Evolutionary Debunking Arguments and Their Challenges to Human Knowledge

A thesis presented to

the faculty of

the College of Arts and Sciences of Ohio University

In partial fulfillment

of the requirements for the degree

Master of Arts

Andrés C. Ruiz

May 2013

This thesis titled

Evolutionary Debunking Arguments and their Challenges to Human Knowledge

by

ANDRÉS C. RUIZ

has been approved for

the Philosophy Department

and the College of Arts and Sciences

James Petrik

Associate Professor of Philosophy

Robert Frank

Dean, College of Arts and Sciences

Abstract

RUIZ, ANDRÉS C., M.A., May 2013, Philosophy

Evolutionary Debunking Arguments and their Challenges to Human Knowledge

Director of Thesis: James Petrik

I critically examine and evaluate the cogency of four kinds of evolutionary debunking arguments in the literature. Specifically, I focus on Alvin Plantinga's Evolutionary Argument Against Naturalism, an argument aimed at establishing the conclusion that naturalism and evolution lead to an epistemic defeater that renders the conjoined belief in both irrational; Michael Ruse's Evolutionary Ethics, aimed at establishing that an evolutionary genealogy of our moral sentiments proves morality to be an illusion; Sharon Street's Darwinian Dilemma Against Realist Theories of Value, which aims to establish the implausibility of natural selection having produced cognitive faculties that accurately track the sort of moral facts posited by moral realists; and Richard Joyce's Evolutionary Debunking of Morality, an argument to show that our moral judgments are unjustified and we ought to therefore adopt moral agnosticism. I argue that Alvin Plantinga fails to prove that the conjunction of naturalism and evolution lead to radical skepticism. Second, I argue that Michael Ruse draws the wrong conclusions from his evolutionary genealogy of morals and as a consequence fails to give a compelling argument against moral realism. By contrast, I defend Richard Joyce's and Sharon Street's arguments against various criticisms and conclude that they present a compelling epistemic challenge to justifications for moral realism.

Acknowledgments

I would like to thank Dr. Petrik for his patience and help throughout the course of writing this thesis. I would also like to thank Dr. Lent for introducing me to the crazy world of Philosophy and being a pivotal influence in my decision to pursue it academically. And of course I would like to thank the entire Philosophy department for providing a great learning environment that has allowed me to hone my thinking and writing skills and giving me a glimpse of the labyrinthine nature of the world of ideas.

Table of Contents

Introduction

In this thesis I propose to examine and critique evolutionary challenges to human knowledge. In particular I concentrate on Alvin Plantinga's Evolutionary Argument Against Naturalism, Michael Ruse's Evolutionary Ethics, Sharon Street's Darwinian Dilemma Against Realist Theories of Value, and Richard Joyce's Evolutionary Debunking of Morality. Further, I survey and critique already existing responses to these evolutionary arguments in the literature. I argue that Plantinga's and Michael Ruse's evolutionary arguments do not succeed in establishing their intended conclusions. By contrast, I defend Sharon Street's and Richard Joyce's arguments against critics and conclude that evolutionary considerations provide moral realism with a compelling epistemic challenge.

Background and Overview

Does the history of human evolution call into question our claims to possess certain kinds of knowledge? Numerous authors from differing fields of thought have advanced theses in the affirmative, including Alvin Plantinga's claim that the conjunction of evolution and naturalism entail global skepticism, Michael Ruse's contention that evolution provides a defeater to moral realism, Sharon Street's claim that Darwinian considerations cast doubt on our moral knowledge if construed in moral realist terms, and Richard Joyce's claim that evolutionary considerations provide us with an epistemic defeater with regards to our moral beliefs. I will examine some of the alleged implications of evolution to epistemology in general, and also, more narrowly, some of the alleged epistemic implications of evolution for ethics. I will also examine the strategies commonly employed in response to these evolutionary challenges. Having

done so, I will then consider a) whether there is a common theme to these responses and b) whether these responses have anything to contribute to larger skeptical worries about human knowledge.

The first author I consider is Alvin Plantinga and his Evolutionary Argument Against Naturalism (EAAN for short). Plantinga argues that the conjunction of naturalism and evolution leads to a skepticism that renders the belief in unguided-evolution self-referentially incoherent. The main idea is this: If the conjunction of naturalism and evolution are true then there is no good reason to trust the reliability of our cognitive faculties. That is because natural selection selects for adaptive behavior, survivability, and it isn't obvious that true beliefs always lead to adaptability and survival. In fact, false beliefs have equal prima facie survival potential. Since it isn't obvious that evolution selects for true beliefs, the reliability of our cognitive faculties is called into question. Thus, the first step in Plantinga's argument is as follows:

P (R/N & E) is either low or inscrutable (Plantinga, 2002, p. 4).

Here "R" stands for reliable cognitive faculties, "N" for Naturalism, and "E" for evolution. The claim being made is that the probability of our having reliable cognitive faculties given the conjunction of naturalism and evolution is low or inscrutable. Step two of Plantinga's argument the attempt to show that anyone who sees the truth of the above and who believes in the truth of N & E has a defeater for his belief in R. If you have a defeater for R then the final step is to argue that you have a defeater for all of your beliefs.

What to make of this argument? It raises many questions. A primary one is the relationship between beliefs and behavior. As Plantinga writes "our having evolved and

survived makes it likely that our cognitive faculties are reliable and our beliefs are for the most part true, only if it would be impossible or unlikely that creatures more or less like us should behave in fitness enhancing ways but nonetheless hold mostly false beliefs" (Plantinga, 2002, p. 5). What then is this relationship? Can a case be made for the likelihood of an organism's survival despite its holding mostly false beliefs?

Whatever the merits of Plantinga's argument, it is clear that evolution plays a central role. If successful, Plantinga's argument establishes that the conjunction of evolution and naturalism entail global skepticism. I will argue that Plantinga fails to provide any plausible scenarios in which creatures possessing thoroughgoing false beliefs about the world nevertheless behave in fitness enhancing ways that are likely to apply to *us*.

The second debunking argument I consider is Michael Ruse's attempt to show that evolutionary considerations reveal morality to be an illusion. Morality is illusory because there are no objective foundations to ground claims like "you ought to maximize personal liberty". According to Ruse, evolution reveals that moral judgments -thinking of certain actions as being called for or obligatory—are convenient falsehoods that evolution programmed into us so as to promote altruistic behavior, which itself is adaptive. Ruse's strategy is to show why evolutionary considerations reveal the illusory nature of morality. Nevertheless, even though illusory, Ruse maintains that morality can be vindicated because it is part of our biological natures to behave morally (most of the time) and it would be practically impossible to rid ourselves of it. Therefore, morality's illusory nature is not something that we should be fearful of. I will argue that Ruse's general strategy has promise, but that he reaches the wrong conclusions from his

evolutionary considerations. An evolutionary genealogy of the moral sentiments does not entail the non-existence of foundations (or truth-makers) to our moral claims; evolutionary considerations can at most call into question the justification for our moral judgments, not render them all false.

Sharon Street formulates an evolutionary challenge of her own that brings moral knowledge into doubt if that knowledge is grounded in a realist metaethics. Street argues as follows: evolutionary forces have played a role in shaping the content of human evaluative attitudes. This raises a question: how can we explain the relation between these evolutionary forces on our evaluative attitudes and the independent truths that moral realism posits? The realist, according to Street, has two alternatives: deny that there is a relationship between the evolutionary forces on our evaluative attitudes and the independent moral truths of realism, or affirm a relationship between the evolutionary forces and the moral truths of realism such that natural selection favored creatures who could grasp these moral truths. Both horns of the dilemma are unsatisfactory according to Street. The first horn leads to a conclusion unpalatable to the realist; namely, most of our evaluative attitudes are off track and as such we have no good reason to trust most of our moral beliefs. The second horn is unacceptable on scientific grounds. It is unacceptable on scientific grounds because the claim that true moral beliefs aid in the reproductive fitness of an organism would be a scientific hypothesis, one that would be in competition with rivals and thus subject to the standard battery of tests good scientific theories must be measured by. Street thinks that rival hypotheses to this claim fare better given that they share the same explanatory power while being much simpler in virtue of not having to appeal to sui generis moral facts. The result is that realist accounts of value cannot

accommodate the corrosive influence of Darwinian evolution on the development of our evaluative attitudes. Evolution thus gives us a defeater for our moral beliefs.

The first step in Street's argument is to show that natural selection has had a powerful influence on our evaluative attitudes. This raises the question of the relationship between the selective forces that have influenced the content of our evaluative judgments on the one hand, and the independent evaluative truths that realism posits on the other. If there is no relation then "we are left with the implausible skeptical conclusion that our evaluative judgments are in all likelihood mostly off-track, for our system of evaluative judgments is revealed to be utterly saturated and contaminated with illegitimate influence" (Street, 2006, p. 122). On the other hand, if there is a relationship, what is it? One such relation might be what Street calls the *tracking relation*. The tracking relation is this: "It is advantageous to recognize evaluative truths; for it surely promotes one's survival (and that of one's offspring) to be able to grasp what one has reason to do, believe, and feel. The evaluative judgments that proved most selectively advantageous to make are, in general, those that are true" (Street, 2006, p. 122).

Unfortunately for realists, it is not obvious that the tracking relation is true. Street offers a number of counterexamples that aim to show how (prima facie) false moral beliefs may nevertheless lead to adaptive behavior. She also notes that a thorough explanation for our having the moral beliefs we do does not need to invoke the truth of these beliefs in order to be exhaustive. I will defend Street's argument against objectors and will conclude that it remains a viable epistemic challenge to moral realism.

Similarly, Richard Joyce's evolutionary argument against moral realism, both the naturalistic and non-naturalistic kinds, focuses on the role natural selection had on

shaping our moral beliefs. He writes, “were it not for a certain social ancestry affecting our biology. . .we wouldn’t have concepts like obligation, virtue, property, desert, and fairness”, we should therefore “cultivate agnosticism regarding all positive beliefs involving these concepts until we find some solid evidence either for or against them” (Joyce, 2006, p. 180). Given that there is no reason to think that natural selection would have produced true moral beliefs, we should be moral skeptics. As with Street, I will defend Joyce’s argument against objectors, though I will conclude that his argument only threatens a certain conception of morality: moral realism.

What is clear from these thinkers is that the story of unguided Darwinian evolution has philosophical import, particularly in the realm of epistemology. This thesis will examine critically the alleged challenges evolution presents epistemology, the common core of these challenges, and will evaluate the viability of the various replies offered to these challenges.

Chapter 1: The Evolutionary Argument Against Naturalism

Alvin Plantinga argues that naturalism and contemporary evolutionary theory together entail a skepticism that undermines their conjunction. Defining naturalism as the thesis that "there aren't any supernatural beings--no such person as God, but also no other supernatural entities" (Plantinga, 2002, p. 1), Plantinga claims that the conjunction of naturalism with the belief that human beings have evolved via processes either as described by, or similar to, contemporary evolutionary theory is self-defeating or self-referentially incoherent. The conjunction of naturalism and evolution give rise to an epistemic *defeater* to whomever holds the belief that human cognitive processes are reliable. If one holds a defeater for the belief that one's cognitive faculties are reliable then one *also* has a defeater for any beliefs *produced* by these cognitive faculties, and among these beliefs is the conjunction of naturalism and evolution itself.

After explaining Plantinga's argument in more detail, I proceed to show that it fails. Though I concede that Plantinga succeeds in demonstrating the possibility of creatures who behave in fitness enhancing ways who yet hold mostly false beliefs, I contend that he fails to establish the *likelihood* of such creatures arising via naturalistic evolution, and much less that *we* are such creatures.

The Argument

In Plantinga's view, a cognitive faculty is reliable if, roughly, "the great bulk of its deliverances are true" *and* there is a true counter-factual of the form: if things were different then our beliefs would accordingly correspond to the altered state of affairs (Plantinga, 2002, p. 2). I take this to mean that a cognitive faculty is reliable if its output

of true beliefs is greater than its output of false beliefs and if it accurately tracks changes in the world so as to produce mostly true beliefs that reflect these changes.

If naturalism is true there is no guarantee that our cognitive faculties meet these criteria. We must ask whether on naturalism there is any good reason to suppose that our cognitive faculties would have developed to furnish us with mostly true beliefs. If the evolution of our cognitive faculties is unguided (in the sense that there is no intelligent being overseeing and directing the process of evolution towards some end) can we trust that it would lead to faculties that on the whole provide more true beliefs than false ones? This is not a new question. Plantinga quotes Darwin expressing a similar worry:

> With me the horrid doubt always arises whether the convictions of man's mind, which has been developed from the mind of the lower animals, are of any value or at all trustworthy. Would anyone trust in the convictions of a monkey's mind, if there are any convictions in such a mind? (Darwin, 315-316)

Patricia Churchland expresses the same worry:

> Boiled down to essentials, a nervous system enables the organism to succeed in the four F's: feeding, fleeing, fighting, and reproducing. The principle chore of nervous systems is to get the body parts where they should be in order that the organism may survive. . .Improvements in sensorimotor control confer an evolutionary advantage: a fancier style of representing is advantageous so long as it is geared towards the organism's way of life and enhances the organism's chances of survival. Truth, whatever that is, definitely takes the hindmost (Churchland, 1987, p. 548).

