A PEIRCEAN RESPONSE TO THE EVOLUTIONARY DEBUNKING OF MORAL KNOWLEDGE

by Gary Slater

Abstract. The evolutionary debunking argument advanced by Sharon Street, Michael Ruse, and Richard Joyce employs the logic of Paul Griffiths and John Wilkins to contend that humans cannot have knowledge of moral truths, since the evolutionary process that has produced our basic moral intuitions lacks causal connections to those (putative) truths. Yet this argument is self-defeating, because its aim is the categorical, normative claim that we *should* suspend our moral beliefs in light of the discoveries about their non-truth-tracking origins, when it is precisely this claim that relies upon the normativity under attack. This article cites Charles S. Peirce (1839–1914) to argue that such self-defeat can be avoided by expanding upon the basic structure of the argument put forth by Griffiths and Wilkins, provided that one embraces a version of realism that corresponds with Peirce's doctrine of final causation. So construed, final causation reconciles real generals (including real moral values) with natural selection and undergirds further speculation of moral facts within values *per se*.

Keywords: Darwinism; epistemology; ethics; final causation; Charles S. Peirce; realism; semiotics; values

---

In his popular work *Darwin's Dangerous Idea*, Daniel Dennett likened Darwin's theory of evolution to a "universal acid," so potent that it "eats through just about every traditional concept, and leaves in its wake a revolutionized worldview, with most of the old landmarks still recognizable, but transformed in fundamental ways" (Dennett 1996, 63). A number of contemporary writers, including Michael Ruse, Sharon Street, and Richard Joyce, have taken this skeptical challenge very seriously. They all conclude that human beings can have *some* knowledge about the world, but only about those aspects of it that bear a certain sort of connection with the evolutionary success of our ancestors. These authors suggest that the only beliefs that can constitute knowledge are beliefs about the sorts of facts

Gary Slater is Doctoral Candidate in Theology and Religion at Oxford University, currently researching at at St. Edwards University, Austin, Texas, He may be contacted at 1510B Newning Avenue, Austin, TX 78704, USA; e-mail: garyslater@gmail.com.

 ISSN 0591-2385

for which natural selection is likely to have equipped us with an accurate awareness. This thoroughgoing evolutionary epistemology threatens to leave significant casualties in its wake. Most notably, its advocates—Street, Joyce, and Ruse—claim that it rules out knowledge of objective facts about morality, which (if there are such facts) would seem to lack the aforementioned connection with reproductive fitness. If one takes it that, despite the contrary proclamations of some philosophers, the great majority of human beings do (intuitively, at least) believe that morality is a matter of objectively binding obligations, it is consequently a matter of great concern if our apparent moral knowledge is debunked.

The aim of this article is to respond to the evolutionary debunking challenge, an effort that comprises three sections. In the first section, I examine precisely how the debunking argument is supposed to work. I suggest that the epistemological principle that guides evolutionary debunking is logically sound, though it is too narrow in its understanding of causality. In the service of antirealism regarding value—and, more specifically, regarding morality—it defeats itself, however. This is because the aim of evolutionary debunking is the categorical, normative claim that we *should* suspend our moral beliefs in light of the discoveries about their allegedly non-truth-tracking origins, when it is precisely this claim that relies upon the normativity under attack. In the second section, I argue that this self-defeat can be avoided without abandoning the logic of evolutionary debunking, on two conditions: that one adopts realism regarding value, and that one expands upon the premises of the evolutionary debunking argument to accommodate real final causation and an anisotropic sense of evolution. The pragmatist logician and philosopher Charles S. Peirce (1839–1914) offers precisely the resources for fulfilling these conditions. A careful reading of Peirce's 1902 paper "Minute Logic" illustrates how causal processes can accord with both real generals and knowledge of normative claims, the effect on this debate being that one can redirect the logic of evolutionary debunking toward purposes antithetical to its initial aim. This is to say that, rather than *debunking* moral knowledge, Peirce provides resources for *reinforcing* realism about morality without severing moral values from causal processes, including (but not limited to) biological processes of natural selection. In the third section, I develop this effort by arguing that objective values are roughly synonymous with final causes, though I speculate that there is a dimension of value that defies human knowledge. I also seek to delimit the place of morality within value *per se*.

Peirce, who was born in Cambridge, Massachusetts, on September 10, 1839 and died in Milford, Pennsylvania, on April 19, 1914, has been recognized as a foundational figure in the pragmatic philosophical tradition and in sign theory, yet his name nonetheless escapes the list of those whom the general public would know as history's most renowned philosophers. Such relative obscurity is due in part to the unfortunate circumstances of

Peirce's life, which were exacerbated by a notoriously difficult personality characterized by substance abuse, marital infidelity, and, very likely, physical violence. Peirce also had a tendency to revise the technical vocabulary in his writings without warning, inventing new terms or changing their definitions abruptly in such a way as to be off-putting to readers unfamiliar with the details of his thought. Yet in terms of the present project, no thinker could be more congenial, as his work combines subtle insights on final causation and natural selection with a staunch commitment to realism regarding value. Although philosophical arguments have been employed to defend commonsense beliefs against modern skepticism since at least the eighteenth century, when Thomas Reid (1710–1796) articulated his commonsense philosophy in response to David Hume (1711–1776), Peirce provides resources whose relevance has not been fully explored regarding the evolutionary debunking of moral knowledge. Some excellent groundwork has been accomplished, however, in the recent secondary literature on Peirce. For example, in his book, *Peirce's Theory of Signs* (2007), Thomas L. Short has shown convincingly that Peirce offers a version of final causation that is also compatible with evolution. Short has also linked these issues to Peirce's mature semeiotic (Peirce's preferred term for sign theory). Another helpful resource is Andrew Robinson, whose work engages biology, hermeneutics, and theologically relevant aspects of Peirce's writings with specific reference to natural selection (Robinson 2010; Southgate and Robinson 2010). Menno Hulswit's book, *From Cause to Causation: A Peircean Perspective* (2002), is also germane in articulating the place of real generals in Peirce's doctrine of final causation.

