We do have something like a perception . . . of the objects of set theory, as is seen from the fact that the axioms force themselves upon us as being true."[31] At least some moral claims similarly force themselves upon us; that one has the right not to be tortured for entertainment is a likely candidate for such a claim. Michael Huemer suggests that in these sorts of cases, we have an intuition (an intellectual appearance) that the moral claim is true and we come to believe the moral claim on the basis of this intuition.[32]

Some thinkers, including Huemer, maintain that the intuition itself

---

in this article, if successful, answers this challenge. Thus, even if reliabilism is false, my model is still helpful in addressing the challenge of the evolutionary debunkers.

28. See Richard Feldman, *Epistemology* (Upper Saddle River, NJ: Pearson Education, 2003), 50–51.

29. See, e.g., G. E. Moore, *Principia Ethica* (Cambridge: Cambridge University Press, 1903), 143; C. S. Lewis, *Miracles: A Preliminary Study* (New York: HarperCollins, 2001), 54; William Lane Craig, "Reply to Objections," in *Does God Exist? The Craig-Flew Debate*, ed. Stan Wallace (Burlington, VT: Ashgate, 2003), 155–88, 172; J. P. Moreland, *Scaling the Secular City: A Defense of Christianity* (Grand Rapids, MI: Baker Books, 1987), 115–17; Russ Shafer-Landau, *Moral Realism: A Defence* (Oxford: Oxford University Press, 2003), 247–66; and Michael Huemer, *Ethical Intuitionism* (New York: Palgrave Macmillan, 2005).

30. Lewis, *Miracles*, 54.

31. Quoted in Leonard Wapner, *The Pea and the Sun* (Wellesley, MA: Peters, 2005), 33.

32. Huemer, *Ethical Intuitionism*, 99–104.

justifies the corresponding moral belief. I will not take a stand on this issue. Regardless of whether intuitions themselves confer justification, it can be the case that some intuition-triggered moral beliefs are produced by reliable processes and that being so produced confers warrant on them.[33] My focus here will be on reliability as a source of warrant, though this focus should not be taken to mean that I reject the possibility of other sources of warrant.

A second possibility is that the proximate process is largely emotion-driven. For instance, perhaps the process involves the disposition to become outraged when treated in certain ways by others (or when facing a threat of being so treated), and this outrage causes one to form the belief that one has certain rights (and that they are, or are in danger of, being violated). The model sketched above is compatible with both of these possibilities, as well as with a host of more complicated options.[34]

However, it is clear enough that whatever the precise nature of the process, it is one that requires the presence of certain cognitive capacities for its occurrence. In order to form the belief that one has certain rights, one must be able to have some grasp of the concept of rights. While there are various theories about the foundation of rights, it is widely agreed that if rights exist at all, their presence is guaranteed by the presence of certain cognitive faculties.[35] The cognitive faculties in question are either the very ones required to form beliefs about rights or are closely linked to such faculties. If you think you possess moral barriers, then you do (assuming such barriers exist at all). Therefore, assuming that rights are real, the processes that ultimately generate, say, the belief that one has a right not to be tortured just for fun are significantly reliable, at least with respect to beliefs of the relevant type. Sufficiently cognitively developed creatures that are products of evolution will possess moral barriers (if such barriers are real) and will also be disposed to believe that they have such barriers. The very cognitive

33. This possibility is suggested by Shafer-Landau (*Moral Realism*, 280 n. 13) and Michael Huemer ("Compassionate Phenomenal Conservatism," *Philosophy and Phenomenological Research* 74 [2007]: 30–55).

34. For useful discussions of some of the various proximate models on offer, see Marc Hauser, *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong* (New York: HarperCollins, 2006), particularly 1–55; Marc Hauser, Liane Young, and Fiery Cushman, "Reviving Rawls's Linguistic Analogy: Operative Principles and the Causal Structure of Moral Actions," in *Moral Psychology*, vol. 2, *The Cognitive Science of Morality: Intuition and Diversity*, ed. Walter Sinnott-Armstrong (Cambridge, MA: MIT Press, 2008), 107–43; and Jonathan Haidt and Fredrik Bjorklund, "Social Intuitionists Answer Six Questions about Moral Psychology," in Sinnott-Armstrong, *The Cognitive Science of Morality*, 181–217.

35. For an overview of the main theories, see Elizabeth Zechenter, "In the Name of Culture: Cultural Relativism and the Abuse of the Individual," *Journal of Anthropological Research* 53 (1997): 319–47. It should be noted that I am proposing only a sufficient condition for the possession of a moral barrier, not a necessary condition.

faculties that lead such beings to believe that they possess moral barriers also entail the presence of those very barriers.