The worry is this: natural selection only "sees" behavior that proves adaptive to an organism. Natural selection does not select for true beliefs simply because they are true, natural selection only cares about behaviors that are *adaptive*. If natural selection is to select for true beliefs it must be because these beliefs *contribute to an organism's survival.* True beliefs are invisible to natural selection *unless* they convey survival advantage. Plantinga takes Churchland's point to be the following: "the objective probability that our cognitive faculties are reliable, given naturalism and given that we have been cobbled together by the processes to which contemporary evolutionary theory calls our attention, is low" (Plantinga, 2002, p. 4). We can put the claim as follows:

P (R/N&E) is low

where 'R' stands for reliable cognitive faculties, N for naturalism, and E for evolution (Plantinga, 2002, p.4). It is clear from this that the probability of our possessing reliable cognitive faculties will depend on the relationship between beliefs and behavior. If on the one hand there is a link between belief and behavior such that true beliefs convey survival advantage to the creatures who hold them then the probability of our being selected for on the basis of reliable cognitive faculties will be high. On the other hand, if there is no link between belief and behavior— for example, false beliefs convey as equal a survival advantage to organisms as true beliefs—then there will be no good reason to believe evolution equipped us with reliable cognitive faculties. This means that

> our having evolved and survived makes it likely that our cognitive faculties are reliable and our beliefs are for the most part true, only if it would be impossible or unlikely that creatures more or less like us should behave in fitness-enhancing ways but nonetheless hold mostly false beliefs (Plantinga, 2002, p. 5).

Can we rule out that it would be impossible or unlikely that creatures more or less like us should behave in fitness-enhancing ways but nonetheless hold mostly false beliefs? Plantinga thinks not. In order to determine the probability of a species possessing reliable cognitive mechanisms given naturalism and evolution we must think about the relationship between beliefs and behavior. Plantinga says there are four mutually exclusive and jointly exhaustive possibilities:

- Epiphenomenalism, the view that behavior is not caused by beliefs.
- Semantic Epiphenomenalism, the view that beliefs have causal efficacy with respect to behavior, but not in virtue of their *content*. Beliefs have causal efficacy with respect to *syntax*, not semantics.
- Beliefs are causally efficacious both semantically and syntactically, but maladaptive.
- Beliefs are both causally connected with behavior syntactically and semantically and are adaptive (Plantinga, 2002, ps 6-7)

On the epiphenomenological view, there is a causal relationship between behavior and beliefs, but that relationship only flows one way. Beliefs are *caused* by something, perhaps by molecular changes in the brain. On this view the physical causally interacts with the mental, but the converse is not also true. Mental states do not cause changes in the physical. If this view is true then according to Plantinga an organism's beliefs would be *invisible* to evolution. Given that these beliefs cause no change in an organism's behavior, whether a belief is true or false will cause no modification in behavior such that natural selection would select for certain types of beliefs over others. If beliefs are invisible to evolution then there is no reason to believe there is a link between true beliefs

and behavior such that true beliefs are likely to be selected for. On this view, the probability of our possessing reliable cognitive mechanisms would be very low. This, therefore, is not an open option to the naturalist (Plantinga, 2002, p. 6).

The second option, semantic epiphenomenalism, fails according to Plantinga because the semantic properties of beliefs are invisible to natural selection (Plantinga, 2002, p. 7). On this view, beliefs have two kinds of properties: semantic and syntactic properties. (Plantinga, 2002, p. 7) A belief's syntactic properties are its *physical* properties,

> the number of neurons involved in the belief, the connections between them, their firing thresholds, the rate and strength at which they fire, the way in which these change over time and in response to other neural activity, and so on. (Plantinga, 2002, ps. 7-8)

In contrast, the semantic content of a belief has to do with what a belief is *about*. "I am currently typing on a computer" is the *content* of my belief. So in addition to the physical properties of the belief in question there will be the semantic properties of the belief that tells us what the belief is about. Under this view, only the *syntactic* properties of a belief are causally efficacious. It is the *physical* properties of a belief that can modify behavior, not its content. If this is true then we have no reason to trust the reliability of our cognitive faculties since truth and falsehood are *semantic* properties of beliefs. If semantic properties aren't causally efficacious then they are invisible to natural selection. And if the truth or falsity of a belief has no role in its being selected for, then the link between true belief and behavior has been severed; therefore once again the probability of our having reliable cognitive faculties will be low (Plantinga, 2002, p. 7).

On the third view, beliefs are causally efficacious both syntactically and semantically yet are maladaptive. Under this view it would be better if organisms had no beliefs at all since, after all, they're maladaptive. Therefore the probability of reliable cognitive faculties will also be low (Plantinga, 2002, p. 8).

The fourth view is the most promising: beliefs are causally efficacious syntactically and semantically and adaptive. In this view beliefs causally interact with the world and have a role in the behavior of an organism and the behavior is adaptive. Plantinga thinks the probability of reliable cognitive faculties under this view isn't as high as one might initially think because desires, not just beliefs, produce behavior. It is belief-desire combinations that are responsible for the modification of behavior in organisms. Since natural selection only "sees" behavior and only "cares" about *adaptive* behavior, it will only select for adaptive behavior. The crucial move in Plantinga's argument is this: "For any given adaptive action, there will be many belief-desire combinations that could produce that action; and very many of those belief-desire combinations will be such that the belief involved is false" (Plantinga, 2002, p. 9). To illustrate:

> Perhaps Paul very much *likes* the idea of being eaten, but when he sees a tiger, always runs off looking for a better prospect, because he thinks it unlikely that the tiger he sees will eat him. This will get his body parts in the right place so far as survival is concerned, without involving much by way of true belief. . .Or perhaps he thinks the tiger is a large, friendly, cuddly pussycat and want to pet it; but he also believes that the best way to pet it is to run away from it. . .or perhaps he thinks the tiger is a regularly recurring illusion, and, hoping to keep his weight

> down, has formed the resolution to run a mile at top speed whenever presented with such an illusion; or perhaps he thinks he is about to take part in a 1600 meter race, wants to win, and believes the appearance of the tiger is the starting signal; or perhaps. . .Clearly there are any number of belief-cum-desire systems that equally fit a given bit of behavior (Plantinga, 2002, p. 8).

Evolution only cares about an organism being in the right place at the right time. Clearly there are many belief-desire combinations that can produce behavior in an organism that will get it away from danger. Since the number of false belief-desire combinations that can prove adaptive seems just as high as true ones, the probability of one of those false belief-desire combinations being selected for is as likely as the true belief-desire combination or is inscrutable. So much for individual beliefs. Is it plausible to think of an organism holding *thoroughgoing* false beliefs that are yet still adaptive? According to Plantinga, the answer is yes:

> Perhaps Paul is a sort of early Leibnizian and thinks everything is conscious (and suppose that is false); furthermore, his ways of referring to things all involve definite descriptions that entail consciousness, so that all of his beliefs are of the form *That so-and-so conscious being is such-and-such.* Perhaps he is an animist and thinks everything is alive. Perhaps he thinks all the plants and animals in his vicinity are witches, and his ways of referring to them all involve definite descriptions entailing witchhood. But this would be entirely compatible with his belief's being adaptive; so it is clear, I think, that there would be many ways in which Paul's beliefs could be for the most part false, but adaptive nonetheless. (Plantinga, 2002, p. 9)

To recap: According to Plantinga, there is no clear connection between the truth of a belief and its ability to promote survival. Clearly some true beliefs promote the continued survival of an organism, *but so can many false ones.* In fact, Plantinga thinks an organism whose beliefs are *all* false can nevertheless behave in adaptive ways that ensure its continued existence. Since natural selection only "cares" about adaptive behavior, it will select for any mechanisms that get a creature's body parts in the right place, and among these mechanisms are any cognitive mechanisms that, while unreliable, nevertheless prove adaptive. Since the probability that evolution selected for a cognitive mechanism that is reliable is either low or inscrutable, we have no reason to believe *we* possess such a mechanism, and so all of our beliefs are undermined. Is there any way to avoid this conclusion? I think so.

Plantinga's argument rests on the weighted probability of our possessing reliable cognitive mechanisms given the four above possibilities. Even if the probability of a reliable cognitive process given, say, the second view, were high, one would still not be justified in believing we possess such reliable cognitive processes because the probability of *that* rests on the probability of said mechanism arising *given* all four views. The naturalist however needn't be too afraid. If one can rule out some of the above then the probability will be the new weighted average given the remaining views. If one can rule out, say, semantic epiphenomenalism, then even if the probability of our having reliable cognitive faculties is low on that view, it would not hurt the weighted average since we've ruled out *that* possibility. One route to attack Plantinga's argument then would be to simply provide independent arguments against the views unlikely to be true. The view

that beliefs have no effect on behavior and the view that beliefs are maladaptive should be ruled out. This is because cognition is very costly.

> The human brain makes up about 2% of body mass, but accounts for about 20% of oxygen consumption. Having beliefs, whether true or false, comes at a high price. The hypothesis that belief has no effect on behavior, or that having beliefs reduces fitness, are non-starters. If this was the case then there would be strong selection for not having beliefs (Griffiths & Wilkins, forthcoming, p. 5).

The probability of our possessing reliable cognitive faculties given semantic epiphenomenalism and on the view that beliefs are maladaptive is low, but the probability that *these* views are correct is also very low. We not only have good *biological* grounds for rejecting them, but even if these biological grounds were shown to be irrelevant, and Plantinga would need to provide some argument as to why we should expect these views to be true given their unlikelihood, we would *still* have the recourse of turning to the arguments in Philosophy of Mind against semantic epiphenomenalism as a means to get rid of this possibility from Plantinga's general probability calculus (R. Taylor, 1963; K. Popper, 1977). Given this, it seems reasonable to conclude that we have good reasons for only worrying about the fourth scenario, and seeing whether we have any good reason to believe we can possess reliable cognitive faculties under *this* particular scenario. I argue that we do.

First, there are two ways to understand the above challenge. Many of Plantinga's critics fall for the first (and wrong) way to understand the offered scenarios and therefore give interesting yet irrelevant detours that leave the argument unscathed. I'll briefly sketch out these responses and then move on to what I think is the right way to approach

Plantinga. The first way to understand the argument is this: Plantinga needs to describe a scenario in which an organism holds *on average* mostly false beliefs about its environment (at least over half of its beliefs must be false) and yet whose behavior proves adaptive. In this first reading of Plantinga's example only *part* of the definite description is false, the part that ascribes consciousness to the objects in question. If the organism is describing a chair in the following way: "That wooden conscious chair is located in my bedroom and is used for sitting", I have made a mostly true claim about the object in question. My beliefs about the chair are still mostly true, only my belief that it is conscious is false. If the organism believes that *everything* is conscious, that organism may have a thoroughgoing false idea about the nature of the universe, but it doesn't follow that most of its beliefs are false and that its cognitive faculties are unreliable. The organism in question will still hold mostly true beliefs about the world so long as its other descriptions of objects are accurate, descriptions about the shape of things, their material composition, their spatio-temporal locations, their origins, etc. As an analogy, if I believe the fundamental nature of the universe resembles that of idealism in Berkeley's sense—i.e., at bottom all that exist are minds and their ideas—and we suppose that materialism is true, then I'll have a thoroughgoing false belief about the nature of the universe. And yet, I will still possess mostly true beliefs about the world. If I see a chair in front of me I will (correctly) believe "There is a chair in front of me." That the underlying metaphysical structure of the chair is material, and I believe it to be of an idea-like substance does not make my belief that there is a chair in front of me false. My beliefs about the chair's shape, its distance from me, who built it, what its used for, etc. will all still be true. On this first reading, what Plantinga has described is a scenario in which for any object X,

we will hold at least one false belief about X. And yet, given the large number of beliefs we have about most objects, as long as those are true, the number of true beliefs we possess will be higher than the number of false beliefs, and so there is no problem. What Plantinga needs to give is a plausible story as to how an organism can possess a higher number of false beliefs about the world than true beliefs, and yet who may nevertheless behave in ways that prove adaptive from the point of view of natural selection. The mere logical *possibility* of such scenarios isn't enough, we need a coherent story *that is likely to actually have occurred*, until then all we're left with is an "IOU".

To put it more succinctly: if I believe a painting is conscious, it does not follow that most of my beliefs about the painting are false. If I believe that *everything* is conscious, it doesn't follow that my beliefs are for the most part false. As a contemporary similar line of argumentation, consider something like the Matrix hypothesis. According to David Chalmers, the Matrix hypothesis is equivalent to the following three propositions:

- Physical Processes are fundamentally computational (Chalmers, 2009, p. 36).
- Our cognitive systems are separate from physical processes (Chalmers, 2009, p. 36).
- Physical Reality was created by beings outside physical space-time (Chalmers, 2009, p. 36).

If it were the case that we were in the Matrix (or any brain in a vat scenario), no skeptical conclusions should be derived. That is because, as Chalmers argues, nothing about this Metaphysical hypothesis is skeptical:

> We still have bodies, and there are still chairs and tables: it's just that their fundamental nature is a bit different from what we may have thought. In this manner, the Metaphysical hypothesis is analogous to a physical hypothesis, such as one involving quantum mechanics. Both the physical hypothesis and the Metaphysical hypothesis tell us about the processes underlying chairs. They do not entail that there are no chairs. Rather, they tell us what chairs are really like (Chalmers, 2009, p. 37).