The Logic of Evolutionary Debunking

In their paper, "When Do Evolutionary Descriptions of Belief Debunk Belief?" (in press), Paul Griffiths and John Wilkins have sought an explicit formulation of an epistemology designed to withstand skepticism based on natural selection. Griffiths and Wilkins argue that skepticism can be resisted for a given type of beliefs *provided* there is, for the type of beliefs in question, what they call a "Milvian Bridge" connecting evolutionary success to the truth of the beliefs in question. At the Battle of the Milvian Bridge in 312 CE, Rome's first officially Christian emperor, Constantine, fought and defeated Maxentius under the banner of Christ and went on to found the Byzantine Roman empire. Constantine claimed that the victory was due to the *truth* of the Christian faith. Utilitarian success—which Griffiths and Wilkins refer to in a nontechnical sense as "pragmatic" success—was thus (allegedly) the result of true belief. Analogously, Griffiths and Wilkins argue, in order for a given type of beliefs (e.g., visual beliefs, logico-mathematical beliefs, moral beliefs) to be capable of constituting knowledge, it must be the case that the type of

beliefs in question contributed to reproductive success in virtue of being *true*:

> We call an argument which links true belief to pragmatic success a "Milvian Bridge" argument. The specific kind of pragmatic success with which we will be concerned is evolutionary success. To defeat evolutionary skepticism, true belief must be linked to evolutionary success in such a way that evolution can be expected to produce organisms which have true beliefs. (Griffiths and Wilkins, in press)

Griffiths and Wilkins thus state their Milvian Bridge criterion for knowledge as follows: "The *X* facts [must be] related to the evolutionary success of *X* beliefs in such a way that it is reasonable to accept and act on *X* beliefs produced by our evolved cognitive faculties" (Griffiths and Wilkins, in press).

It will help to illustrate in a more concrete manner what exactly is the thought motivating this evolutionary criterion for knowledge. The following pair of examples will make more lucid the central principle that drives it. First, consider a fictional pair of early humans, Fred and George. Fred, let us suppose, has perceptual faculties which frequently result in false beliefs about medium-sized physical objects in his vicinity. Fred often mistakes fire for flowers, cliff-edges for gentle streams, and dangerous predators for cuddly companions. George, on the other hand, has perceptual faculties that deliver him mostly true beliefs about his physical surroundings. All else being equal, then, George will likely live significantly longer than Fred, and will enjoy significantly more opportunities to reproduce than Fred will. George's offspring will likely outnumber Fred's. And provided the traits which determine the accuracy of their respective perceptual abilities are to some extent heritable by their offspring, George's offspring will tend to have more accurate perceptual faculties than Fred's offspring. After hundreds of generations of this evolutionary winnowing, the population of creatures of which Fred and George were once members can reasonably be expected to consist overwhelmingly of creatures whose perceptual beliefs are, on the whole, true. This, at any rate, is the picture that Griffiths and Wilkins seem to have in mind as they venture the following:

> It is overwhelmingly likely that commonsense beliefs are produced by cognitive adaptations that track truth . . . At the heart of that explanation will be the fact that animals can increase their fitness by detecting states of affairs in the world and matching their actions to those states of affairs. (Griffiths and Wilkins, in press)

Consider a second example. Suppose now that Fred and George are equal with respect to the accuracy of their perceptual beliefs, so that natural selection can't sift them on that basis. They differ, however, with respect to the accuracy of their beliefs about the invisible and intangible objects that—suppose for the sake of the example—litter their environment. Fred

believes that the invisible-intangible objects in the nearby area are large and spherical, and about this he is correct, let us suppose. George, on the other hand, mistakenly believes that the invisible-intangible objects in the vicinity are small and cubic. In this case, however, it seems that the difference in accuracy between Fred and George's beliefs *won't* give rise to a corresponding difference in reproductive success. Thus, all else being equal, George and his offspring will get along in the evolutionary game just as well as Fred and his offspring, *despite* the former's radically false invisible-intangible object beliefs. And so, the thought goes, we have *no* reason to suppose that after hundreds of years of natural selection, the population of creatures of which Fred and George were once members will consist mostly of creatures whose invisible-intangible-object beliefs are, on the whole, true.

To adjust this example slightly, suppose that having the belief that invisible-intangible objects are large and round (as is in fact the case in the fictional world under consideration) tends to make a creature more socially adept than does the belief that invisible-intangible objects are small and cubic. This adjustment should make invisible-intangible object beliefs relevantly analogous to moral beliefs. Now, it looks like Fred's and George's differing invisible-intangible object beliefs *will*, after all, make a difference to their reproductive prospects. All else being equal, Fred and his kin can be expected to fare somewhat better than George and his kin, and so on for the subsequent generations. Eventually the population will consist mostly of creatures that believe, correctly, that invisible-intangible objects are large and spherical. But unlike with the perceptual beliefs in the first example, it isn't due to the *truth* of these invisible-intangible-object beliefs that the tendency to have these beliefs was evolutionarily favored, but rather, due to the benefits conferred by a trait to which those beliefs gave rise—benefits that would have been conferred even if those beliefs had been wholly false. In considering whether the evolutionary outcome would have been otherwise had the truths about invisible-intangible-object beliefs been *different* than they actually are, it seems that, despite such a difference in the invisible-intangible-object facts, the evolutionary outcome would have been *no* different.