The model I propose here is similar to one presented by David Enoch. Enoch offers what he calls a "third-factor explanation" to account for the correspondence between moral facts and human moral beliefs.[36] He explains: "It is possible that the explanation of a correlation between the two factors A and B is in terms of a third factor, C, that is (roughly speaking) responsible both for A-facts and for B-facts."[37] This description captures the essence of my own strategy, though the third factor I identify is different from Enoch's third factor. Enoch's third factor is the goodness of survival or reproductive success.[38] My third factor is certain cognitive faculties. Such faculties are responsible for the presence of moral rights in that the presence of the relevant faculties entails the presence of rights. Thus, the connection between the cognitive faculties and moral rights is logical. These same faculties also generate moral beliefs, including the relevant beliefs about rights. The connection between the cognitive faculties and beliefs about moral rights is causal. In this way, the relevant cognitive faculties are responsible for both moral rights and beliefs about those rights, and so the cognitive faculties explain the correlation between moral rights and beliefs about those rights.

Do evolutionary processes tend to produce true beliefs in general? If not, things become messier for my approach, since we would face a situation in which evolutionary processes are reliable with respect to a certain class of moral beliefs but not reliable more generally. On the other hand, if evolutionary processes are generally reliable, then we seem to be on firm ground in claiming that according to my model, at least some moral beliefs are produced by reliable processes.

Peter Carruthers has argued that there is good reason to believe that evolutionary processes are generally reliable belief-forming processes (when they produce beliefs at all). Carruthers offers two main considerations to support this view. First, "organisms (of the sort that act on beliefs) will only survive, in general and in the long run, if they base their actions on beliefs that are true, or at least close to the truth."[39] Second, many of the proximate belief-forming mechanisms that evolutionary forces have bestowed upon us are generally reliable; Carruthers offers perception and memory as examples. The idea here is that evolutionary

36. Enoch, "Epistemological Challenge," 429. I came across Enoch's paper late in the process of writing this article. Thus, despite the similarities between our two strategies, my proposal was developed largely independently of Enoch's.

37. Ibid., 429–30.

38. Ibid., 430.

39. Carruthers, *Human Knowledge*, 112.

processes are reliable in virtue of the fact that the proximate belief-forming processes that occur as a result of evolutionary processes are generally reliable.[40] So, assuming that rights (or something like them) are real, there is good reason to believe both that (i) when evolutionary processes produce beings that think they have rights, those beings generally do have rights and (ii) when evolutionary processes produce other kinds of beliefs, those beliefs tend to be true. Evolutionary processes are thus generally reliable and, according to my model, are also reliable with respect to the particular case of beliefs regarding one's own moral barriers. Therefore, if my model is correct, at least some moral beliefs are produced by reliable processes and hence possess warrant.

To this point we have been considering the reliability of beliefs about one's own rights. Let us turn now to consideration of moral beliefs about the moral barriers of beings distinct from oneself. According to the model sketched above, at least some such beliefs are generated by the disposition to reason in accordance with the principle that entities that are alike with respect to their known properties are alike with respect to their unknown properties (the Likeness Principle). Given the laws of nature that hold in our universe, this belief-forming process has a high degree of conditional reliability. That is, when it takes true beliefs as inputs, it tends to produce true beliefs as outputs.[41] So, suppose that a being who believes correctly that she has certain rights encounters a second being similar to herself with respect to its known properties. If the process of reasoning in accordance with the Likeness Principle leads her to form the belief that the second being has the same rights that she has, then this belief is formed by a conditionally reliable process operating on true beliefs.

Therefore, if we assume that beliefs that are produced by highly reliable processes (including highly conditionally reliable processes operating on true beliefs) have a degree of warrant sufficient for knowledge, then the model sketched in the previous section implies that at least some moral beliefs constitute knowledge (again, assuming that moral barriers exist). If the evolutionary explanation of our beliefs about moral barriers sketched earlier is correct, and those barriers are real, then many of us know that we (and others) possess such barriers. My account does not assume any particular metaethical theory about the nature of these moral barriers beyond that they are real and that their

40. Ibid., 113.

41. For this understanding of conditional reliability, see Alvin Goldman, "A Priori Warrant and Naturalistic Epistemology: The Seventh Philosophical Perspectives Lecture," *Noûs* 33, suppl., "Philosophical Perspectives: Epistemology" (1999): 1–28; and Feldman, *Epistemology*, 93.

presence is entailed by the presence of certain cognitive faculties. For all I have said, it could be that facts about these barriers turn out to be natural facts, supernatural facts, or nonreducible sui generis moral facts.

With this account of moral knowledge in hand, we are in a position to see why epistemological evolutionary debunking arguments fail. In the next section, I explain the basic structure of such arguments and describe my general strategy for resisting them. In subsequent sections, I examine the arguments of Ruse, Street, and Joyce in detail.