While Chalmers is not directly responding to Plantinga, I bring him into consideration given how common it has been for critics of Plantinga to read him *as if* he's arguing that the falsity of an underlying metaphysical ontology makes our *everyday* beliefs false.[1] Critics have then replied to Plantinga by adopting a similar line of argumentation as Chalmers. This is a misreading. If we pay close attention we see that what is doing all the work in Plantinga's early Leibnizian scenario is the creature's use of *definite descriptions* to refer to objects in such a way that all of its utterances and beliefs entail a falsehood. The scenario is set up in a way such that every proposition is "contaminated" by the false ascription of creaturehood to every object the Leibnizian individual has any beliefs about. The problem is therefore not one of *layers* in which false beliefs about bottom level phenomena contaminate upper level beliefs, but rather about *language* contaminating the entire infrastructure. The source of the deep error then is not merely having a false belief about the bottom level of reality that then "infects" the upper levels. The source of systematic error in Plantinga's scenario is the use of *definite descriptions that entail the falsehood* of the propositions in question. In what follows I

[1] See Ramsey (2002), page 26 and Fodor (2002), page 34.

will present what I take to be the right reading of Plantinga's challenge and will then critique it by adopting William Ramsey's strategy invoking the causal theory of reference.

As a clarification of his position designed to eliminate these kinds of misreadings, Plantinga writes:

> Suppose naturalism is true and in fact there is no such person as God. Now several naturalists have argued that belief in God, while false, is nonetheless adaptive. So suppose a tribe of cognitively gifted creatures believe that everything (except God Himself) has been created by God; they therefore think everything is a *creature*, something created by God. Suppose further that their only way of referring to the various things in their environment is by way of such definite descriptions as 'the tree creature before me' or 'the tiger creature approaching me'. Suppose still further that all their beliefs are properly expressed by singular sentences whose subjects are definite descriptions expressing properties that entail the property of creaturehood—such sentences as 'The tiger creature approaching me is dangerous' or 'The tree creature before me is full of apple creatures'. Suppose, finally, that their definite descriptions work the way Bertrand Russell thought definite descriptions work: 'The tallest man in Boston is wise', for example, abbreviates 'There is exactly one tallest man in Boston, and it is wise.' Then, from the naturalist perspective all their beliefs are false. Yet these can still be adaptive: all they have to do is ascribe the right properties to the right creatures (Plantinga, 2002, p. 260).

In a footnote, Plantinga writes

> It is worth remembering that the question here is whether it is possible that there be creatures like us whose beliefs are for the most part false but nonetheless adaptive. The question is not, therefore, whether we use definite descriptions in the Russellian way, or whether our predicates express properties including 'being created by God' (Plantinga, 2002, p. 261).

It isn't clear why Plantinga doesn't think it relevant whether these creatures use language in a similar way to us. What is Plantinga's motivation in talking about creatures *like* us rather than ourselves? He writes

> To try to guard against interspecific chauvinism, I suggested that we think, not about ourselves and our behavior, but about a population of creatures a lot like us on a planet a lot like earth (Darwin suggested we think about monkeys in this connection). We imagine furthermore that they and their cognitive systems have evolved by way of the mechanisms to which contemporary evolutionary theory direct our attention, unguided by the hand of God or anyone else. Now what is the P(R/N&E), specified, not to us, but to them? (Plantinga, 2002, p. 5)

Further, he writes

> Now return to Darwin's doubt, and observe that if this is the sensible attitude to take to P(R/N&E) specified to that hypothetical population, then it will also be the sensible attitude towards P(R/N&E) specified to us. We are relevantly like them in that *our* cognitive faculties have the same kind of origin and provenance as *theirs* are hypothesized to have (Plantinga, 2002, ps. 10-11).

Notice here something very important: the group of creatures Plantinga imagines using Russellian definite descriptions that entail creaturehood for any object they talk about do

not possess cognitively unreliable faculties because of their origins and provenance. Their beliefs are false because *their language use* entails the falsity of the things they talk about. Contrary to Plantinga, it *is* relevant to point out that *our* language does not function this way. At least if we accept the causal theory of reference. The proper way to address Plantinga then is to deny that our language works in the way his imagined early Leibnizian's language works:

> Consider first a sentence that might be uttered by someone, such as 'fire is dangerous.' Following the work of Saul Kripke and others, there is now widespread agreement that in such an utterance, the reference of the term 'fire' is due in large measure to causal connections that exist between the speaker and actual instances of fire. An important consequence of this view is that the speaker can refer to fire, and can predicate truths about fire, even if she suffers from profound ignorance and/or false beliefs about the nature of fire itself. . .It seems quite obvious that a person can truthfully say fire is dangerous—and really be referring to fire—even though the person mistakenly thinks that fire is, say, essentially a substance (Ramsey, 2002, p. 27).

The creatures in question have defeaters for their beliefs *given the way their language works.* If one were to "fix" their language then they would not possess a systematic defeater for all of their beliefs. Now, given that *our* language does *not* work in this way then the origin of *their* defeaters, their language, does not carry over to *us*. Jerry Fodor agrees:

> But the question nonetheless arises, how, in our case, did evolution go about getting what it wanted? Darwinism is, remember, a *historical* thesis about *us*. If

> it's to jeopardize Scientism, it's got to underwrite the claim that our minds actually evolved under selection that favored us *because* our beliefs were generally successful. . . (2002, p. 37)

and

> what's required, to repeat, is a plausible historical scenario according to which *our* minds were selected because we had lots of behaviorally successful false beliefs (2002, p. 37).

Of course it *is* possible that such creatures exist, but mere logical possibility isn't enough here. Perhaps there is only one coherent and consistent scenario in which the bulk or the whole of a creature's beliefs turn out to be false and nevertheless proves adaptive, and Plantinga has merely stumbled upon that one. Plantinga needs to give a *whole range* of scenarios in which the *bulk* of a creature's beliefs turn out to be false and yet somehow the creature behaves in ways that prove adaptive. The Russellian scenario is only one possibility, and yet we know that one does not apply to *us*. How many other scenarios are there? Plantinga provides very few. It must be remembered that Plantinga's argument for belief-desire combinations proving adaptive for *individual* beliefs gets off the ground only given that there are *many* such possible combinations. He writes that "for any given adaptive action, there will be many belief-desire combinations that could produce that action; and very many of those belief-desire combinations will be such that the belief involved is false" (Plantinga, 2002, p. 9). If you were to put all the possible combinations, true and false, into a hat and draw randomly from that hat, the probability that you would draw a false belief-desire combination would be high.

The same is then necessary for belief forming *mechanisms.* Plantinga needs to show that there are *many* scenarios in which false belief forming mechanisms can arise and lead to adaptive behavior. Another scenario he provides is equally suspect: he writes that

> Michael Rea argues that naturalism implies an ontology of gunk, an ontology according to which there really aren't any *objects* (although there is a sort of continuous gunk or goo which may, as they say, be 'propertied' differently in different places.) Suppose Rea is right: then since most of our beliefs imply that there are objects, most of our beliefs will be false. Still, our natural way of cutting up the world into objects could be adaptive; if so, most of our beliefs would be false but adaptive (Plantinga, 2002, p. 261).

Previously I argued that objections to Plantinga that are similar to Chalmers' objections to Matrix-type skeptical arguments are based on a misreading of Plantinga and therefore fail to engage with the most important component of his argument, the use of definite descriptions. Those types of responses miss the mark when employed against the *Early Leibnizian* scenario offered by Plantinga. However, it *is* appropriate to reply in a Chalmers fashion against *this* particular scenario. This is because this scenario is similar to the Matrix Hypothesis that Chalmers argued does not lead to wide skepticism. A "gunky" metaphysics merely entails that the "bottom level" of reality is a certain way: namely, that there *isn't* one, you can keep "chopping" each level in into smaller and smaller pieces *ad infinitum*. There are two ways of interpreting Plantinga here. One is that the *semantics* of everyday life presuppose objecthood, and therefore, if there are no objects, all sentences presupposing their existence are false. The second reading is that

our *beliefs* presuppose objecthood and therefore if the bottom layer of reality turns out to be gunky then beliefs about the "upper" levels of reality that presuppose the falsehood of gunk will themselves be false. I will take these in turn.

If our semantics about every day objects *presupposed* the truth of objecthood *then* the fact that reality turned out to be "gunky" would render all talk about everyday objects false. But Plantinga can't simply help himself to the assumption that our language *does* presuppose objecthood. This semantic thesis has been denied by some (Van Inwagen, 1990) and is far from a benign assumption to grant. Plantinga needs to *argue* for that semantic thesis.

The second interpretation can be challenged by arguing, as Ramsey does, that our talk of everyday objects does not depend on any given metaphysical thesis being correct. Just as the thesis of Quantum Mechanics did not, and does not, threaten our talk about everyday objects, a gunky topology does not either. There are still chairs, and I still exist. Plantinga's gunk scenario therefore falls prey to Chalmers and Ramsey style replies.

What Plantinga needs to argue is that there are a great number of different ways in which belief forming mechanisms leading to thoroughgoing false beliefs can prove adaptive, or at least sketch out an adaptive scenario the probability of which is as great as that of the combined probability of its rivals. Given that Plantinga provides only a few scenarios where thoroughgoing false beliefs still prove adaptive, and given how suspect each of these scenarios is, it is not unreasonable to remain unpersuaded by his line of argumentation. Plantinga owes us more detailed accounts than he provides. If we were to throw in all the false belief forming mechanisms along with the true ones into a hat, you don't get a high probability of picking a false one at random given that we know of a very

few number of such possible scenarios (Plantinga has provided a few, and in my view only the early Leibnizian is a scenario in which the creatures do have thoroughgoing false beliefs, and even then, we know this scenario does not apply to *us*).[2]

It may not be enough to simply point out how questionable Plantinga's examples are. Merely showing that he has not provided an account that would establish the soundness of his argument is not the same as establishing the falsity of his conclusion. Evan Fales provides an example that should make us doubt whether *any* scenarios of thoroughgoing false beliefs could nevertheless prove adaptive. Suppose that:

> Freddy who is carrying a heavy rock he falsely believes to be light and soft, nearly steps on a puff adder. Believing that being hit by something light and soft will be fatal for the adder (also false), he quickly drops the rock on it, and lives to see another day. So far so good for Freddy. Continuing on with his rock, Freddy encounters an angry warthog on the trail. Still believing the rock to be light as a feather, and believing (falsely) that dancing upon something light deters warthogs, Freddy proceeds to do a two-step on top of the rock directly in the path of the charging pig. The moral of this fable is plain: there are no effective algorithms connecting false belief to appropriate action, as there are when the input is true beliefs and the rules of inference employed are valid or inductively sound. Intelligent action is hard enough for a brain to manage; burdening it with ever-changing completely arbitrary principles would make the task impossible. Freddy

---

[2] Of course this presupposes the causal theory of reference, something I cannot argue for here. If this is too contentious, we need not assume any *positive* thesis about the way our language works, we may simply assume a *negative* thesis, namely that our language does not work in a Russellian fashion, something Plantinga would be likely to grant.
See Kripke, S. (1980), *Naming and Necessity*, Cambridge: Harvard University Press

may survive the adder, but he'll not survive long. Nor will his genetic heritage (Fales, 2002, p. 51).

The key claim in the above is that "there are no effective algorithms connecting false belief to appropriate action, as there are when the input is true beliefs and the rules of inference employed are valid or inductively sound" (Fales, 2002, p. 51).[3]

Let us suppose that Plantinga can provide such an account however. Then if he is correct, the probability of our possessing reliable cognitive faculties given naturalism and evolution "will be the weighted average of P(R/N&E&Pi) for each of the four possibilities Pi --weighted by the probabilities, on N&E, of those possibilities" (Plantinga, 2002, p. 9).

The formula is thus:

P(R/N&E) = (P(R/N&E&P1) x P(P1(N&E)) + (P(R/N&E&P2) x P(P2/N&E)) + (P(R/N&E&P3) x P(P3/N&E)) + (P(R/N&E&P4) x P(P4/N&E)). (Plantinga, 2002, p. 9)

And yet, as we have seen, one crucial step in Plantinga's attempt to demonstrate that the probability of our possessing reliable cognitive faculties given naturalism and evolution is to show how widespread false beliefs may nevertheless prove adaptive and show that this is a *likely consequence* of evolution. I have argued that Plantinga has not given us any good reason to believe this is true.

Next we will turn to another species of evolutionary debunking argument. While Plantinga's evolutionary argument against naturalism is meant as a *far reaching* skeptical argument threatening the reliability and trustworthiness of *all* of our beliefs, the

---

[3]Recently Stephen Law has argued this point at more length, though I do not have the space or time to discuss the technical details (S. Law, 2011).

arguments I will now be examining will be more modest in their goals: their conclusions will merely entail the untrustworthiness of a *restricted* domain of human knowledge, the *moral* domain.

Chapter 2: Michael Ruse's Evolutionary Ethics: A Phoenix Arisen

Evolutionary debunking arguments come in different forms. To better appreciate the differences between the arguments I will be tackling, it will be helpful to sketch out Richard Joyce's taxonomy of the three kinds of debunking in the literature. The first kind of debunking is *truth debunking*. This particular type of debunking aims to show that "all moral assertions are untrue" (Joyce, 2013, p. 3). As an example of such an extreme form of debunking Joyce points to Mackie's moral error-theory. Adherents of this theory hold that the purpose of moral language is to pick out certain features of the world, but as it happens, there are no such features, and therefore moral utterances are all false. "Softer" versions of truth debunking are versions that do not aim to establish the falsity of *all* moral utterances but of *restricted domains* of such utterances, as in perhaps aiming to establish that all talk of "rights" is false and misguided.