And this is what Griffiths and Wilkins think is the crucial difference between mundane perceptual or "commonsense" beliefs, on the one hand, and moral beliefs, on the other, when they claim that "contemporary evolutionary explanations of morality, just like Darwin's own explanation, do not involve any adaptive advantages produced by detecting and acting in accordance with objective moral facts" (Griffiths and Wilkins, in press). In short, the debunkers argue that natural selection is sensitive to facts about visible and tangible physical objects, but *in*sensitive to facts about objective morality (which relevantly resemble invisible-intangible object beliefs). As Michael Ruse puts it: "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" (Ruse 2006, 24).

It is with something very much like this picture in mind that Ruse, Street, and Joyce have formulated versions of what has become known in the literature as an "evolutionary debunking argument" against moral knowledge. This argument is aimed at undermining the idea that we have knowledge of objective truths of morality, where "objective truths of morality" are to be understood as truths about the morally relevant properties instantiated in the world that hold true independently of anyone's point of view. As Joyce and Street have relied most directly on something like the Milvian Bridge Principle, it is from these two authors that citations will primarily be drawn. Joyce and Street claim that such objective moral truths are incapable of *causing* anything in the world in virtue of their intangible nature, and hence they form no part of the scientific explanation of our evolved moral judgments. As Street puts it: "A creature obviously can't run into such truths or fall over them or be eaten by them. In what way would it have promoted the reproductive success of ancestors to grasp them?" (Street 2006, 130–31). Joyce expresses a closely related thought:

> We have an empirically confirmed theory about where our moral judgments come from (we are supposing). This theory doesn't state or imply that they are true; it doesn't have as a background assumption that they are true, and, importantly, their truth is not surreptitiously buried in the theory by virtue of any form of moral naturalism. This amounts to the discovery that our moral beliefs are products of a process that is entirely independent of their truth, which forces the recognition that we have no grounds one way or the other for maintaining these beliefs (Joyce 2006, 211).

It may be suggested, then, that the debunking argument against moral knowledge can be stated most simply as follows:

> **Milvian Bridge premise:** If humans have knowledge of *X* facts, then *X* facts played a causal role in the evolution of our ancestors in such a way that, had the *X* facts been different, natural selection would have favored correspondingly different *X* beliefs.
>
> **Evolutionary premise:** Beliefs about objective morality did *not* play a causal role in the evolutionary success of our ancestors in such a way that, had the objective moral facts been different, natural selection would have favored correspondingly different beliefs about objective morality.
>
> **Conclusion:** Therefore, humans do not have knowledge of objective moral facts.

In its current form, this argument is self-defeating.

In reviewing the argument, a question presents itself: on what basis are the premises of evolutionary debunking selected? With regard to Joyce and Street's arguments, the selection of premises for the

evolutionary debunking argument is carried out in the service of values that are explicitly named. For example, in addressing the question of why natural selection ought to be seen as the overriding determinant of knowledge, Street's answer is that in epistemology an evolutionary account is "clearer," "more parsimonious," and "sheds more light on the explanandum in question" (Street 2006, 129). Yet the values of clarity, parsimony, and the shedding of "light" make in themselves no reference to an organism's reproductive success. One finds a similar tendency in Joyce's exhortation that "we should, initially, cultivate an open mind in order to go and find some other, more reliable grounds for either believing or disbelieving moral propositions" (Joyce 2006, 211). But by what rule should one follow Joyce's advice? The answer is through extolling open-mindedness as a value. In itself, however, open-mindedness confers no particular adaptive benefit. Since such values are offered categorically as well as through utility, and considering the aim of evolutionary debunking to deny realist theories of value, this is a contradiction that works against the debunking aim. Ironically, Street notes something similar in her challenge to realism:

> Exactly why would it promote an organism's reproductive success to grasp the independent evaluative truths posited by the realist? The realist owes us an answer here. It is not enough to say, ''because they are true.'' We need to know more about why it is advantageous to apprehend such truths before we have been given an adequate explanation. (Street 2006, 130)

According to logic of the Milvian Bridge premise, the values that are extolled in evolutionary debunking are themselves debunked. There is nothing necessarily wrong with an argument stemming from the claim that something is categorically good. Yet the irony of Joyce and Street venturing normative claims in order to debunk epistemological access for whole categories of normative claims is compounded by their obfuscation regarding the values that inform their own arguments.