V. THE BASIC EVOLUTIONARY DEBUNKING ARGUMENT

Recall the thesis presented at the beginning of this article :

> The Evolutionary Debunking Thesis (EDT): If S's moral belief that P can be given an evolutionary explanation, then S's moral belief that P is not knowledge.

Many would-be debunkers emphasize that evolutionary explanations of moral beliefs imply that the beliefs in question can be explained without appealing to the truth of those beliefs.[42] Let us suppose that this is correct; it is certainly correct in the case of the evolutionary story sketched in Section III of this article. That account does not require the truth of the moral beliefs in question (though of course it is consistent with their truth).

Gilbert Harman was perhaps the first contemporary philosopher to outline a case against moral knowledge based on the claim that human moral beliefs can be explained without appealing to any moral truths.[43] His worry was roughly this: if moral facts do not explain moral beliefs, then they do not explain anything at all, and if they do not explain anything at all, then we should conclude that they do not exist.[44] If there are no moral facts, then there is no moral knowledge.

This argument has been the subject of much debate. I agree with Russ Shafer-Landau's view that this debunking rests on an implausible criterion for existence (what Shafer-Landau calls the "strong causal test of ontological credibility").[45] However, as this argument is a metaphysical

42. See, e.g., Ruse, *Taking Darwin Seriously*, 252; and Joshua Greene, "From Neural 'Is' to Moral 'Ought': What Are the Moral Implications of Neuroscientific Moral Psychology?" *Nature Reviews: Neuroscience* 4 (2003): 847–50; Joyce, *Evolution of Morality*, 183; and Michal Ruse, "Is Darwinian Metaethics Possible (and If It Is, Is It Well Taken)?" in Boniolo and Anna, *Evolutionary Ethics*, 13–26, 23.

43. See Gilbert Harman, *The Nature of Morality: An Introduction to Ethics* (Oxford: Oxford University Press, 1977), 3–26.

44. Joyce sketches this line of reasoning; see Joyce, *Evolution of Morality*, 186–87.

45. See Shafer-Landau, *Moral Realism*, 110–14.

rather than epistemological debunking and my focus here is on epistemological debunkings, I will put this argument aside.

Still, many epistemological evolutionary debunking arguments can be understood as variations on Harman's basic idea. Such arguments can be understood as different ways of supporting the following line of reasoning:

The Basic Evolutionary Debunking Argument

P1. If S's moral belief that P can be given an evolutionary explanation, then S's moral belief that P can be explained without appealing to the truth of P.

P2. If S's moral belief that P can be explained without appealing to the truth of P, then S's moral belief that P lacks warrant.

P3. If S's moral belief that P lacks warrant, then S's moral belief that P is not knowledge.

C. Therefore, if S's moral belief that P can be given an evolutionary explanation, then S's moral belief that P is not knowledge (EDT).

I deny P2 on the grounds that (i) being produced by a sufficiently reliable process can bestow warrant on a belief sufficient for knowledge and (ii) a belief that P can be produced by a sufficiently reliable process even when the belief can be explained without appealing to the truth of P. The account sketched earlier in this article shows that (ii) is true. Claim (i) is more controversial, but it shows at least that a defender of P2 must engage with certain debates in epistemology in order to establish P2.

Joshua Greene has remarked: "Understanding how we make moral judgements might help us to determine whether our judgements are perceptions of external truths or projections of internal attitudes."[46] The remark suggests that there are just two alternatives: either (*a*) our moral beliefs are perceptions of (and presumably caused by) moral facts and so may constitute knowledge or (*b*) our moral beliefs are mere projections of our attitudes and so do not constitute knowledge. I favor a third alternative: (*c*) despite not being caused by moral facts, at least some of our moral beliefs are reliably formed and hence may constitute knowledge.[47] Explaining how *c* might be the case was one of the central burdens of the earlier sections of this article. As we will see, this basic approach offers a way of resisting the debunkings of Ruse, Street, and Joyce. We begin with Ruse.

46. Greene, "From Neural 'Is' to Moral 'Ought'," 848.

47. The possibility of reliability without causation when it comes to moral facts and beliefs is also defended by Colin McGinn in *Ethics, Evil, and Fiction* (Oxford: Oxford University Press, 1997), 43; and Shafer-Landau in *Moral Realism,* 275.

VI. MICHAEL RUSE'S EVOLUTIONARY DEBUNKING OF MORALITY

The following remarks by Ruse suggest an argument for P2:

> You would believe what you do about right and wrong, irrespective of whether or not a "true" right and wrong existed! The Darwinian claims that his/her theory gives an entire analysis of our moral sentiments. Nothing more is needed. Given two worlds, identical except that one has an objective morality and the other does not, the humans therein would think and act in exactly the same ways.[48]

The first sentence of this passage suggests the following principle:

> (1) IF S's moral belief that P can be explained without appealing to the truth of P, THEN: If P were false, then S would still believe P.