The second kind of debunking is *theory debunking*. This particular kind of debunking is "aimed at showing that certain *interpretations* of moral judgments are false" (Joyce, 2013, p. 3). This particular brand of debunking aims to tackle what the *truth-makers* for certain moral utterances are by showing that they are false in favor of some other truth-maker. For example, Joyce explains, when one says

> "keeping promises is morally good", one may hold that it is good "because God commands promise-keeping, or because it maximizes happiness, or because it follows from the optimally coherent equilibrium of our intuitions". To debunk one of these theories would probably leave the truth value of the claim 'keeping promises is morally good' intact (Joyce, 2013, p. 4).

The third form of debunking is what Joyce calls *justification debunking*. This type of debunking does not aim to show that all or a specific domain of our beliefs are *false* but rather aims to show that even if by chance our moral judgments turned out to be true they would still be *unjustified.* It is this third type of debunking that Joyce is engaged in. Joyce's argument is one that aims to show that evolutionary considerations put us in a position to call into question the *justification* of our moral judgments.

I will start my survey of evolutionary debunking arguments aimed at the debunking of moral truth by examining Michael Ruse's account. Ruse believes that evolutionary considerations are all that are needed to demonstrate why morality is an illusion and that moral claims are therefore false. Even if illusory, Ruse contends that there is no reason to worry, as morality is such an integral part of our biological natures that it is unlikely, or perhaps impossible, to shed ourselves of it upon discovering this fact. I will argue that Ruse's two goals: to show why morality is an illusion, and to show why even if it is an illusion morality is vindicated, both fail. I will use Ruse's account as an example of an evolutionary debunking argument that does not work.

Michael Ruse's Evolutionary Account:

Ruse argues that the human propensity for moral altruism, which he defines as "a conscious being helping others because it is right and proper to do so" (Ruse, 1986, p. 98) is grounded in *evolutionary* altruism, "the working together for biological payoff" (Ruse, 1986, p. 98). While there is no necessary link between the two—ants engage in evolutionary altruism but not in moral altruism—Ruse claims that moral altruism is *one way* that evolution could have secured our propensity to engage in evolutionary altruism. He writes that "humans are the kinds of animals which benefit biologically from

cooperation within their groups, and literal, moral altruism is the way in which we achieve that end" (Ruse, 1986, p. 98). The basic idea is that believing that some cooperative actions are "called for" morally speaking and other actions as uncalled for will make it easier to motivate individuals to behave in ways that are conducive to biological payoff.

There are reasons why evolution would imbue us with moral concepts. As Ruse notes, there are multiple avenues by which evolution could promote altruistic behavior. One avenue is by programming us in the same way ants are programmed, by genetic hardwiring. Ants are "machine like, working in their nests according to innate dispositions, triggered by chemicals (pheromones) and the like" (Ruse, 1986, p. 98). There are advantages to this approach: "it eliminates the need for learning, it cuts down on the mistakes, and much more" (1986, p. 98). These advantages come at a cost however, as flexibility is compromised, "if circumstances change, individual ants cannot respond" (1986, p. 98). In the case of ants these costs are acceptable because (biologically speaking) ants are rather cheap to produce. Humans on the other hand require much more biological investment and so "apparently the production of 'altruism' through innate, unalterable forces, poses too much of a risk" (1986, p. 98).

Another route evolution could take to achieve altruistic behavior in individuals stands at the other extreme; namely, the evolution of super-brains: "rationally calculating at each point if a certain course of action is in our best interests" (1986, p. 98). This is a curious example given that altruism usually means *selfless* behavior. It seems that what Ruse means by altruism here is not selfless acts that are beneficial to others but rather a more general "any act that helps others". Either way, nothing much hinges on this, this is

clearly not the way we evolved, and this route itself introduces great biological costs since “by the time I have decided whether or not to save the child from the speeding bus, the dreadful event has occurred” (1986, p. 99).

According to Ruse we are somewhere in the middle of these two. He writes:

> To make us cooperate for our biological ends, evolution has filled us full of thoughts about right and wrong, the need to help our fellows, and so forth. We are obviously not totally selfless. Indeed, thanks to the struggle for reproduction, our normal disposition is to look after ourselves. However, it is in our biological interests to cooperate. Thus we have evolved innate mental dispositions inclining us to cooperate, in the name of this thing which we call morality. We have no choice about the morality of which we are aware. But, unlike the ants, we can certainly choose whether or not to obey the dictates of our conscience. We are not blindly locked into our courses of action like robots. We are inclined to behave morally but not predestined to such a policy (Ruse, 1986, p. 99).

Our moral sense is therefore an adaptation:

> a feature helping us in the struggle for existence and reproduction—no less than hands and eyes, teeth and feet. It is a cost effective way of getting us to cooperate, which avoids both the pitfalls of blind action and the expense of a superbrain of pure rationality (Ruse, 1986, p. 99).

Before proceeding, it should be noted that Ruse’s strategy is a strange one. His main goal is to show that evolutionary considerations reveal “objective morality” to be an illusion, while at the same time aiming to show that the evolutionary history of our moral

sentiments can help *reinforce* our commitment to moral values and obligations. He writes:

> Ultimately, there is no reasoned justification for ethics in the sense of foundations to which one can appeal in reasoned argument. All one can offer is a causal argument to show why we hold ethical beliefs. But once such an argument is offered, we can see that this is all that is needed (Ruse, 1986, p. 102).

If Ruse thinks his account can *epistemically* reinforce our commitment to moral values then he is mistaken and contradicting himself. If he means *psychological* commitment then he also fails because an individual questioning whether there is *anything at all* that he *ought* to do might find little comfort in learning the evolutionary history of what led him to ask this sort of question in the first place. Evolutionary considerations might be just as likely to lead someone to jettison all the moral obligations one thinks one has in favor of self-interested behavior as they are to reinforce them (L. Young, 2013).

When Ruse moves on to consider the implications of this evolutionary story about our moral sentiments, he starts by saying that "the whole point of today's approach is that we transcend a rugged struggle for existence—in thought and deed" (1986, p. 99). He writes that "what excites the evolutionist is the fact that we have feelings of moral obligation laid over our brute biological nature, inclining us to be decent for altruistic reasons" (1986, p. 100). He then claims that if we're looking for moral guidance, looking for principles which we might adopt to conduct our behavior, we can look at the precedent set by John Rawls. He quotes the following line from Rawls approvingly:

> The guiding idea is that the principles of justice for the basic structure of society are. . .the principles that free and rational persons concerned to further their own

> interests would accept in an initial position of equality as defining the fundamental terms of their association. These principles are to regulate all further agreements; they specify the kinds of social cooperation that can be entered into and the forms of government that can be established. This way of regarding the principles of justice I shall call justice as fairness (Rawls, 1971, 11).

Ruse thinks the Rawlsian and the "evolutionist" (the term he uses for someone who embraces the views he's promoting) are in similar positions.

> For both the biologist and the Rawlsian, the question is that of how one might obtain right action from groups of people whose natural inclination is (or rather, of whom one would expect the natural inclination to be) that of looking after themselves. In both cases the answer is found in a form of enlightened self-interest. We behave morally because, ultimately, there is more in it for us than if we do not (Ruse, 1986, p. 101).

It is unclear if this last line is meant to be taken as a descriptive or normative claim. The key question to ask is whether we're engaging in Rawlsian reflection in order to find action guidance that simply *makes sense* for self-interested individuals to engage in or if the deliberation is supposed to result in an output of normative rules that we would all be *obligated* to follow. Are the prescriptions of the Rawlsian veil of ignorance merely *pragmatic* for the evolutionist? If so then they're not *moral requirements,* they're mere arrangements that make sense for us to follow in order to prevent society from collapsing, but there would be no reason to believe we are *obligated* to act in these ways. This brings us to the main question at hand: what *metaethical* implications does Ruse's view have?

According to Ruse, the metaethical implication of evolution is something akin to an error-theory of morality. He writes:

> In particular, the evolutionist argues that, thanks to our science, we see that claims like "You ought to maximize personal liberty" are no more than subjective expressions, impressed upon our thinking because of their adaptive value. In other words, we see that morality has no philosophically objective foundation. It is just an illusion, fobbed off on us to promote biological altruism (Ruse, 1986, p. 102).

He continues:

> The evolutionist is no longer attempting to derive morality from factual foundations. His/her claim now is that there are no foundations of any sort from which to derive morality—be these foundations evolution, God's will or whatever. Since, clearly, ethics is not nonexistent, the evolutionist locates our moral feelings simply in the subjective nature of human psychology. At this level, morality has no more (and no less) status than that of the terror we feel at the unknown—another emotion which undoubtedly has good biological adaptive value (Ruse, 1986, p. 102).

This is far too quick. Nothing about Ruse's evolutionary explanation for the origins of our moral sense has done anything to call into question whether or not there actually exists an ontological *grounding* for claims like "you ought to maximize personal liberty". Ruse seems to jump from giving a plausible descriptive account about the origins of a moral sense to the conclusion that all moral claims are groundless and mere expressions of our subjective feelings. The problem is that this is a *non-sequitur* without the aid of supporting premises which Ruse does not supply. A further problem with

Ruse's conclusion is that his evolutionary story has done nothing to illuminate what the *function* of moral utterances is. It is difficult to know what he means by "subjective expressions", is he here embracing an emotive theory of ethics according to which the function of moral utterances is to express *attitudes* rather than beliefs (R. Brandt, 1950)? It isn't clear from the above.

A further worry is that Ruse merely *asserts* that "there are no foundations of any sort from which to derive morality—be these foundations evolution, God's will, or whatever" (Ruse, 1986, p. 102). But has he shown this? An evolutionary account of the development of our moral sense might tell us why we have a propensity to make moral judgments, why we view the world through normative lens and the like, but it does *not* by itself show that there can be no foundation (and I take it to mean that by "foundation" Ruse means truth-makers) from which to derive morality. Ruse is making an unwarranted leap from a descriptive account of the genealogy of our moral sense to a *metaphysical* claim about the *ontological grounding* of the moral judgments we make through our moral sense. I agree with Street and Joyce that such a genealogy can cast doubt on and call into question the *justification* of our current moral beliefs. However, once we have reflected on whatever the best available evolutionary account of the development of our moral sense tells us is the reason why we tend to make moral judgments, we might then come to realize that a question like "but is it *really* obligatory for me to maximize utility?" remains open. A genealogy of the development of our moral sense might render my belief in the principle of utility *unjustified,* but it does *not* show that there *cannot* be a truth-maker ("grounding", "foundation", or what have you) for the principle. Maybe I don't have *access* to this truth-maker, but again, a genealogy of the moral sense does not

entail this negative existential. Once we have the best story science can offer us, we may *then* say something like "Okay, now we have an empirically confirmed story on hand as to why human beings tend to make the normative judgments they do, so perhaps our *pre-philosophical* moral judgments might be unjustified, but that does not mean reason will not help us uncover a justification for the principle of utility or the categorical imperative". As we will see in the next chapter, Sharon Street and Richard Joyce don't think further reflection will help us arrive at justified moral beliefs once we know our starting judgments have been contaminated by the corrosive influence of evolutionary pressures, so for them there is no way to get uncontaminated justified moral judgments from an unreliable starting pool of contaminated ones, but as a response to Ruse, it is a legitimate way to respond.

On Ruse's view, an individual equipped with this evolutionary genealogy will come to realize that his belief in the principle of utility is false. It is false because there are no foundations (or truthmakers) upon which it can be grounded. And yet my argument is that the genealogy hasn't shown *that.* As we shall see in the next chapter, Richard Joyce draws a different conclusion. An individual equipped with an evolutionary genealogy should come to see that his belief in the principle of utility is *unjustified.* This follows much less problematically than Ruse's claim and is the reason I'm sympathetic with Joyce and object to Ruse.

Christine Korsgaard has made a similar point, arguing that these types of evolutionary accounts confuse *third-person* explanations as being sufficient to answer normative questions that can only be answered from the *first-person* point of view. She writes:

> The case of the evolutionary theory shows that a theory could be adequate for the purposes of explanation and still not answer the normative question. And there is an important reason for this. The question how we explain moral behavior is a third-person, theoretical question, a question about why a certain species of intelligent animals behaves in a certain way. The normative question is a first-person question that arises for the moral agent who must actually do what morality says. When you want to know what a philosopher's theory of normativity is, you must place yourself in the position of an *agent* on whom morality is making a difficult claim. You then ask the philosopher: must I really do this? Why must I do it? And his answer is the answer to the normative question (Korsgaard, 1996, p. 16).

There is an argument Ruse brings up later that sheds light on why he thinks evolutionary considerations might undermine the search for foundations to ethical claims. He asks us to consider morality as analogous to Ouija boards, writing:

> During the First World War, many bereaved parents turned to spiritualism for solace. Down the Ouija board would come the messages: "It's alright Mum. I've gone to a far better place. I'm just waiting for you and Dad." I take it that these were not in fact the words of the late Private Higgins, speaking from beyond. Rather they were illusory—a function of people's psychology as they projected their wishes. . .The moral to be drawn from this little story is that we do not need any further justificatory foundation for "It's alright Mum" than that just given. At this point, we do not need a reasoned underpinning to the words of reassurance. ("Why is it alright?" "Because I'm sitting on a cloud, dressed in a bedsheet,

> playing a harp.") What we need is a causal explanation of why the bereaved "heard" what they did. The evolutionist's case is that something similar is very true of ethics. (Ruse, 1986, p. 102)

But of course this story is not analogous to the case of ethics in the least. No one would ask for a "foundation" to the claim "It's alright Mum". What would be asked for is a *causal* explanation, and the two competing explanations are that either there was something supernatural behind the communication or that we can give a satisfactory explanation solely in terms of psychology and projection. On the other hand, when we ask for a *justification* for moral claims we're looking for more than a mere *causal* explanation as to why someone would come to *believe* they're obligated to do something. Since an obligation *demands* something from the obligated, we want to know if there really *is* something that *grounds* that obligation such that the person in question *should* actually be *moved* by it.