A truly constructive response to evolutionary debunking is in fact *reconstructive*, as it entails retaining such values as clarity and open-mindedness while reexamining the logical rules whose effect is to ghettoize moral knowledge. Using resources from Peircean logic and philosophy, as well as from such key commentators on Peirce as Short, Robinson, and Hulswit, this article adopts a strategy of arguing that self-defeat can be avoided if one substitutes realism for antirealism regarding value. Such a shift requires revisions of both the Milvian Bridge and evolutionary premises, since in their current forms these premises do in fact foreclose the possibility of realism regarding value. In the case of the Milvian Bridge premise, revisions include inserting the term "values" alongside "facts," expanding the notion of "causal" to include final as well as efficient causation, and complementing the emphasis on one's biological ancestors with objectively general possibilities whose concrete existence has yet to come about, yet whose bearing on

natural selection is real. In the case of the evolutionary premise, revisions represent what amounts to a wholesale reversal, expressible as follows: "It is likely (though not necessary) that beliefs about objective morality *did* play a causal role in the evolutionary success of our ancestors in such a way that, had the objective moral facts/values been different, natural selection would have favored correspondingly different beliefs about objective morality, so long as 'causal' is understood to include final causation and evolution is understood anisotropically." The ensuing conclusion would then read: "Therefore humans can have knowledge of objective moral facts." I concede that such revisions render the debunking premises virtually unrecognizable from their previous iterations. Yet certain key points of the Milvian Bridge premise do remain, the most important of which are that knowledge is indispensably causal and that it is bound up with natural selection (though not reducible to it). The result is a continuous line of reasoning that extends from causal processes through nonmechanistic natural selection to values that are independently real. These elements are mutually reinforcing and often overlapping, but for present purposes what matters most is that the argument I seek to articulate does not defeat itself, and that it also points toward more precise descriptions of moral facts.

Revising the Debunking Premises

Before discussing the premises of evolutionary debunking, it helps to establish the appropriate context for engaging with Peirce's notion of final causation (the citations regarding which are all cited from *The Collected Papers of Charles Sanders Peirce*, henceforth CP). As Short has noted, Peirce's writings on final causation arose at the turn of the twentieth century as part of a broader effort to devise a philosophical architectonic (Short 2007, 64). The architectonic was an outline of the overarching structure of relations among disciplines. Although Peirce had been interested in such a project since grappling with Kant's *Critique of Pure Reason* in the 1860s, he fell away from the architectonic only to resurrect it in the 1890s, and in 1902 he undertook major revisions that bore fruit with his 1903 lectures on pragmatism. The architectonic encompassed everything from metaphysics and chemistry to mathematics and aesthetics, yet as it appeared circa 1902, the specific reasons for its existence lay in what that Peirce at that time called "phenomenology" (Short 2007, 60–61). Peirce's phenomenology (which he would later call "phaneroscopy") bears little relation to the better-known project of Edmund Husserl, though it does resemble Husserl's in that it purports to be a science of appearances. Short makes the case that Peirce wanted to expose "the experiential roots of the ideas of externality" without assuming realism in such a way as to beg the question it was meant to address, which led him to revise his architectonic as a means of navigating among its disciplines to find an appropriate descriptive vocabulary

(Short 2007, 64). The doctrine of final causation arose within this effort as a means to provide Peirce with an account of explanation in accordance with realism and amenable to the sort of system-building that the architectonic represented. In 1902, Peirce composed a paper, "Minute Logic," that contains some of his most compelling insights on final causation. It is this article that is the chief text in my effort to revise the premises of the evolutionary debunking argument.

Recall that, in its original form, the Milvian Bridge premise reads:

> **Milvian Bridge premise:** If humans have knowledge of *X* facts, then *X* facts played a causal role in the evolution of our ancestors in such a way that, had the *X* facts been different, natural selection would have favored correspondingly different *X* beliefs.

In light of Peirce's final causation, this premise can be revised to look like this:

> **Revised Milvian Bridge premise:** If humans have knowledge of *X* facts/values, then *X* facts/values *perhaps* played a causal role in the evolution of our ancestors in such a way that, had the *X* facts been different, natural selection would have favored correspondingly different *X* belief, and *certainly* played a causal role in representing the objectively general possibilities for which selection is made and by which it is explained.

In the revised premise, the inclusion of values alongside facts corresponds with the recognition of final causation alongside efficient causation. In "Minute Logic," Peirce defined final causation as follows:

> We must understand by final causation that mode of bringing facts about according to which a general description of result is made to come about, quite irrespective of any compulsion for it to come about in this or that particular way. (CP 1.211)

Peirce's definition bears expounding upon for a moment, as it reflects various features that distinguish it from the versions of final causation typically associated with classical or scholastic philosophy. First, this definition does not entail that some concrete future event magically reaches backwards through time and interferes with present events. Rather, a final cause, though real, is general and unrealized. Although not all final causation entails purpose, purpose *does* entail a final cause, and hence is general:

> A purpose is an operative desire. Now a desire is always general; that is, it is always some *kind* of thing or event which is desired; at least, until the element of will, which is always exercised upon an individual object upon an individual occasion, becomes so predominant as to overrule the generalizing character of desire. (CP 1.205)

In his book, *From Cause to Causation: A Peircean Perspective*, Menno Hulswit provides an incisive account of how Peirce allows one to view the interaction between such "operative desires" and concrete present events. As Hulswit puts it:

> Whenever someone wants to realize an idea, this idea functions as a principle of selection in the choice of the appropriate means whereby that idea is to be realized. The idea is general. In terms of the relationship between final and efficient causation, efficient causation, considered apart from its final causational component, is a dyadic relation between two concrete individual events or facts, and final causation is a triadic relation between the general final cause, the concrete efficient cause, and its concrete effect. (Hulswit 2002, 80)