If we combine (1) with

> (2) IF: If P were false, then S would still believe P, THEN: S's belief that P lacks warrant,

we can infer

> (P2) If S's moral belief that P can be explained without appealing to the truth of P, then S's moral belief that P lacks warrant.

It is widely held that if moral properties are exemplified at all, they supervene on nonmoral properties.[49] This supervenience thesis entails that any two possible worlds that are identical with respect to their nonmoral properties are identical with respect to their moral properties. If this supervenience thesis is correct, then Ruse's argument fails. The central task of the rest of this section is to explain how the supervenience of the moral on the nonmoral makes trouble for Ruse's reasoning in the passage quoted above.

Let us define a *Ruse world* as a possible world in which there are no objective ethical facts and which is otherwise as similar as possible to the actual world. Ruse's claim is that in the nearest Ruse world, we have exactly the same moral beliefs that we have in the actual world.

To evaluate this claim, let us initially suppose that our world does include some objective ethical facts; specifically, let us suppose that it

48. Ruse, *Taking Darwin Seriously*, 254; also see Richard Joyce, *The Myth of Morality* (New York: Cambridge University Press, 2001), 163. It is possible to interpret Ruse as advancing a metaphysical debunking argument here. The essence of that argument is that since we do not need to posit moral facts to explain our moral sentiments, we should conclude that there are no moral facts. I think this sort of argument is not successful, but a full evaluation of this argument falls outside the scope of the present article.

49. See, e.g., Huemer, *Ethical Intuitionism*, 202 and 280 n. 7; Shafer-Landau, *Moral Realism,* 77–78; and Graham Oppy, *Arguing about Gods* (Cambridge: Cambridge University Press, 2006), 353.

includes at least the existence of the moral barriers discussed earlier. This assumption, together with the supervenience thesis, implies that the nearest Ruse world is not identical to the actual world with respect to its nonmoral properties. Drawing on the account of moral knowledge sketched in Section IV above, we can see what some of these nonmoral differences might be. In Section IV, I suggested that if you think you possess moral barriers, then you do (assuming such barriers exist at all). So, in the actual world, I believe that I have certain moral rights and (in accordance with our assumption that some moral properties are exemplified in the actual world) I do have such rights. Given supervenience, in the nearest world in which I lack those rights, I lack the cognitive capacities to grasp such concepts as selfhood and rights. This means that in the nearest Ruse world I lack moral beliefs altogether. Thus, there is good reason to reject Ruse's claim that I "would believe what [I] do about right and wrong, irrespective of whether or not a 'true' right and wrong existed." Given the supervenience of the moral on the nonmoral and the evolutionary account sketched above, there are at least some moral beliefs for which (1) does not hold. Assuming that I actually possess the right not to be tortured merely for entertainment, it follows that if I lacked this right, I would similarly lack the belief that I have such a right. Premise (1) above is false.

There is another way of understanding what Ruse is up to in the passage quoted above. Consider again the last sentence of that passage: "Given two worlds, *identical* except that one has an objective morality and the other does not, the humans therein would think and act in exactly the same ways."[50] I defined a Ruse world as a possible world in which there are no objective ethical facts and which is otherwise as similar as possible to the actual world. Let us define a *Ruse* world* as a world in which there are no objective ethical facts and which is otherwise identical to the actual world.[51] Perhaps Ruse intends to claim that in the nearest Ruse* world, we have exactly the same moral beliefs that we have in the actual world. This claim seems true. However, its truth does not support (1) but rather a somewhat different claim:

> (1*) IF S's moral belief that P can be explained without appealing to the truth of P, THEN: If P were false *and everything else was the same*, then S would still believe P.

To move from (1*) to (P2), we need the following revised version of (2):

50. Ruse, *Taking Darwin Seriously*, 254 (emphasis added).

51. Presumably this claim should be qualified so as not to stipulate that we have all the same moral beliefs in a Ruse* world that we have in the actual world; otherwise the claim is trivial.

(2*) IF: If P were false *and everything else was the same*, then S would still believe P, THEN: S's belief that P lacks warrant.