Here's a structural disanalogy. Ruse assumes from the outset that there are no ghosts. Operating under this assumption we don't need to ask for a truth-maker to the ghost's claim that everything is alright because there was no such utterance. *Since we know or assume that ghosts don't exist* we therefore look for a *causal* explanation about what the mother claims to have heard in terms of psychology and projection. But now consider the case of ethics. We can't simply *assume* from the beginning that moral facts don't exist and therefore look for causal explanations as to why people make the moral judgments they do. In the case of the ghost, the mother hears something, but since we know ghosts don't exist we do not demand for something to ground the ghost's claim, since there was no such claim. In the case of ethics, an individual may have a belief in the

principle of utility, and we may have an evolutionary genealogy on hand that explains why this individual believes as he does. However, we *cannot* assume that there are no moral facts to ground this individual's belief in the principle of utility. So *it makes sense* to ask for a possible foundation to that individual's belief in the principle of utility given that *there might very well be one*. In the case of the ghost we don't look for a foundation to the ghost's claim because the non-existence of the ghost was already *decided from the outset*. In the case of ethics we have *not* decided on the non-existence of moral facts from the outset, the evolutionary genealogy is *neutral* on their possible existence and therefore *it is* worthwhile to wonder what, if anything, exists out there that could ground the individual's belief in the principle of utility.

Consider Korsgaard's point: suppose you're an individual reasoning as follows: "Yes, I know that my belief that I should maximize utility is the result of this evolutionary process", it would still make perfect sense for this individual to then ask "But *should I* maximize utility?" The question remains *open* for the ethical inquirer. Now consider the grieving mother reasoning in a similar fashion: "Yes, I know that what I heard is a product of psychological projection and didn't really happen…but is what he said really *true*? *Why* did he say it was alright?" This second question seems very odd to ask once you *know* what you heard was an illusion. The question shouldn't even arise. In the first case the normative question remains *open* regardless of the evolutionary genealogy on hand. Ruse claims that

> Ultimately, there is no reasoned justification for ethics in the sense of foundations to which one can appeal in reasoned argument. All one can offer is a causal

> argument to show why we hold ethical beliefs. But once such an argument is offered, we can see that this is all that is needed (Ruse, 1986, p. 102).

The problem here is that whether there *is* some reasoned justification for ethics in the sense of foundations to which one can appeal in reasoned argument *is the very question at hand*. Why should we believe the *collective illusion* hypothesis in the first place? Further, Ruse's claim that once such a causal argument to show why we hold ethical beliefs is offered, we can see that this is all that is needed is simply false. Giving a causal story about why you came to believe that you are obligated to do some action can just as easily serve to *undermine* your belief that you are obligated to do that action in the first place. It is very hard to see what Ruse could mean when he says "that this is all that is needed" in one sentence while in the very next sentence claiming that "the evolutionist's case is that ethics is a collective illusion of the human race, fashioned and maintained by natural selection in order to promote individual reproduction" (Ruse, 1986, p. 102). If ethics is a "collective illusion" then presumably my belief that I'm obligated to X is an illusion, and if we are to understand the word "illusion" in its normal use, this means that my belief is *false*.

Furthermore, Ruse's analogy breaks down when we consider the difference between moral explanations and supernatural explanations. In the Ouija board scenario we have two competing explanations: natural and supernatural. A supernatural explanation would explain the apparent communication of spirits through the board in terms of entities that are not currently explainable by nor postulated by the natural sciences. Such entities could be ghosts, spirits, angels, God, and the like. In order to provide a *supernatural* explanation for the event we'd have to introduce entities that

don't mesh well with the basic methodological assumptions of science, given its commitment to naturalism. Of course this doesn't mean that these explanations are *bad* or a priori *false*; it simply means that if we have to explain a given phenomenon by introducing extra entities into our ontology we should make sure to have thoroughly exhausted all competing explanations that have the virtue of being more ontologically conservative. In the case of Ouija boards we have very good explanations that don't involve introducing extra-ontological entities, one of which involves the *ideomotor effect* (Burgess, Kirsch, Shane, Niederauer, Graham, Bacon, 1998). Moral explanations need not be anything "spooky" or non-natural. Richard Joyce makes a similar point when he writes:

> It is easy enough to make 'moral objectivity' sound spooky, but it need not be. A utilitarian, for example, identifies moral facts with whatever produces the most happiness; and the question of which action available to a person produces the most happiness is an objective matter (it's something we could all be wrong about). Therefore it is not clear that accepting the existence of moral objectivity requires believing in categories of things above and beyond those we already accept. Thus it is not obvious that using a principle of parsimony to eliminate moral objectivity from the world is quite as straightforward as Ruse seems to think (Joyce, 2013, p. 6).

Similarly, David Brink has argued against views that want to "shave off" moral facts using Ockham's razor by appealing to a *constitution relation* between natural facts and moral facts such that if natural facts can explain some phenomenon then moral facts, in virtue of being constituted by those natural facts, can *also* have explanatory power. They

inherit the explanatory power of their constitutive building blocks. In regards to racial injustice in South Africa, he writes:

> The ethical naturalist claims that moral facts are constituted by, and so supervene on, natural facts. It is just such social, political, and legal arrangements that constitute South Africa's racial injustice. It is impossible that these natural facts should remain as they are but that racial injustice should fail to exist (Brink, 1989, p. 191).

And

> The ethical naturalist claims that moral facts not only supervene on, but actually consist in, natural facts. Because the natural facts, which Harman concedes help explain South African instability and protest against apartheid, are what constitute the racial injustice of apartheid, their explanatory power ensures the explanatory power of the injustice that they realize (Brink, 1989, p. 191).

Or so that's the story ethical naturalists want to tell. But what about non-naturalist moral realists? It turns out the difference between the two might be a *semantic* rather than substantive one. Russ Schafer-Landau's moral metaphysics turns out to be virtually identical to that of the moral naturalist's:

> Almost everyone who has thought about the matter has agreed that the moral, too, supervenes on the non-moral. Somehow, in a way yet to be explained, the non-moral features of a situation *fix* its moral status. From which it follows that a thing's moral status cannot change without some correlative change in its non-moral features (Shafer-Landau, 2003, p. 77).

And:

> According to the sort of ethical non-naturalism that I favour, a moral fact supervenes on a particular concatenation of descriptive facts just because these facts realize the moral property in question. Moral facts necessarily co-vary with descriptive ones because moral properties are always realized exclusively by descriptive ones. Just as facts about a pencil's qualities are fixed by facts about its material constitution, or facts about subjective feelings by neurophysiological (and perhaps intentional) ones, moral facts are fixed and constituted by their descriptive constituents. A pencil's length or weight at a time is fixed and constituted by a particular molecular composition, though the same length or weight may, at other times, be realized differently. An instance of pain is exhaustively made up of a set of neural (and perhaps intentional) events, though the same kind of pain (i.e. the same property) can be realized differently at other times and in other people. So, too, the admirability of an action or motive may be realized by different sets of descriptive facts, but on any given occasion, the moral features are fixed by the descriptive ones that compose them at that time. This is the relevant constitution relation; ethical non-naturalists can invoke this relation to give meaning to the idea that descriptive features fix the moral ones (Shafer-Landau, 2003, p. 77).

Ruse's analogy therefore breaks down for two reasons: first, because the kind of justification we want in his example is not the kind of justification we're looking for in ethics. Second, because supernatural explanations introduce entities that violate the principle of parsimony whereas ethical explanations in terms of moral facts do not introduce any special ontological "stuff" (they might introduce sui generis *properties* of

matter, moral ones, but they don't introduce new sui generis *substances*, which is what supernatural explanations tend to invoke). This means that Ruse's main argument for the thesis that morality is a collective illusion fails.

Ruse's second thesis, that even if morality *is* an illusion it can still be vindicated, also fails. Ruse writes:

> What is really important to the evolutionist's case is the claim that ethics is illusory inasmuch as it persuades us that it has an objective reference (Ruse, 1986, p. 103).

He continues:

> "Love little children" is not like "My favorite vegetable is spinach." The latter is just a matter of subjective preference. If you do not like spinach, then nothing ensues. But we do not take the former (moral) claim to be just a matter of preference. It is regarded as objectively binding upon us—whether we take the ultimate source of this objectivity to be God's will, or (like G.E. Moore) apprehension of nonnatural properties, or whatever (Ruse, 1986, p. 103).

The source of the "illusion" for Ruse lies in the fact that we regard certain claims as objectively binding upon us:

> The evolutionist's claim, consequently, is that morality is subjective—it is all a question of human feelings or sentiments—but he/she admits that we "objectify" morality, to use an ugly but descriptive term. We think morality has objective reference even though it does not. Because of this, a causal analysis of the type offered by the evolutionist is appropriate and adequate, whereas justification of

> moral claims in terms of reasoned foundations is neither needed nor appropriate (Ruse, 1986, p. 103).

The second part of the above is yet another non-sequitur. If you think you've uncovered the mechanism by which human beings reason morally by appealing to our evolutionary past and claim to have shown that we "objectify" morality even though (somehow) you know that we really *aren't* obligated to do anything, then it simply doesn't follow that "a causal analysis of the type offered by the evolutionist is appropriate and adequate". Adequate and appropriate for what exactly? If you think you are obligated to X, but you have an empirically confirmed story that shows *why* you think you're obligated to X, even though you really *aren't* obligated, then it simply follows that you have *no reason* to *care* about morality. What you have is a *defeater* for your belief that you're obligated to X. Far from being an "appropriate and adequate" story what we have is an *undermining* story that *liberates* individuals from their beliefs about the sorts of things they think they're obligated to do. Ruse finally says:

> Furthermore, completing the case, the evolutionist points out that there are good (biological) reasons why it is part of our nature to objectify morality. If we did not regard it as binding, we would ignore it. It is precisely because we think that morality is more than mere subjective desires, that we are led to obey it (Ruse, 1986, p. 103).

Quite so. We have a *descriptive* account here as to why we tend to think about morality in terms of its being objectively binding on all of us. If we didn't, we would simply disregard it. Once we are armed with this story though, *if* Ruse could *show* that there really are no objectively binding obligations on all of us (something he has *not*

shown) then we would see morality for what it really is, an illusion that we can disregard. Ruse thinks that he *has* shown morality to be a collective illusion, but instead of accepting what that would entail, our license to disregard what we believe to be our moral obligations, he thinks we have good reason instead to *continue* behaving as if morality *did* have an objective foundation. To the contrary, it is clear that an individual undergoing first person rational deliberation, when armed with this evolutionary story, might come to see that he/she really has no *reason* to do what morality tells him/her to do. There might be *pragmatic* reasons for behaving as if morality weren't a fiction, but there aren't any *substantive* reasons to behave morally, at least not any reasons to refrain from behaving immorally when one thinks one can get away with it.

There is a second argument Ruse gives for the conclusion that ethics is without foundation. He writes:

> Imagine two worlds, identical except that one has an objective ethics (whatever that might mean) and one does not. Perhaps, in one world God wants us to look after the sick, and in the other He could not care less what we do. The evolutionist argues that, in both situations, we would have evolved in such a way as to think that, morally, we ought to care for the sick. To suppose otherwise, to suppose that only the world of objective ethics has us caring about the sick, is to suppose that there are extrascientific forces at work, directing and guiding the course of evolution. And this is a supposition that is anathema to the modern biologist. (Ruse, 1986, p. 108)

Of course this objection fails for reasons already covered. Moral realists of the natural and non-natural variety both hold that descriptive facts *fix* moral facts.[4] There just *aren't* any worlds that are identical in their descriptive facts but which vary in their moral facts. Moral realists hold that all worlds identical in descriptive respects will be identical in their moral facts. Theists also hold that God's commandments necessarily flow from his fixed nature, and God would not prescribe different behaviors in worlds that are identical in their descriptive circumstances.[5] Ruse invokes the existence of parallel worlds identical in their physical facts but differing in their moral facts so as to show that in both worlds we would hold similar or identical moral judgments. He thinks that the only explanation as to why we would hold different moral judgments in physically identical worlds is the presence of extra-scientific forces at play. Ruse claims that it is wrong to think that only a world in which moral facts exist would produce creatures with moral sentiments like ours, and he thinks this is false because the evolutionary story he has presented us with gives us a causal story that shows how such sentiments can come about without appealing to moral facts. Given that both stories explain the phenomena in question, and one of them invokes the existence of extra-scientific forces whereas the second does not, considerations of simplicity should force us to accept the latter, all else being equal. Later I will argue a similar point but differ in my conclusion. I will argue that the fact that we would believe some moral claim is true *even if it were false* gives us reason to call into question the justification of this belief. The difference is that Ruse thinks we would believe certain moral propositions even if there were no moral facts to

---

[4] This *would,* however, work against a realist who denies the supervenience relation.
[5] With the possible exception of certain extreme divine command theorists, such as René Descartes.

ground them, and further we can give an evolutionary account as to why we believe such propositions that does not invoke their existence, therefore we should abandon our belief in moral facts on the grounds of parsimony. Ruse's conclusion is *ontological.* My own, and Richard Joyce's conclusion will be *epistemic*.