As they appear in the texts of Street and Joyce, the values that guide evolutionary debunking function much like final causes as Hulswit identifies them, which is to say that they do not *follow from* the Milvian Bridge so much as are put forth as *intrinsically worthy* reasons for accepting the principle. Yet Peircean final causation goes beyond what is expressed by Hulswit. Just as it does not require an identifiable purpose, neither does final causation require an identifiable source of causal agency. As Peirce put it:

> A final cause may be conceived to operate without having been the purpose of any mind: that supposed phenomenon goes by the name of *fate*. The doctrine of evolution refrains from pronouncing whether forms are simply fated or whether they are providential; but that definite ends are worked out none of us today any longer deny. (CP 1.204)

The ability to articulate real final causes without presupposing a specific source of agency reinforces a naturalistic account of value. It is important, however, to note that Peirce distinguished between the terms "reality" and "existence" in a way that many contemporary thinkers do not. As he put it in a separate text (but in a manner that bears on these issues), reality is broader than existence:

> I myself always use *exist* in its strict philosophical sense of "react with the other like things in the environment." . . . The word "reality," on the contrary, is used in ordinary parlance in its correct philosophical sense . . . I define the *real* as that which holds its characters on such a tenure that it makes not the slightest difference what any man or men may have *thought* them to be, or ever will have *thought* them to be. (CP 6.495)

Final causes are real without being existent, and, as I will argue, so are moral values.

How does Peirce's doctrine of final causation guide the revisions of the Milvian Bridge premise? One important upshot is that in spite of its requiring neither agency nor purpose, Peircean final causation *does* entail

realism regarding value. This is what warrants the inclusion of the term "values" alongside "facts" in the revised premise. For on the understanding of a final cause as a metaphysically general possibility whose reality is enacted in processes of selection that point toward it, it is reasonable to interpret value in such a way as to include values as final causes. Such, at any rate, is the sense in which values are manifested in the texts of Street and Joyce, serving to justify as much as explain the selection of their arguments, and whose presentation in categorical terms only makes sense if one understands them as general. There is also the second revision of the Milvian Bridge premise, which reads:

> *X* facts/value *perhaps* played a causal role in the evolution of our ancestors in such a way that, had the *X* facts been different, natural selection would have favored correspondingly different *X* belief, and *certainly* played a causal role in representing the objectively general possibilities for which selection is made and by which it is explained.

This revision can be explained in light of Peirce's perspective on the relationship between final and efficient causation. Peirce saw the two types of causation as real and as complementary, yet not at all equal in terms of epistemology. As Peirce put it, "Final causation without efficient causation is helpless. . . . Efficient causation without final causation, however, is worse than helpless, by far; it is mere chaos" (CP 1.220). When the original Milvian Bridge premise refers to a "causal role" in its claim that "had the *X* facts been different, natural selection would have favored correspondingly different *X* belief," this role is to be understood with respect to efficient causation, which is to say with reference to a concatenation of concrete events that has led directly from past to present and supplied the present with its brute actuality. Conversely, when the revised Milvian Bridge premise refers to a "causal role" in the sense of "representing the objectively general possibilities for which selection is made and by which it is explained," this role is to be understood with respect to final causation. To the extent that knowledge of some fact or value entails an explanation for that fact or value, final causation is instrumental in the formation of knowledge. This is the reason for including the word "certainly" into the revised Milvian Bridge premise. As for the inclusion of the word "perhaps," this simply reflects the plausible claim that although the relative adaptive success of our biological ancestors has very likely determined some or even most of what humans take to be commonsense belief, not all knowledge is reducible to biology.

As for the second premise in the evolutionary debunking argument, the evolutionary premise, its initial form was given as follows:

> **Evolutionary premise:** Beliefs about objective morality did *not* play a causal role in the evolutionary success of our ancestors in such a way that, had the

> objective moral facts been different, natural selection would have favored correspondingly different beliefs about objective morality.

And here is the revised version:

> **Revised evolutionary premise:** It is likely (though not necessary) that beliefs about objective morality *did* play a causal role in the evolutionary success of our ancestors in such a way that, had the objective moral facts/values been different, natural selection would have favored correspondingly different beliefs about objective morality, so long as 'causal' is understood to include final causation and evolution is understood anisotropically.

In what virtually amounts to a reversal of the evolutionary premise (and insofar as the phrase "though not necessary" obtains, even an abandonment of it), there are several changes in language between the evolutionary premise and its revised version. To make sense of these changes, three revisions in particular bear greater discussion: the reference to anisotropic explanation alongside final causation, the inclusion of the word "likely," and the inclusion of the parenthetical phrase "though not necessary."

Among the revisions to the evolutionary premise, the reference to evolution understood anisotropically—anisotropy being the property of directional dependency—extends most directly out of the distinction between final and efficient causation already discussed. In "Minute Logic," Peirce wrote:

> An efficient cause, detached from a final cause in the form of a law, would not even possess efficiency: it might exert itself, and something might follow *post hoc,* but not *propter hoc;* for *propter* implies potential regularity. Now without law there is no regularity; and without the influence of ideas there is no potentiality. (CP 1.213)

Peirce's comments provide insight on the entailments of evolution understood anisotropically or directionally. Were evolution to reduce to efficient causation, this would be a purely mechanistic process that has "no influence of ideas" and no variation, creativity, or growth. On the other hand, evolution as informed by final causation is characterized by "potential regularity," includes an identifiable *propter* along with the *post*, and is understood anisotropically. Citing from Elliott Sober's book, *The Nature of Selection* (1984), Short notes that Sober's distinction between "selection of" and "selection for" is useful for explaining how Peircean final causation bears on evolution:

> Eliminating the unfavorable is selecting the favorable. Implicit in the theory—indeed, in the idea of selection *simpliciter*—is a distinction between 'selection of' and 'selection for'. . . . If we are to defend what Peirce did with Darwin, not merely as reasonable in his day but with continuing pertinence in our own, then we need first to restate the Darwinian idea, not

> in Darwin's terms but in contemporary, neo-Darwinian terms, with some care, relying on the best treatments of it. (Short 2007, 128)

The goal to restate the neo-Darwinian idea with care is partly responsible for the analysis of the logic of evolutionary debunking that comprised the first section of this article. For Short, such an effort involves construing Sober's selection for/selection of distinction respectively in terms of abstractness and concreteness:

> Being concrete, what is selected has many features and it has each of those features in a quite specific way. But what is selected is selected *for* just one of its features—more precisely, for a *type* of feature that it exemplifies or results in exemplifications of. (Short 2007, 129)

Evolution that is understood with respect to general types for which selection is made cannot be understood mechanistically. This point tends to get overlooked by contemporary proponents of evolutionary debunking. Note the following from Street:

> Now of course there are radical differences between the mechanism of a reflex response and the "mechanism" of an evaluative judgment. The former is a brute, hard-wired physical mechanism, while the latter is a conscious mental state subject to reflection and possible revision in light of that reflection. But this does not change the fact that there is a deep analogy between their functional roles. From an evolutionary point of view, each may be seen as having the same practical point: to get the organism to respond to its circumstances in a way that is adaptive. Something like a reflex mechanism does this through a particular hard-wiring of the nervous system, while an evaluative judgment—or a more primitive evaluative experience such as some other animals are likely to have—does this by having the organism experience a particular response as called for, or as demanded by, the circumstance in question. (Street 2006, 128)

By highlighting only a *particular* response, Street misses out on the sense in which a final cause as a *type* of response plays a role in explaining the circumstance in question. This is a distinction that Peirce in his day criticized Darwin's proponents for overlooking. As he put it in "Evolutionary Love" (1893):

> Natural selection, as conceived by Darwin, is a mode of evolution in which the only positive agent of change in the whole passage from moner to man is fortuitous variation. To secure advance in a definite direction chance has to be seconded by some action that shall hinder the propagation of some varieties or stimulate that of others. (CP 6.296)

Although there may appear to be little difference between Peirce's take on natural selection here and that of Street, note that Peirce acknowledges that a winnowing principle is necessary to understand evolution according to a "definite direction," which is to say anisotropically. Peirce in 1890

had also recognized "the principle of hereditary transmission" (CP 1.398), which rounded out what he saw as the principles required to grasp Darwin's central idea, and which requires an element of anisotropic explanation by which hereditary continuity is discerned across multiple generations.

As for the inclusions of the word "likely" and phrase "though not necessary" in the revisions to this premise, these changes reflect recent research suggesting that interpretation that makes reference to a final cause *can* be empirically shown to affect evolutionary outcomes, and yet also that such outcomes need not exhaust the possibilities for what counts as knowledge. In other words, it is *possible* for adaptive behavior undertaken in relation to some value to affect the outcome of evolution understood mechanistically, though at the same time it is also possible for knowledge of evaluative facts to make no reference to adaptive success whatsoever. This defies the central claim of the evolutionary premise, which is that beliefs about objective morality did *not* play a causal role in the evolutionary success of our ancestors in such a way that, had the objective moral facts been different, natural selection would have favored correspondingly different beliefs about objective morality. It is helpful at this point to call attention to recent biosemiotic research that has drawn from Peirce's work. Andrew Robinson, for example, has argued that Peircean semiotics "may contribute to understanding how humans evolved" (Robinson 2010, 9), and has sought to show empirically how even primitive organizations exhibit adaptive behavior in ways that are irreducible to purely mechanistic explanation. In a *Zygon* paper coauthored with Christopher Southgate, "Interpretation and the Origin of Life" (2010), Robinson argues that "Peirce's account of the triadic, teleological, and fallible nature of semiotic processes can be developed into a tightly formulated general definition of interpretation capable of covering all instances and levels of interpretation," one that "provides an additional diagnostic property of protobiotic entities" (Southgate and Robinson 2010, 358). Careful to seek a definition of interpretation that "does not presume the concepts it seeks to test" (Southgate and Robinson 2010, 346), the authors contend that protobiotic entities exhibit adaptive behaviors that do not reduce to mechanistic explanations, and that such behaviors can be defined as interpretations. Entailing both anisotropic explanation and final causation, the authors define interpretation with precision in the following terms (Southgate and Robinson 2010, 348):

A response, R, of an entity is an interpretant [significate effect] of some X as a sign of some object O if and only if:

(1) The entity has a property, Q, of undergoing change of state S in response to some X, where R is any actual instance of such a response;

(2) And

a. R tends to increase the probability of an effect of a certain general type, P;
b. This tendency of R depends on a relation between X and O, where the occurrence of X does not necessarily imply the occurrence of O;

(3) The property Q has been selected for the tendency of instances of R to actualize effects of general type P.