While the problem with the argument that has (1) and (2) as its premises is that (1) is false, the problem with the argument that has (1*) and (2*) as its premises is that (2*) is false. Assuming as before that nihilism is false in the actual world and that the moral supervenes on the nonmoral, the nearest Ruse* world is in fact an impossible world. This is so because the nearest Ruse* world is a world that is identical to the actual world with respect to its nonmoral properties but differs with respect to its moral properties. It is hard to see why the fact that we have false moral beliefs in an impossible world should threaten the idea that we have moral knowledge in the actual world. My model of moral knowledge implies that my belief that I have rights tracks the truth in nearby worlds, and that is enough. That there is an impossible world in which I lack rights but still believe that I have them is no threat to my knowing that I have rights in the actual world.[52]

Therefore, each of the two interpretations of Ruse's reasoning that I have considered fails. The first interpretation fails because (1) is false; the second fails because (2*) is false. Given the supervenience of the moral on the nonmoral, the claim that I would believe I have rights even if I lacked them is false. It is true that I would believe I had rights if I lacked rights and everything else was the same, but this claim in no way supports the conclusion that I do not know that I have rights in the actual world.

Ruse could perhaps salvage his argument by helping himself to either the assumption that nihilism is true in the actual world or the assumption that the moral does not supervene on the nonmoral. But there is little reason for a critic of epistemological debunking arguments to grant either of these controversial assumptions.

VII. SHARON STREET'S EVOLUTIONARY DILEMMA

Unlike Ruse, Sharon Street does not seek to provide an unconditional debunking of morality. Her intended conclusion is rather that any realist theory of value (a theory according to which there are at least some evaluative truths that hold independently of all our evaluative attitudes) is subject to this sort of debunking, whereas at least some nonrealist theories of value cannot be so debunked.[53] Because I intend the account of moral knowledge sketched earlier to be compatible with realist theories of value, I must address Street's challenge. The following para-

52. This discussion of the second interpretation of Ruse's argument draws on some points made by Keith DeRose; see Keith DeRose, "Solving the Skeptical Problem," *Philosophical Review* 104 (1995): 1–52.

53. Street, "Darwinian Dilemma," 110–12, 152–56.

graph from the concluding section of her paper offers the clearest statement of her overall argument:

> Ultimately, the fact that there are *any* good scientific explanations of our evaluative judgments is a problem for the realist about value. It is a problem because realism must either view the causes described by these explanations as distorting . . . or it must enter into the game of scientific explanation, claiming that the truths it posits actually play a role in the explanation in question. The problem with the latter option, in turn, is that they don't. The best causal accounts of our evaluative judgments, whether Darwinian or otherwise, make no reference to the realist's independent evaluative truths.[54]

The two horns of Street's dilemma are (i) accept that all human moral beliefs can be explained without appealing to their truth or (ii) reject this claim. The problem with (ii), according to Street, is that it is implausible on scientific grounds. The alleged problem with (i) is that it implies moral skepticism. I shall take the first horn of this dilemma.

Street's idea is that if our moral beliefs can be given evolutionary explanations and there is no causal connection between our moral beliefs and objective moral facts, then "the forces of natural selection must be viewed as a purely distorting influence on our evaluative judgments."[55] She argues for this as follows:

> On this view, allowing our evaluative judgments to be shaped by evolutionary influences is analogous to setting out for Bermuda and letting the course of your boat be determined by the wind and tides: just as the push of the wind and tides on your boat has nothing to do with where you want to go, so the historical push of natural selection on the content of our evaluative judgments has nothing to do with evaluative truth. . . . Of course it's *possible* that as a matter of sheer chance, some large portion of our evaluative judgments ended up true, due to a happy coincidence between the realist's independent evaluative truths and the evaluative directions in which natural selection tended to push us, but this would require a fluke of luck that's not only extremely unlikely, in view of the huge universe of logically possible evaluative judgments and truths, but also astoundingly convenient to the realist.[56]

54. Ibid., 155.

55. Ibid., 121.

56. Ibid., 121–22. A similar argument is suggested by Joshua Greene's remarks about "a hodgepodge of evolutionary forces" in Joshua D. Greene, "The Secret Joke of Kant's Soul," in *Moral Psychology*, vol. 3, *The Neuroscience of Morality: Emotion, Brain Disorders, and Development*, ed. Walter Sinnott-Armstrong (Cambridge, MA: MIT Press, 2008), 35–80, 72.

The heart of Street's reasoning in this passage is that if moral facts do not explain the moral beliefs of human beings, then those beliefs being correct would involve a lucky coincidence that is incompatible with genuine knowledge. This suggests that Street's reasoning depends on the following pair of principles:

> (3) If S's moral belief that P can be explained without appealing to the truth of P, then S's moral belief that P being true would be a lucky coincidence.
>
> (4) If S's moral belief that P being true would be a lucky coincidence, then S's moral belief that P lacks warrant.

From (3) and (4) we can infer:

> (P2) If S's moral belief that P can be explained without appealing to the truth of P, then S's moral belief that P lacks warrant.