There is, however, a point made by Ruse that is promising, though he fails to develop it at length, and the conclusion he draws from it is wrong. He writes:

> Suppose, instead of evolving from savannah-living primates (which we did), we had come from cave dwellers. Our nature and our morality might have been very different. Or, take the termites (to go to an extreme example from a human perspective). They have to eat each other's feces, because they lose certain parasites, vital for digestion, when they molt. Had humans come along a similar trail, our highest ethical imperatives would have been very strange indeed. What this all means is that, whatever objective morality may truly dictate, we might have evolved in such a way as to miss completely its real essence. We might have developed so that we think we should hate our neighbors, when really we should love them. Worse than this even, perhaps we really should be hating our neighbors, even though we think we should love them! Clearly, this possibility reduces objectivity in ethics to a mass of paradox (Ruse, 1986, p. 108).

The interesting point about the above is that it brings out a question about the relationship between our moral beliefs and whether we have any good reason to trust that they accurately track moral facts. If our moral beliefs are products of contingent circumstances, and given certain other contingent events these beliefs could have changed substantially from what we currently believe, then we need to have a good story

on hand that can allow us to be justified in believing that our moral judgments accurately track and properly correspond to the relevant moral facts.

Unfortunately Ruse draws the wrong conclusion from this, that objectivity in ethics is reduced to a mass of paradox. It is hard to know what he means by this, but I think one plausible interpretation is the simple conclusion that there isn't or can't be any objectivity in ethics. Of course this just doesn't *follow* from the epistemic concerns Ruse raises, at least not in an easily discernible straight line. Ruse needs to provide more premises to really bring out the full implications of this epistemic worry. Unfortunately, he does not do this. We will see in the next chapter how the two most promising evolutionary debunking arguments against morality take up these epistemic worries and develop them more carefully than Ruse, drawing the appropriate epistemic conclusions from them.

In this chapter I have argued that Michael Ruse's evolutionary debunking strategy fails on both counts. His evolutionary considerations do not entail what he claims they do. We have also seen that his vindicating strategy that aims to show how these considerations can help *strengthen* our moral convictions are equally wide of the mark. Ruse raises an interesting epistemic worry for moral realists but fails to draw the right conclusion from it and does not develop it in in enough detail for us to be able to draw the right inferences from it. For these reasons Michael Ruse's evolutionary debunking argument fails. In this and the previous chapter we have considered two evolutionary debunking strategies and have attempted to show why both fail. Plantinga thinks evolutionary considerations about the development of our cognitive faculties call into question *all* of our beliefs. Ruse on the other argues that these considerations only call

into question the *moral* domain. In the next chapter I will examine two more evolutionary debunking arguments and will defend them from objections. I will argue that these evolutionary arguments pose a real threat to moral realists.

Chapter 3: Sharon Street And Richard Joyce's Evolutionary Debunking Arguments

Sharon Street has argued that given the fact that evolutionary forces have played an integral role in the development and shaping of our human evaluative attitudes, an explanation is needed for the relationship between these evolutionary forces on the one hand and the independent evaluative truths that non-natural moral realism posits on the other. According to Street, moral realism cannot give a satisfactory account of this relation. In this chapter I will reconstruct Street's argument and critically assess its strength in light of the many objections that have been raised against it in the literature. In the process I will also explain Richard Joyce's similar evolutionary debunking argument and will show why objections to *that* argument fare no better. I will conclude that evolutionary considerations pose a serious challenge to moral realism that has yet to be answered in a satisfactory manner. If the only two options are moral realism or moral agnosticism, then we ought to be moral agnostics. There are, however, more options and I will briefly explore one such option.

A Darwinian Dilemma For Realist Theories of Value

Sharon Street's argument is posed in the form of a dilemma. Granting the role that evolution has played in the development of our evaluative attitudes, an explanation is needed as to the relationship between the unguided development of these evaluative attitudes and the independent moral truths that moral realism posits. According to Street, a moral realist has only two avenues by which to approach this issue: a) deny that there is a relationship between the shaping of our moral attitudes and the truths that moral realism posits on the one hand or b) affirm that there is a relationship between evolutionary

influences and the independent evaluative truths of moral realism, where this relationship is understood as a *tracking* relationship. Both horns of the dilemma lead to unacceptable results we have no good reason to believe our moral judgments accurately track the moral truths that are said to exist independently of our judgments about them. The first horn thus leads to moral skepticism. The second horn is unacceptable, Street argues, on scientific grounds. Either way, moral realism faces an undercutting defeater because of evolutionary considerations as to how our evaluative judgments have been shaped in light of evolution.

The target of Street's argument are realist theories of value entailing "that there are at least some evaluative facts or truths that hold independently of all our evaluative attitudes" (Street, 2006, p. 110). These evaluative facts or truths are of the form "X is a normative reason to Y, that one should or ought to X, that X is good, valuable, or worthwhile, that X is morally right or wrong, and so on" (Street, 2006, p. 110). Evaluative attitudes are states such as

> desires, attitudes of approval and disapproval, unreflective evaluative tendencies such as the tendency to experience X as counting in favor of or demanding Y, and consciously or unconsciously held evaluative judgments, such as judgments about what is a reason for what, about what one should or ought to do, about what is good, valuable, or worthwhile, about what is morally right or wrong, and so on (Street, 2006, p. 110).

Street's target are moral realists who are united in the view that "there are evaluative facts or truths that hold independently of all our evaluative attitudes" (Street, 2006, p. 111) and who also hold non-naturalism about these evaluative facts, a view in which

"evaluative facts or truths are not reducible to any kind of natural fact, and are not the kinds of things that play a role in causal explanations; instead, they are irreducibly normative facts or truths" (Street, 2006, p. 111).

So much for her target. Before laying out the dilemma, Street notes that her argument comes with a caveat similar to Alan Gibbard's in his *Wise Choices, Apt Feelings*. "*If* the evolutionary facts are roughly as I speculate, here is what might be said philosophically" (Street, 2006, p. 113). Street's dilemma arises only *if* some evolutionary account of the origin of our evaluative attitudes roughly like the one she outlines is true. Though the precise details may be fuzzy and unknown to us now, or perhaps will always remain so, as long as the general *outline* of an evolutionary account is certain enough, then the philosophical implications of this outline will be worth working out because, as Street believes, there is good reason to believe an account roughly like hers *is* true.

With the target and caveat out of the way we can now focus on Street's argument. The first premise of her argument is the following: "The forces of natural selection have had a tremendous influence on the content of human evaluative judgments" (Street, 2006, p. 114). This is of course not to deny that many other factors other than natural selection have helped shape our evaluative attitudes; genetic drift, social, cultural, and historical forces have played similar and equally important roles. The claim is simply that "our system of evaluative judgments is thoroughly saturated with evolutionary influence" (Street, 2006, p. 114).

As support for this Street asks us to consider first the great costs and benefits, measured in the Darwinian currency of reproductive success, of accepting some value judgments over others. It is clear "how fatal to reproductive success it would be to judge

that the fact that something would endanger one's survival is a reason to do it, or that the fact that someone is kin is a reason to harm that individual" (Street, 2006, p. 114). Creatures who accepted such value judgments would not last very long as they would "run off cliffs, seek out. . .predators, and assail (their) offspring, resulting in (their) speedy elimination and (their) evaluative tendencies from the world" (Street, 2006, p. 114). It is therefore clear how beneficial it would be to judge "that the fact that something would promote one's survival is a reason in favor of it, or that the fact that something would assist one's offspring is a reason to do it" (Street, 2006, p. 114). It would seem then that evaluative tendencies are not or could not be completely blind to the selective pressures of natural selection given how different tendencies can have drastically different outcomes in the likelihood of a creature's survival. We have, then, some reason to believe that there would have been

> overwhelming pressure in the direction of making those evaluative judgments which tended to promote reproductive success (such as the judgment that one's life is valuable), and against making those evaluative judgments which tended to decrease reproductive success (such as the judgment that one should attack one's offspring) (Street, 2006, p. 115).

This is a very rough hypothesis, but Street thinks evidence for this can be found in the patterns of evaluative judgments that we find across time and cultures. That the following evaluative judgments are commonplace across cultures provides evidence for the above hypothesis:

- The fact that something would promote one's survival is a reason in favor of it (Street, 2006, p. 115).

- The fact that something would promote the interests of a family member is a reason to do it (115).
- We have greater obligations to help our own children than we do to help complete strangers (115).
- The fact that someone has treated one well is a reason to treat that person well in return (p115).
- The fact that someone is altruistic is a reason to admire, praise, and reward him or her (115).
- The fact that someone has done one deliberate harm is a reason to shun that person or seek his or her punishment (115).

According to Street, what explains the wide acceptance of the above moral judgments is that "*these* sorts of judgments about reasons tended to promote survival and reproduction much more effectively than the alternative judgments" (Street, 2006, p. 116). For example, the rough explanation for the first bullet point is obvious:

> creatures who possessed this general evaluative tendency tended to do more to promote their survival than those who, say, had a tendency to view the fact that something would promote their survival as counting *against* it, and so the former tended to survive and reproduce in greater numbers (Street, 2006, p. 116).

Explanations for the other bullet points may be found by drawing upon theories of kin selection and theories of reciprocal altruism, though Street herself does not go into much detail (Street, 2006, p. 116).

In contrast, if our evaluative judgments had *not* been shaped by evolutionary forces then we would expect more widespread acceptance of evaluative judgments like the following:

- The fact that something would promote one's survival is a reason against it (Street, 2006, p. 116).
- The fact that something would promote the interests of a family member is a reason not to do it (116).
- We have greater obligations to help complete strangers than we do to help our own children (116).
- The fact that someone has treated one well is a reason to do that individual harm in return (116).
- The fact that someone is altruistic is a reason to dislike, condemn, and punish him or her (116).
- The fact that someone has done one deliberate harm is a reason to seek out that person's company and reward him or her (116).

Since we do *not* find widespread acceptance of such moral judgments across cultures and time, which is what we'd expect if Darwinian forces had *not* greatly influenced our evaluative attitudes, we have further evidence for Street's premise, unless there is a plausible alternative explanation for the presence of the evaluative attitudes we actually have. We find deeply and widely held judgments "with exactly the sort of content one would expect if the content of our evaluative judgments had been heavily influenced by selective pressures" (Street, 2006, p. 117). And so "the observed patterns in the actual

content of human evaluative judgments provide evidence in favor of the view that natural selection has had a tremendous influence on that content" (Street, 2006, p. 117).

Further evidence is provided by the strong continuity in our evaluative judgments and more basic evaluative judgments found in other creatures across the animal kingdom, particularly in our closest relatives like chimpanzees. Chimpanzees for example seem to experience at more basic level "actions that would promote survival or help their offspring as in some way 'called for'" (Street, 2006, p. 117). Further, at some basic level chimpanzees

> seem to experience the fact that another chimpanzee has helped them, whether by sharing food, grooming them, or supporting their position within the group hierarchy, as 'counting in favor of' assisting that other individual in similar ways (Street, 2006, p. 117).

We have to remember at this point that none of this counts as a fully descriptive theory about the origins of our human moral capacities. That job is for evolutionary psychologists to sort out. All that Street needs is for the general *outline* of this story to be correct. In line with this, Street has helped herself to what may be some of the least controversial evidence at hand.

The influence of Darwinian selective pressures on the content of our evaluative judgments is most plausibly taken to be indirect. For one, it is implausible to think that the reason some creature affirms any given moral judgment over another is because of some genetic hardwiring that is itself "visible" to natural selection. For example, it is implausible to think that an organism's affirming of "one ought to help those who help you" is due to some genetic hardwiring that was selected for by natural selection. Rather,

Street thinks there different “levels” of evaluative judgments. More refined judgments (like the above) are not likely to be what are selected for by natural selection. Rather, *basic* evaluative tendencies are the proper objects of selection. Rather than propositions that we can reflectively endorse (like the above), these basic evaluative tendencies are best understood as unreflective, non-linguistic, motivational tendencies to experience something as “called for” or “demanded” in itself, or to experience one thing as “calling for” or “counting in favor of” something else (Street, 2006, p. 119).

Examples of these may be a bird that experiences some kind of motivational “pull” in the direction of feeding offspring or “a more sophisticated version might be possessed by a chimpanzee who has a motivational and perhaps emotional or proto-emotional experience of certain behaviors as ‘called for’ by certain circumstances” (Street, 2006, p. 11). Selection on our “upper level” moral judgments can be understood as *indirect.* Natural selection acts on our basic evaluative judgments, which are themselves the basic building blocks for our “upper level” more sophisticated judgments. This in turn implies that “had the general content of our basic evaluative tendencies been very different, then the general content of our full-fledged evaluative judgments would also have been very different, and in loosely corresponding ways” (Street, 2006, p. 120).

All that needs to be granted for Street’s dilemma to get off the ground is this: “The content of human evaluative judgments has been tremendously influenced by the forces of natural selection, such that our system of evaluative judgments is saturated with evolutionary influence. The truth of some account very roughly along these lines is all that is required for the Darwinian Dilemma to get off the ground” (Street, 2006, p. 121).