In spite of the technical nature of this definition, Robinson and Southgate's are comparatively straightforward in their bearing on revisions to the debunking premise. By extending their research to include even single-celled amoebae, the authors cast their explanatory net wide enough to extend final causation and anisotropic explanation (e.g., "R tends to increase the probability of effects of a general type P") into many areas previously thought only explainable with respect to mechanistic natural selection. To put it perhaps crudely, if an amoeba in some primordial past displayed behavior whose adaptive success is only explainable with reference to some final cause as a general type of outcome, then the totality of concrete outcomes that *did* come about between that time and the present owe something to that amoeba's change in behavior, that is, its having interpreted.

So it is that, with reference to the past, final causation can be shown to be like efficient causation, or at least effect a chain of changes of state explainable with respect to efficient causation. Robinson and Southgate write that "it is hard to imagine a living organism that did not possess properties Q of responding to signs X of features O of their environment by undergoing changes of state, each Q having been selected for its tendency to further outcomes of general type" (Southgate and Robinson 2010, 353), and so in matters human it is certainly possible to imagine changes of state brought about on the basis of belief in some value, including some moral value. This is the reason for my inclusion of the word "likely" in the revision of the evolutionary premise. At the same time, it is also possible for knowledge of value to make *no* reference to adaptive success whatsoever, which is the basis for the revised premise to include the parenthetical phrase "but not necessary." One of the most significant drawbacks of the debunking position is its implication that epistemology is chained to primordial struggles for survival, the parameters for human knowledge borne like boats ceaselessly back into the past. For example, Street claims that even "our present-day ability to do astrophysics is presumably a refined extension of more basic abilities to discover and model the physical features of the world around us" (Street 2006, 144). This entails that biology and its pressures for reproduction and survival continue to dominate human judgment above all or most other factors. Robinson and Southgate have

drawn from Peircean semiotics to counter this claim, though perhaps the most eloquent explanation comes from Short. In relation what he calls the "emancipation of purpose from biology" (Short 2007, 148), he puts the matter as follows:

> It is often supposed that pleasure and avoidance of pain are the ultimate purposes for which these other purposes [e.g. morality, science and art, etc.] are adopted as means. But that is to overlook a fact we have been at pains to establish: that a purpose formed is independent of the conditions that explain its formation . . . Besides, pleasure and pain have turned out to be highly malleable. We learn to take pleasure in things—caviar, alcohol, hard work—initially unpleasant. Art appreciation is taught. We have a moral duty to take pleasure in doing our duty and in exercising self-restraint. But most importantly, our capacity to diagram and symbolize means that we can formulate possible purposes independently of any motive to adopt them. Sometimes, we then adopt them, arbitrarily or for reasons not well considered. (Short 2007, 149)

In light of this point, and in light of the relation between final causation and natural selection as found in his Peirce's "Minute Logic" of 1902, it is possible to state the following revision of the debunking premises, whose changes entail a different conclusion:

> **Revised Milvian Bridge premise:** If humans have knowledge of *X* facts/values, then *X* facts/values *likely* played a causal role in the evolution of our ancestors in such a way that, had the *X* facts been different, natural selection would have favored correspondingly different *X* belief, and *certainly* played a causal role in representing the objectively general possibilities for which selection is made and by which it is explained.
>
> **Revised evolutionary premise:** It is likely (though not necessary) that beliefs about objective morality *did* play a causal role in the evolutionary success of our ancestors, so long as 'causal' is understood to include final causation and evolution is understood as anisotropic and nonmechanistic.
>
> **Revised conclusion:** Therefore, humans can have knowledge of objective moral facts.

Interrogating Value and Morality

The logic of the Milvian Bridge has now been augmented to accommodate final causation and anisotropic understandings of evolution, which warrants a realist epistemology regarding value that includes value as the object of moral knowledge. There are, however, outstanding questions, and the effort to raise these questions points to a much larger, more difficult project. One such question is the nature and location of real values, which would likely require a detailed account of the means by which values pervade cognition, perhaps even perception, and yet also possess an independent reality. Another question concerns the place of moral values within the

larger category of value as a whole, which would likely require not only clear definitions for what makes morality distinct from value *per se*, but also an account of how morality can be both informed by personal, social, historical, and perhaps even causal contexts and yet also remain irreducible to such contexts. Finally, underlying all of these questions is the further question of which resources in Peirce's texts can (or cannot) be of use to these projects, with the inevitable secondary issue of where these resources appear in Peirce's career and how they fit in more broadly with what Peirce tried to do. The present section represents only a sketch. It also hews less closely to the source material in Peirce than what has come so far. But the task is worth undertaking, even if only as speculation following on from the more localized concerns about evolutionary debunking preceding it.

When it comes to tracing out the nature, place, and epistemological dimensions of real values, Street has issued a relevant challenge. She writes: "The challenge for realist theories of value is to explain the relation between these evolutionary influences on our evaluative attitudes, on the one hand, and the independent evaluative truths that realism posits, on the other" (Street 2006, 109). My answer is that a value is synonymous with a final cause. As a final cause, a value can get itself thought or enacted in actuality, and it is in this way that a value enters the realm of the phenomenal and makes itself available for empirical study. As Peirce rather inspiringly put it, "Ideas are not all mere creations of this or that mind, but on the contrary have a power of finding or creating their vehicles, and having found them, of conferring upon them the ability to transform the face of the earth" (CP 1.217). At the same time, the reality of a value transcends the existent present, not only in terms of vague potentialities for future embodiment, but also through the explanatory capacity of an already actualized past. There is perhaps even a dimension of value that transcends what might conceivably enter the imagination, its relation to the imagination being analogous to the way that one's imagination is continuous with a world that is also beyond it. Although there appears to be little in Peirce's texts that bears on this sort of value, Short has suggested that Peirce in 1907 arrived at the insight that an ultimate interpretant is not part of the "hermetic circle of words interpreting words and thoughts interpreting thoughts" (Short 2007, 59), but rather is, in Peirce's works, "the living definition that grows up in the habit" (CP 5.491). I interpret Short in that such a dimension of value is real precisely because it is not *merely* an object of the imagination, but is more often the very thing that, in guiding its purposes, in being out of the "hermetic circle," makes imagination possible. Value so construed escapes even the revised Milvian Bridge premise, and is almost certainly not synonymous with morality.