The weak link of this argument is (3). The model of moral knowledge I sketched earlier shows how a moral belief that P can be true when the belief is not explained by P without the truth of the belief being a lucky coincidence. According to that model, the presence of the very cognitive faculties that cause (or at least causally contribute to) my belief that I have certain rights also entails that I have those very rights. Furthermore, the process that produces this belief is a reliable one, and hence the belief constitutes knowledge. Something similar is true of my belief that other people have the same rights that I have. Those others have certain nonmoral features that make them relevantly like me. Those features both causally contribute to my belief that they have rights and entail that they have those same rights. And, once again, the process that produces this belief is a reliable one.

Street's mention of a "huge universe of logically possible evaluative judgments and truths" conjures up the notion of two sets, one whose members are all the logically possible moral beliefs and the other whose members are all the logically possible moral claims, with some members of each set being selected at random.[57] Given the large number of members in each set, it is improbable that the selected members of each set will correspond such that for each randomly selected moral belief that P it will also be the case that the moral claim that P is randomly selected.

My model of moral knowledge shows that this way of looking at things is not entailed by the presence of evolutionary explanations for human moral beliefs. In my model, certain nonmoral features of the world both entail certain moral facts and causally contribute to the

57. Street, "Darwinian Dilemma," 122.

presence of moral beliefs that correspond to those moral facts. On this model, it is not at all unlikely that moral beliefs and moral facts will correspond. Given the truth of my model, it would be true to say, as Steven Pinker does, that "our moral sense has evolved to mesh with an intrinsic logic of ethics."[58]

David Enoch suggests that the fundamental challenge Street's argument poses for the normative realist "is that of coming up with an explanation of a correlation between our relevant beliefs and the relevant truths."[59] Some comments Street makes in a later paper support this interpretation. For instance, in the later paper Street declares that "the [Darwinian] dilemma notes the striking coincidence between the independent normative truths posited by the realist and the normative views that evolutionary forces pushed us toward, and challenges the realist to explain this coincidence."[60] The realist is hamstrung in this task by the fact that there is no causal connection between moral facts and moral beliefs. As I noted earlier, Enoch and I both propose a "third-factor" explanation to account for the correspondence between moral facts and human moral beliefs. My third factor is certain cognitive faculties: the relevant cognitive faculties secure a correlation between moral rights and beliefs about moral rights because they entail the presence of moral rights and generate beliefs about such rights.[61]

It might be suggested that this explanation itself involves an unlikely coincidence. Specifically, we might wonder about the remarkable fact that the presence of the cognitive faculties that cause me to believe that I have rights also entails that I do have those rights. Isn't this an improbable fluke?[62]

Consider the facts that generate the coincidence. On the one hand are the facts about which moral properties supervene on which nonmoral properties. On the other hand are facts about what causes what.

58. Pinker, "Evolution and Ethics," 151.

59. Enoch, "Epistemological Challenge," 426.

60. Sharon Street, "Reply to Copp: Naturalism, Normativity, and the Varieties of Realism Worth Worrying About," *Philosophical Issues* 18 (2008): 207–28.

61. In her reply to Copp, Street distinguishes between "internal explanations of reliability that provide *good (internal) reason* to think that we are reliable, and internal explanations of reliability that provide *no good reason whatsoever* to think that we are reliable" (Street, "Reply to Copp," 216). Street declares that explanations of the second sort are of no use in answering the challenge posed by the Darwinian Dilemma. I believe that my explanation falls into the first category. I do not merely assert the truth of our commonsense moral beliefs and then simply note that these are the very views evolutionary forces that pushed us toward (see ibid., 214). Instead, I seek to explain the reliability of our moral beliefs (about rights) by positing a connection between moral rights and our beliefs about such rights that is grounded in our cognitive faculties.

62. Enoch (independently) poses the same sort of objection to his own explanation; see Enoch, "Epistemological Challenge," 433.

It is plausible that these causal facts depend in turn on what the laws of nature are. So the alleged coincidence that we are investigating is generated by (i) the existence of certain supervenience relationships and (ii) the obtaining of certain laws of nature. Does the correspondence between these two groups of facts require an explanation and, if so, what might that explanation be?

There are various possible answers to this question. One is that there is a personal God who intentionally set things up so that the correlation between moral beliefs and facts posited by my model would hold.[63] This God would establish both the supervenience relationships and the laws of nature, seeing to it that the two correspond so as to produce sufficiently reliable moral beliefs on the part of human beings. A second possibility is that instead of a personal God, there is a fundamental truth or law in our universe that implies that the supervenience relationships and the laws of nature will correspond in the right way. Perhaps, as Thomas Nagel puts it, "the capacity of the universe to generate organisms with minds capable of understanding the universe is itself somehow a fundamental feature of the universe."[64]