The Dilemma

According to Street:

> The basic problem for realism is that it needs to take a position on what relation there is, if any, between the selective forces that have influenced the content of our evaluative judgments, on the one hand, and the independent evaluative truths that realism posits, on the other. Realists have two options: they may either assert or deny a relation (Street, 2006, p. 121).

Street takes on each horn in reverse order. On the second horn of the dilemma, in which one denies a relation, "the forces of natural selection must be viewed as a purely distorting influence on our evaluative judgments, having pushed us in evaluative directions that have nothing whatsoever to do with the evaluative truth" (Street, 2006, p. 121). As an analogy

> 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 every now and then, the wind and tides might happen to deposit someone's boat on the shores of Bermuda. Similarly, every now and then, Darwinian pressures might have happened to push us toward accepting an evaluative judgment that accords with one of the realist's independent evaluative truths. But this would be purely a matter of chance, since by hypothesis there is

> no relation between the forces at work and the "destination" in question, namely evaluative truth (Street, 2006, p. 122).

This conclusion is unpalatable to the moral realist because it leads to the conclusion that our evaluative judgments are unreliable because "they are revealed to be contaminated with illegitimate influence. We should have been evolving towards affirming the independent evaluative truths posited by the realist, but instead it turns out that we have been evolving towards affirming whatever evaluative content tends to promote reproductive success" (Street, 2006, p. 122). The realist may here object that a full explanation of why we hold the evaluative judgments we do is not complete until we take into consideration non-evolutionary forces, such as sociological and reflective influences that allow us to reach judgments that are not simply inherited. As Street puts the objection "we are not unthinking beings who simply endorse whatever evaluative tendencies were implanted in us by evolutionary forces" (Street, 2006, p. 123). And so "just as a compass and a little steering can correct for the influence of the wind and tides on the course of one's boat, so rational reflection can correct for the influence of selective pressures on our values" (Street, 2006, p. 123).

This objection is correct in that it gives us a fuller picture of the types of influences on our moral thinking. Where it goes wrong according to Street is the fact that what rational reflection about evaluative matters involves "inescapably, is assessing some evaluative judgments in terms of others" (Street, 2006, p. 124). It must "proceed from some evaluative standpoint; it must work from some evaluative premises; it must treat some evaluative judgments as fixed, if only for the time being, as the assessment of other evaluative judgments is undertaken" (Street, 2006, p. 124). Further:

> In rational reflection, one does not stand completely apart from one's starting fund of evaluative judgments: rather, one uses them, reasons in terms of them, holds some of them up for examination in light of others. The widespread consensus that the method of reflective equilibrium, broadly understood, is our sole means of proceeding in ethics is an acknowledgment of this fact: ultimately, we can test our evaluative judgments only by testing their consistency with our evaluative judgments, combined of course with judgments about the (non-evaluative) facts. Thus, if the fund of evaluative judgments with which human reflection began was thoroughly contaminated with illegitimate influence and the objector has offered no reason to doubt *this* part of the argument—then the tools of rational reflection were equally contaminated, for the latter are always just a subset of the former. It follows that all our reflection over the ages has really just been a process of assessing evaluative judgments that are mostly off the mark in terms of others that are mostly off the mark (Street, 2006, p. 124).

Reflective equilibrium then cannot help us escape the conclusion that our evaluative judgments are contaminated and off the mark because the process of this kind of moral reflection unavoidably *uses* the judgments in question to assess others. It must take some judgments as givens and from there proceed to test for *coherence* and helps us assess what types of judgments imply others, but it cannot help us assess what types of judgments are *contaminated* with illegitimate influence. The first horn of the dilemma then is unpalatable to the moral realist.

The Second Horn

The second option available to the moral realist is to assert that there *is* a relation between the forces of natural selection and the evaluative truths we've inherited. The realist then is burdened with the task of specifying the nature of this relationship. It is implausible to think this relationship is one of sheer chance, that natural selection *just happened* to supply us with correct moral beliefs. So then what is this relationship? The most attractive option is *the tracking relation*. This can be understood in the following way: "it is advantageous to recognize evaluative truths; surely it promotes one's survival (and that of one's offspring) to be able to grasp what one has reason to do, believe, and feel" (Street, 2006, p. 125). So, according to this view

> Our ability to recognize evaluative truths, like the cheetah's speed and the giraffe's long neck, conferred upon us certain advantages that helped us to flourish and reproduce. Thus, the forces of natural selection that influenced the shape of so many of our evaluative judgments need not and should not be viewed as distorting or illegitimate at all. For the evaluative judgments that it proved most selectively advantageous to make are, in general, precisely those evaluative judgments which are true (Street, 2006, 125).

According to Street, to take this particular route is to accept a *scientific hypothesis.* Accepting the above view entails that the realist accepts a specific hypothesis about why we hold the type of judgments we do and why natural selection took this particular route. This hypothesis "says that the presence of these judgments is explained by the fact that these judgments are true, and that the capacity to discern such truths proved advantageous for the purposes of survival and reproduction" (Street, 2006, p. 126). So when we ask

why we take our survival and that of our offspring to be valuable "the tracking account answers that these judgments are true, and that it promoted reproductive success to be able to grasp such truths" (Street, 2006, p. 126).

Taking this horn of the dilemma leads to problems for the realist. As a scientific hypothesis, this account "renders itself subject to all the usual standards of scientific evaluation, putting itself in direct competition with all the other scientific hypotheses as to why human beings tend to make some evaluative judgments rather than others" (Street, 2006, p. 126). And as a scientific hypothesis, according to Street, it fares poorly when compared to competing hypotheses. There are competing hypotheses that better explain why we hold the judgments we do, ones that do not involve assuming the truth of these judgments as an explanation for why we hold them. Street has in mind a competing account which she calls the *adaptive link account*. According to this account

> tendencies to make certain kinds of evaluative judgments rather than others contributed to our ancestors' reproductive success not because they constituted perceptions of independent evaluative truths, but rather because they forged adaptive links between our ancestors' circumstances and their responses to those circumstances, getting them to act, feel, and believe in ways that turned out to be reproductively advantageous (Street, 2006, p. 127).

This account suggests that there are selective advantages for certain behaviors that have nothing to do with the truth or falsity of the beliefs in question. Street gives the example of helping someone who has helped one before:

> Just as we may see a reflex mechanism as effecting a pairing between the circumstance of a hot surface and the response of withdrawing one's hand, so we

> may view this evaluative judgment as effecting a pairing between the circumstance of one's being helped and the response of helping in return. Both of these pairings of circumstance and response, at least if the evolutionary theory of reciprocal altruism is correct about the latter case, are ones that tended to promote the reproductive success of ancestors who possessed them (Street, 2006, p. 127).

We can better appreciate the differences between the two competing account by the following:

> Consider, for instance, the judgment that the fact that something would promote one's survival is a reason to do it, the judgment that the fact that someone is kin is a reason to accord him or her special treatment, and the judgment that the fact that someone has harmed one is a reason to shun that person or retaliate. Both the adaptive link account and the tracking account explain the widespread human tendencies to make such judgments by saying that making them somehow contributed to reproductive success in the environment of our ancestors. According to the tracking account, however, making such evaluative judgments contributed to reproductive success because they are true, and it proved advantageous to grasp evaluative truths. According to the adaptive link account, on the other hand, making such judgments contributed to reproductive success not because they were true or false, but rather because they got our ancestors to respond to their circumstances with behavior that itself promoted reproductive success in fairly obvious ways: as a general matter, it clearly tends to promote reproductive success to do what would promote one's

> survival, or to accord one's kin special treatment, or to shun those who would harm one (Street, 2006, p. 128).

The second horn of Street's dilemma then places the realist in the position of having to defend the tracking account by means of offering it as an inference to the best explanation. This approach, however, simply will not work, as for every explanation offered for any given evaluative attitude X by the realist, Street will offer an equally compelling account that appeals only to the survival adaptability of said evaluative attitude, an account that Street believes is superior to the tracking relation with respect to parsimony, clarity, and in virtue of shedding much more light on the explanandum.

The adaptive link account is more parsimonious according to Street because the tracking account posits entities to explain a given phenomenon whereas the adaptive link account does not. The tracking account posits non-natural moral facts to explain why individuals make certain evaluative judgments. On the other hand, the adaptive link account explains our propensity to make certain types of judgments "by pointing out how they got creatures who made them to act in ways that tended to promote reproductive success" (Street, 2006, p. 129). So the adaptive link account explains our propensity to make evaluative judgments "without any need to posit a role for evaluative truth" (Street, 2006, p. 129).

In terms of clarity, the adaptive link account wins out because the tracking account is rather obscure. The tracking account is a claim that "making certain evaluative judgments rather than others promoted reproductive success *because these judgments were true*" (Street, 2006, p. 129). What could this mean? On the face of it, the claim seems straightforward. Creatures that grasp truths behave in ways that lend themselves to

their continued survival. It is easy to see how *some* truths lend themselves to adaptive behavior such as what kinds of predators, if any, are nearby, what kinds of food are edible, etc. However, the tracking account posits that grasping *moral* truths also convey such a survival advantage. This claim isn't immediately obvious.

Irreducibly normative truths as postulated by Nagel, Dworkin, Scanlon, and Shafer-Landau, says Street, are not the kinds of things that an organism *runs into* in its environment. So the question then becomes *in virtue of what* would an organism's grasp of such truths convey a survival advantage? "To say that these truths could kill you or maim you, like a predator or a fire, would be one kind of answer, since it makes clear how recognizing them could be advantageous" (Street, 2006, p. 130). But of course this is not an answer available to the realist. It's difficult to see what kind of answer a value realist could give here. Street's adaptive link account doesn't face such problems because the answer is so simple: "ancestors who judged that they should care for their offspring met with greater reproductive success simply because *they tended to care for their offspring* and so left more of them" (Street, 2006, p. 131). In terms of clarity then, the adaptive link account scores better than the tracking relation of the realist.

Third, the adaptive link account does a better job at shedding light on what is in need of explanation. If we're asking questions such as "Why do we tend to judge that we have special obligations to care for our children, rather than strangers or distant relatives? Why do we tend to view the killing of other human beings as a much more serious matter than the killing of plants or other animals?" (Street, 2006, p. 132) The answer is quite straightforward under the adaptive link account: Making judgments of these kinds led to reproductive success over counterparts. As Street writes, "it is quite clear why creatures

who judged their survival to be valuable would do much better than those who did not, and so on" (Street, 2006, p. 132). Comparing the adaptive link account to the tracking account, it becomes obvious which answer actually *explains* the phenomenon in question. The tracking account tries

> to answer these same questions by saying that these judgments are *true:* that survival *is* valuable, that we *do* have special obligations to care for our children, that the killing of human beings *is* more serious than the killing of plants or other animals (Street, 2006, p. 132).

According to Street, the tracking account cannot explain why so many of the moral truths it posits coincide perfectly with precisely the kinds of judgments we'd expect to see if our judgments had been selected for merely on the basis of their forging adaptive links between circumstances and response. The tracking account also has a difficult time in explaining why we hold certain pre-critical moral judgments, such as the fact that someone is a member of an out-group being a reason to value him/her less, that upon reflection turn out to be false. The realist cannot say that we developed this belief because it was *true*. The adaptive link account can explain this quite simply. The realist however will be forced to admit that many of our pre-critical moral judgments are false. Here I believe Street undersells the strength of her argument. If they are false then they must have been selected for something other than truth. If this is so then the realist must grant that many of our evaluative judgments are thoroughly contaminated with illegitimate influence; illegitimate here being construed in terms of selective pressure leading to judgments that were selected for on some basis *other than* truth. This leads to a strange situation in which the realist must accept that some of our evaluative judgments were

selected for because they were true, and some others were selected for merely in virtue of their tendency to contribute to adaptive behavior. Why should this be, and do we have any means by which to ascertain which judgments belong to which group?

To sum, Sharon Street offers the adaptive link account as a hypothesis that explains the

> disparate mishmash, ranging across all kinds of unrelated spheres and reflecting all kinds of unrelated values -some self-interested, others family-related, still others concerning how we should treat non-relatives and other forms of life, and so on. The power of the adaptive link account is that it exposes much of this seeming unrelatedness as an illusion; it illuminates a striking, previously hidden unity behind many of our most basic evaluative judgments, namely that they forge links between circumstance and response that would have been likely to promote reproductive success in the environments of our ancestors (Street, 2006, p. 134).

It is a hypothesis that beats its competitor in several important respects. Because of this, the second horn of the dilemma poses a serious problem to those who would affirm moral realism and also affirm our evolutionary history such that there is some kind of relationship between the forces of natural selection and the moral truths we posit. How then can a realist respond to this Dilemma?