Morality, although certainly laden with values, is distinct from this more astral reading of value just described. Joyce holds that a moral judgment entails a mix of cognitivism, which is the "view that moral judgments do

not express beliefs but rather perform some other kind of speech act" (Joyce 2006, 53) and noncognitivism, which is the view that "to say that someone is evil is simply to ascribe to him a property without thereby announcing how one feels about the person" (Joyce 2006, 56). He further clarifies his definition of moral judgments as "ways of expressing conative attitudes, such as approval" that "also express beliefs, i.e. they are assertions," which "purport to be deliberative considerations irrespective of the interests/ends of those to whom they are directed," be "inescapable," "transcend human conventions," "centrally govern interpersonal relations," "imply notions of desert and justice," and involve "the emotion of guilt is an important mechanism for regulating one's moral conduct" (Joyce 2006, 70). These descriptions are insightful. Peirce, for his part, distinguished aesthetics from morality by arguing that "esthetic goodness . . . may be possessed, and in some degree must be possessed, by any kind of representamen—rhema, proposition, or argument" (CP 5.140), whereas "moral goodness, or veracity, may be possessed by a proposition or an argument, but cannot be possessed by a rhema" (CP 5.141). Esoteric terminology aside (in Peirce's vocabulary, a "representamen" is simply a sign that stands for something in some respect, and a "rhema" is a simple representation—e.g., goodness, whiteness—without separate parts), the key to this distinction is that knowledge of moral goodness, while owing its reality some final cause or value, must also reside in some cognitively reflective form available form such as an argument or proposition. Without disagreeing with either Joyce's or Peirce's descriptions, I would suggest a refinement of Peirce's claim by asserting a sequential process by which a value is felt, cognized, and acted upon.

Conclusion

Peirce was a sprawling and diffuse thinker, and in terms of the issues that have motivated the evolutionary debunking debate as well as those issues that are implicated by it, there are several further resources within his thought that could be marshaled in support of worthwhile insights. Candidates within Peirce's texts include his logic of abduction, the logic of relations, connection between purpose and thought, and considerations on process, each of which is in some way consistent with a realist position regarding value, and as such could be called upon to defend the possibility of inquiry into morally relevant propositions that the debunking advocates attack. A particularly compelling option, one that has been initiated by Short and Robinson and developed in relation to such longstanding career projects within the field of Peircean religious studies such as Robert C. Neville's axiology of thinking or Michael Raposa's theosemiotic, is to navigate the taxonomies of Peirce's semeiotic to articulate a metaphysics of religious experience within which real moral values are included, and

which would bear on science and religion questions beyond even the implications of evolutionary debunking. These questions, however, extend beyond the purview of the current article, which has shown the evolutionary debunking argument to be self-defeating and responded to this self-defeat by drawing from Peirce's doctrine of final causation to expand and repair the debunking premises.

References

Dennett, Daniel. 1996. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. Harmondsworth, UK: Penguin.

Griffiths, Paul E., and John S. Wilkins. In press. "When Do Evolutionary Explanations of Belief Debunk Belief?" In *Darwin in the 21st Century: Nature, Humanity, and God*, ed. P. R. Sloan. Notre Dame, IN: Notre Dame University Press.

Hulswit, Menno. 2002. *From Cause to Causation: A Peircean Perspective*. Dordrecht, The Netherlands: Kluwer Academic.

Joyce, Richard. 2006. *The Evolution of Morality*. Cambridge, MA: MIT Press.

Peirce, Charles S. *Collected Papers of Charles Sanders Peirce*, vols. 1–8, 1931–53, vols. 1–6 edited by Charles Hartshorne and Paul Weiss; vols. 7–8 edited by Arthur W. Burks. Cambridge, MA: Harvard University Press.

Robinson, Andrew. 2010. *God and the World of Signs: Trinity, Evolution, and the Metaphysical Semiotics of C. S. Peirce*. Leiden, The Netherlands: Brill.

Ruse, Michael. 2006. "Is Darwinian Metaethics Possible? (And If It Is, Is It Well Taken?)." In *Evolutionary Ethics and Contemporary Biology*, ed. Giovanni Boniolo and Gabriele De Anna, 13–26. Cambridge: Cambridge University Press.

Short, Thomas L. 2007. *Peirce's Theory of Signs*. Cambridge: Cambridge University Press.

Sober, Elliott. 1984. *The Nature of Selection: Evolutionary Theory in Philosophical Focus*. Chicago: Chicago University Press.

Southgate, Christopher, and Andrew Robinson. 2010. "Interpretation and the Origin of Life." *Zygon: Journal of Religion and Science* 45:345–60.

Street, Sharon. 2006. "A Darwinian Dilemma for Realist Theories of Value." *Philosophical Studies* 127(1):109–66.