A third possibility is that the relevant supervenience relationships and laws of nature are brute facts that have no further explanation. This is perhaps most plausible in the case of the supervenience relationships as these are necessary relationships. I have argued elsewhere that the view that these supervenience relationships have no deeper explanation is no less plausible than the view that their truth is somehow grounded in or secured by a personal God.[65] The issue of the laws of nature is a bit trickier. However, some have suggested that the most fundamental laws of nature are brute features of the universe; thus, physicist Steven Weinberg defends "the ancient search for those principles that cannot be explained in terms of deeper principles."[66] And some have suggested that the laws of nature are in fact necessary truths.[67] If the fundamental laws of nature are brute, necessary truths, then it follows that neither the supervenience relationships nor the laws of

63. For suggestions along these lines, see, e.g., Robert Adams, *Finite and Infinite Goods* (New York: Oxford University Press, 1999), 70; Larry Arnhart, "The Darwinian Moral Sense and Biblical Religion," in *Evolution and Ethics: Human Morality in Biological and Religious Perspective*, ed. Philip Clayton and Jeffrey Schloss (Grand Rapids, MI: Eerdmans, 2004), 204–20; and William Wainwright, *Religion and Morality* (Burlington, VT: Ashgate, 2005), 66–67.

64. Thomas Nagel, *The Last Word* (New York: Oxford University Press, 1997), 132.

65. Erik Wielenberg, "In Defense of Non-natural, Non-theistic Moral Realism," *Faith and Philosophy* 26 (2009): 23–41.

66. Steven Weinberg, *Dreams of a Final Theory* (New York: Pantheon, 1992), 18.

67. See, e.g., Alexander Bird, "The Dispositionalist Conception of Laws," *Foundations of Science* 10 (2005): 353–70.

nature could have been different than they actually are, in which case the correspondence posited by my model is also a necessary feature of the universe. If this is the case, then the correspondence in question can hardly be characterized as resting on a lucky coincidence: where there is no contingency, there are no coincidences.

Each of these proposals is somewhat speculative; however, it is clear enough that Street's reasoning, as it stands, does not establish her intended conclusion. Street's claim that "realism must either view the causes described by [evolutionary] explanations as distorting . . . or it must enter into the game of scientific explanation, claiming that the truths it posits actually play a role in the explanation in question" constitutes a false dilemma.[68] My model provides a third option according to which moral facts do not cause moral beliefs, yet evolutionary explanations of our moral beliefs need not be viewed as distorting influences on those beliefs.

VIII. RICHARD JOYCE'S EVOLUTIONARY DEBUNKING OF MORALITY

In *The Evolution of Morality*, Richard Joyce offers an extended debunking of morality.[69] Like the other purported evolutionary debunkings I have considered so far, Joyce's challenge is grounded in Harman's 1977 challenge to morality. Indeed, Joyce explicitly acknowledges the connection between his argument and that of Harman.[70] I will focus my remarks here on the part of Joyce's debunking that is particularly relevant to the model of moral knowledge I sketched earlier—a section in which Joyce considers an attempt to "vindicate moral beliefs" based on reliabilism and argues that "the availability of an evolutionary moral genealogy does in fact pose a serious challenge for such [an endeavor]."[71]

In developing my reliabilist account of moral knowledge earlier, I drew on some of Peter Carruthers's ideas from his book *Human Knowledge and Human Nature*. Carruthers observes that holding false beliefs can sometimes be evolutionarily advantageous. He gives the example of a "belief in the magical properties of a particular plant, which in fact contains a powerful medicine" and notes that such a belief "might prove very useful to those who live in the region where that plant flourishes."[72] Joyce considers this very example and argues as follows:

> If there were an innate belief that certain plants have magical powers, and this belief were the product of a dedicated psychological mech-

68. Street, "Darwinian Dilemma," 155.
69. See Joyce, *Evolution of Morality*, 179–219.
70. Ibid., 184–90.
71. Ibid., 211.
72. Carruthers, *Human Knowledge*, 113.

> anism with a distinct evolutionary history . . . we would have an empirically confirmed hypothesis of how this belief-formation mechanism works which does not require that any of the beliefs be even approximately true, [so] we would have to conclude that any such innate beliefs are products of an *unreliable* process. . . . We can now apply this lesson to the case of innate moral beliefs. What seems clear is that in the crucial respect such beliefs are like the imaginary belief about the plant's magical properties. We have seen that nowhere does the evolutionary hypothesis outlined in earlier chapters assume that moral beliefs are or were true. . . . Thus I conclude that by the process reliabilist's own lights a certain plausible view of how innate moral beliefs may have evolved leads naturally to the conclusion that such beliefs are epistemically unjustified.[73]

In this passage Joyce appeals to the following claims:

> (5) If S's moral belief that P can be explained without appealing to the truth of P, then S's moral belief that P is a product of an unreliable process.
>
> (6) If S's moral belief that P is a product of an unreliable process, then S's moral belief that P lacks warrant.[74]

These claims together entail

> (P2) If S's moral belief that P can be explained without appealing to the truth of P, then S's moral belief that P lacks warrant.

The crucial claim here is (5). It is not clear to me whether Joyce intends to provide an argument for (5) or simply takes it as obviously true. If there is an argument for (5) here, it is an argument from analogy. The argument runs as follows. The belief that a certain plant is magical was formed by an unreliable process and human moral beliefs are relevantly like this plant-belief, so human moral beliefs, too, were formed by unreliable processes. The relevant respect in which human moral beliefs are like the plant-belief is that they can be explained without appealing to their truth.

The weakness of this reasoning lies in the fact that the plant-belief has a relevant feature that we cannot simply assume human moral beliefs possess, namely, falsehood. Moreover, it is this feature of the plant-belief that makes it clear that if there is a psychological mechanism devoted to producing these sorts of beliefs, that mechanism is unreliable. That the mechanism routinely produces false beliefs is what makes it clear that the mechanism in question is unreliable, not the fact that the beliefs it produces can be explained without appealing to their truth. So the

73. Joyce, *Evolution of Morality*, 215.

74. It is not clear whether Joyce himself accepts (6); his point is that reliabilists accept it and hence they should conclude that human moral beliefs are unjustified.

analogy between the plant-belief and human moral beliefs is not particularly strong—unless we know that human moral beliefs are false. Of course, if we are in a position to know that, there is little reason to bother with evolutionary debunkings of morality at all. Thus, the analogical argument for (5) is unconvincing.

Perhaps Joyce thinks (5) is obviously true. I think my model shows that (5) is false. To see this, consider a Westerner's belief that she possesses certain moral rights. The evolutionary explanation for this belief I sketched above does not appeal to the truth of this belief. Yet the model says that moral rights supervene on the cognitive faculties that generate this belief, and this thus implies that, even though moral rights in no way help to explain the corresponding belief about moral rights, the belief in question is nevertheless true. This third-factor explanation is a possibility that Joyce does not consider. In this way, his argument that reliabilists should conclude that human moral beliefs are produced by unreliable processes is too fast. My model is a counterexample to (5).[75]

IX. CONCLUSION

Having argued that three prominent epistemological evolutionary debunkings of morality are unsuccessful, I would like to conclude by saying a bit about why some might be seduced by such debunkings as well as comment on a project that is somewhat similar to such debunkings but which is, by my lights, much more reasonable. I suspect that some evolutionary debunkings gain an illegitimate air of plausibility by exploiting many people's moral skepticism (or at least skepticism about certain moral claims). This is perhaps clearest in the case of Joyce's claim of a parallel between moral beliefs and the belief that a certain plant is magical. Those already inclined to view morality as being, like magic, a kind of fantasy or illusion, may be less disposed to notice the question-begging nature of Joyce's comparison. Or, consider the first few sample human evaluative judgments provided by Sharon Street in her paper:

(1) The fact that something would promote one's survival is a reason in favor of it.
(2) The fact that something would promote the interests of a family member is a reason to do it.
(3) We have greater obligations to help our own children than we do to help complete strangers.[76]

75. For another critical discussion of Joyce's debunking, see Peter Carruthers and Scott M. James, "Evolution and the Possibility of Moral Realism," *Philosophy and Phenomenological Research* 77 (2008): 240–44.

76. Street, "Darwinian Dilemma," 115.

These moral claims coincide neatly with how we might want morality to work; indeed, don't they coincide suspiciously well with such desires? Any skepticism about such claims engendered in the mind of a reader primes her for a debunking. But if my criticism of Street's debunking above is correct, any plausibility her debunking might seem to have is merely an illusion.

For all I have argued in this article, there might well be good reason to be skeptical of certain moral claims, or even of morality altogether. If there is good reason to think that morality, in part or as a whole, is an illusion, then perhaps evolutionary theory can help us understand why human beings are susceptible to this sort of illusion.[77] This sort of project is perfectly sensible. Kant remarks that we possess a "propensity to quibble with these strict laws of duty, to cast doubt upon their validity . . . and to make them, where possible, more compatible with our wishes and inclinations."[78] If Kant is right about this, then perhaps evolutionary theory can shed some light on the presence of such a propensity in human nature. But if human moral knowledge is to be rejected as mere fantasy, this rejection should not be based on any of the evolutionary debunking arguments considered here.

77. See Hallvard Lillehammer, "Debunking Morality: Evolutionary Naturalism and Moral Error Theory," *Biology and Philosophy* 18 (2003): 567–81.

78. Immanuel Kant, *Grounding for the Metaphysics of Morals*, 3rd ed., trans. James W. Ellington (Indianapolis: Hackett, 1993), 17.