Abraham Graber

Abraham Graber argues that Street's Dilemma can not only be resisted but can also be deflected such that it poses a challenge to *anti-realism.* Graber objects to Street on several fronts. His first attack on Street is the claim that Street's first horn is question begging. He grants that it would indeed be absurd to expect to land on Bermuda without

a compass and a map. Your expectations to land there, *before taking off*, would be irrational. However, this says nothing about the kind of evidence you can have as to whether or not you've arrived there *once you've set foot on land.* For example, even though the probability of landing on Bermuda is rather low, if you find yourself on shore and "written in large letters on a thirty-foot billboard are the words 'Welcome to Bermuda'" (Graber, 2012, p. 591) then you'd have very good reason to believe you've actually arrived. He therefore thinks that "the first horn of Street's argument is only effective if one takes oneself not to have any evidence regarding the point of arrival of moral psychology from the seas of evolution" (Graber, 2012, p. 591). Graber argues that since realists presumably *do* hold that we have some good reasons to believe we possess justified beliefs about ontologically robust moral properties, it would be question begging to assert otherwise. He writes:

> Any reason that the realist has for thinking that we track evaluative truths is a reason to think that we evolved in such a way as to be able to track evaluative truths, just as any reason we have for believing we've arrived in Bermuda is a reason to think that the tides and the winds have brought us to Bermuda (Graber, 2012, p. 591).

This objection to Street's first horn fails. While we would all agree that a billboard with the words "Welcome to Bermuda!" would give one good reason to suppose they are indeed in Bermuda, it is far more difficult to see what would count as an analogous case when looking for evidence that our moral beliefs have accurately tracked the truth. There is not very much that we can point to that all parties involved, or even most, would agree gives us very compelling reasons to believe our moral judgments are

true. There are no "moral billboards" to which can compare our moral judgments and see how well they match. Perhaps a divine command theorist would have such a resource available in virtue of God's possessing perfect knowledge, including moral knowledge, which he could relay to us via his commandments, but since Street's target is non-natural, non-reductive moral realism, such a moral billboard is not on hand. Graber himself does not provide us with any of the alleged reasons realists give for supposing our moral judgments accurately track moral facts, nor does he cite any in the literature. It is therefore very difficult to take seriously his "Welcome to Bermuda!" analogy since he has not pointed to any such epistemic equivalent in the moral realm.

Graber is correct in pointing out that Street's first horn does not rule out the possibility of attaining independent evidence that we've arrived in Bermuda. The problem, however, is *showing us* what that evidence is when it comes to the moral realm. For this reason I think his attack on the first horn fails. It fails because he *suggests* that we can have independent reasons for supposing we've arrived at our desired destination, and he correctly points out that nothing in Street's argument prevents this from being the case. But of course a mere *suggestion* that such independent evidence could exist does not make for a compelling reason to believe that we *do* have independent reason to believe that our evaluative judgments are on track. Moreover, is very difficult to see how we could know *that* without reference to the very same judgments in question in the first place. Reflective equilibrium will not work, for reasons already spelled out by Street; so, given that we can't find an *internal* way of sorting out truth tracking judgments from those that aren't, and given that we do not have access to some *external* source, it is hard to see just what this evidence is supposed to be. Street's argument is not, therefore,

question begging. It merely declined to take seriously the possibility of our finding independent evidence that we've arrived at our desired destination. But since there is no such independent evidence, at least none provided by Graber, then the first horn remains a live possibility.

For the sake of argument however I'll assume that we *do* have good reason to believe we have at least *some* justified moral judgments. Unless the realist wants to say that our possessing true moral beliefs is a result of cosmic coincidence, the fact that we possess true moral beliefs would count as evidence that there *was* some kind of truth tracking relation between the forces of natural selection and of our moral judgments. This would then put Graber in a position to confront Street's *second* horn.

Graber's move here is to argue that Street is wrong to present her argument as a dilemma. On the first horn, it turns out that if our evaluative judgments turn out to be true, then they are true merely by chance. Graber then claims that Street's second horn turns out to be the same as the first. He writes

> Suppose that Street is correct and the best explanation of our evaluative judgments makes no reference to evaluative truth. It follows that, if our evaluative judgments track evaluative truth, it is only by chance. But this is exactly the same problem that was suppose to arise on Street's first horn (Graber, 2012, p. 594).

Since according to Graber the first horn begs the question, and the second horn turns out to be the same as the first, then the latter also begs the question. Since I take myself to have shown that the first horn *does not* beg the question, then this move loses its bite and as such Graber's objections to Street's argument fail. Graber claims that "Street has not, in fact, offered a dilemma. Instead, we ought to view Street's 'dilemma'

as a long abductive argument" (Graber, 2012, p. 594). On the contrary, it *is* a dilemma for realism. The dilemma is quite straightforward, affirming or denying a relation between the forces of natural selection and our moral judgments leads to the unpalatable result for the realist that *if* we possess true moral judgments then we do so by cosmic chance, and of course the realist can't affirm this. If this does not count as a dilemma then what does?

Knut Olav Skarsaune

Knut Olav Skarsaune's proposed solution to Street's dilemma gains traction from its intuitive plausibility. He asks us to consider the following evaluative thesis: P: Pleasure is usually good and pain is usually bad. *If* P is true then, Skarsaune argues, there is a relation between evolutionary pressures and evaluative facts. If P is false then there is no relation. He then reasons that if P is true then "realists *can* give an account of how. . .there is a relation between Darwinian forces and the evaluative facts" (Skarsaune, 2011, p. 233). If P is *false* then we have good reason to believe that "many or most of our evaluative beliefs are false" (Skarsaune, 2011, p. 233). So, if P is true then the first horn is tenable. If P is false then the second horn is tenable.

Skarsaune's first argument aims to establish the following conditional: "if pleasure is usually good and pain usually bad, there does indeed exist a relation between evolutionary pressures and the evaluative facts, a relation which is *truth-conducive* in the sense that it would tend to bias our evaluative beliefs toward the truth" (2011, 2011, p. 233). The reason for this is quite straightforward,

> Evolution has caused us to value reproductively beneficial things by making us such that we take pleasure in these things, and caused us to disvalue

> reproductively harmful things by making us such that these things cause us pain (Skarsaune, 2011, p. 233).

From the above premise it should be easy to see how to proceed given the truth of P. If pleasure is usually good then

> to the extent that evolution has influenced our evaluative beliefs through the mechanism just described, that influence has been truth-conducive. For if pleasure is usually good, then the activities and states of affairs evolution has caused us to value through this mechanism tend to *be* good—because they are pleasurable. Hence, if P is true, there is a relation between reproductive enhancement and goodness after all (Skarsaune, 2011, p. 234).

Skarsaune calls this the "pre-established harmony account" according to which "if pleasure is (usually) good, then it was ordained ahead of time, as it were, that almost whatever evolution should happen to make us value through this mechanism, it would thereby also imbue with value" (Skarsaune, 2011, p. 235). Straightforwardly, if pain is usually bad, and the mechanisms of evolution make it such that creatures who experience pain will disvalue it and consider it as bad, then the mechanisms of evolution are truth-conducive since these mechanisms produce the right kinds of reactions and beliefs about pain.

What if P is *not* true? Then, says Skarsaune, we need not worry about the Darwinian dilemma at all since we're already in big trouble prior to any Darwinian considerations. If pain is not usually bad then many, if not most, of our ethical beliefs turn out to be misguided. Beliefs of the kind "that we have a reason not to torture people,

that we have reason to give aspirin to people with headaches, etc." (Skarsaune, 2011, p. 237) turn out to be false. He writes:

> Beliefs of these kinds make up large sections of our total body of evaluative beliefs. If pleasure isn't usually good, then, or pain isn't usually bad, large portions of our body of evaluative beliefs contain many falsehoods. But then many of our evaluative beliefs are false. Depending upon just how large a fraction of our evaluative beliefs are tied up with pleasure and pain, and upon how seldom pleasure and pain in fact are good/bad, the upshot could even be that *most* of our evaluative beliefs are false (Skarsaune, 2011, p. 237).

If pleasure isn't usually good and pain usually bad then a large portion of our moral judgments will turn out to be false. If we're wrong about pleasure or pain then there is no relation between evolutionary pressures and the evaluative facts of realism. If there is no relation, then most of our evaluative judgments are probably false. But then, says Skarsaune, we'd be lost anyway. Darwinian dilemma or not.

Now, Skarsaune claims that if you accept Street's framing of the dilemma and you take on the first horn the only option available to the realist is the tracking account. But as he has shown, there is at least one more option, his own "pre-established harmony" account. Fair enough. However, if the realist takes this route then new problems arise. I believe there are two objections lurking in the vicinity.

The first objection is this: assuming the "pre-established harmony" account is correct, our evaluative judgments having to do with pleasure or pain are largely correct. And if you're a utilitarian this would be good news since these are the only kinds of judgment that matter in the long run anyway. But one could counter that value judgments

having to do with pleasure or pain are not *exhaustive* of the judgments we make in the moral domain. While considerations as to whether actions lead to pleasure or pain figure strongly in our moral assessments, they are not the only kinds of things we look for when assessing some actions as virtuous or as "called for" by ethics. It is not at all *obvious* that the reason we think actions like lying are wrong is because of their *consequences*. Street gives several other examples that aim to show that considerations about pleasure or pain aren't exhaustive: dispositions toward reciprocity, the appreciation of altruism, and dispositions to accord moral status according to social status are such examples (Street, 2006, ps. 113, 115-116). It is at least prima facie plausible therefore that pleasure/pain considerations aren't exhaustive of our value-theorizing. While Skarsaune grants this, he does not believe such value judgments represent a large enough sample such that we should be too afraid of the skeptical implications his harmony account would have on our moral knowledge. He writes

> if these kinds of evolutionary influence are all the would-be anti-realist debunker can point to besides the pleasure-pain mechanism, I can see no reason why the realist should be forced to the sceptical conclusion that evolutionary pressures probably have left us hopelessly off track in our evaluative beliefs (Skarsaune, 2011, p. 236).

I believe the anti-realist debunker can point to further moral wrongdoings or moral goods that aren't obviously right or wrong in virtue of their being conducive to pleasure or pain. Consider using people as means rather than as ends. Many hold that it is wrong to do so even if doing so would bring you pleasure while the other individual would be unaware of his being used. I've already mentioned telling the truth, even if the

truth leads to bad consequences. While of course there is ample *disagreement* about this, the fact that there is ample disagreement seems to show that there are competing intuitions at play within society. Our judgments of the wrongness of killing *even if* there is no pain involved in the death of another individual are yet another example. And yet another one are cases of individuals who watch child pornography. It is perfectly possible given today's technology to watch child pornography via online video streaming or piracy without financially contributing to the enterprise. Given this, it would be wrong to say that an individual streaming child pornography would be committing a moral wrong in virtue of his financial contribution to the suffering and exploitation of children. However, even if an individual is not causally contributing to the suffering of innocent children, it seems correct to say that watching this type of material is wrong, regardless of whether doing so *causes* suffering. As another example of moral judgments many of us share that have little to do with considerations of utility is the following story:

> In October, 1958, appelant (Ms. Angelynn York) went to the police department of Chino for the purpose of filing charges in connection with an assault upon her. Appellee Ron Story, an officer of that police department, then acting under color of his authority as such, advised appellant that it was necessary to take photographs of her. Story then took appellant to a room in the police station, locked the door, and directed her to undress, which she did. Story then directed appellant to assume various indecent positions, and photographed her in those positions. These photographs were not made for any lawful or legitimate purpose.

> Appellant objected to undressing. She stated to Story that there was no need to take photographs of her in the nude, or in the positions she was directed to take, because the bruises would not show in any photograph. . .
>
> Later that month, Story advised appellant that the pictures did not come out and that he had destroyed them. Instead, Story circulated these photographs among the personnel of the Chino police department. In April, 1960, two other officers of that police department, appelle Louis Moren and defendant Henry Grote, acting under color of their authority as such, and using police photographic equipment located at the police station, made additional prints of the photographs taken by Story. Moreno and Grote then circulated these prints among the personnel of the Chino police department (Rachels, 2010, ps. 112-113).

Now, as an attack against classical utilitarianism it would be fair to point out that even *if* the combined pleasure of the police officers outweighed the psychological damage Angelynn York suffered from the experience, it would *still* have been wrong to do this. It would have been wrong *even if Angelynn York never found out* the pictures were circulated. It seems to me this is a pretty widely held intuition, that *there's something further* that settles the permissibility or impermissibility of the situation. And appeals to the *something further* are as common in our culture as appeals to mere utility.

There may be more examples, but an exhaustive examination of moral judgments tied to pain and pleasure contrasted with those that aren't would be too ambitious an undertaking here. I will therefore leave the issue by noting that I am skeptical of Skarsaune's claim that pain/pleasure considerations drive most of our moral theorizing and that *if* we could secure *those* then all would be well. Of course this doesn't *refute*

Skarsaune, but it should be enough to make realists uncomfortable with following his lead. If we accept Skarsaune's approach, how much moral knowledge would be lost, and how much are realists *willing* to let go of?

My second objection to Skarsaune regards the *justification* of our moral judgments if his account is correct. If evolution *programmed us* to regard pain as bad and pleasure as good, and Skarsaune seems to believes this as he writes that "we *can't help* but think that it is bad to starve, to have our arms and legs broken, or to be shunned by everyone we know" (Skarsaune, 2011, 235), then it becomes an interesting question as to how our beliefs about such matters could be *justified.* A thought experiment might help.

Suppose an individual comes to know that while he slept a scientist planted a chip in his brain that would produce the belief that there is and always has been a house at the end of the street. Suppose further that he finds a note explaining the nature of the chip, its function, and informing the individual that said chip is currently implanted in his brain. This particular individual has no knowledge about the moral character of the scientist, he has no way of knowing if this scientist is playing a practical joke by making him have a false belief or if for some strange reason he has implanted a chip that forces him to have a true belief about something he would have easily believed had he merely looked out the window anyway. Let us further stipulate that this individual, were he to look outside, would see a house, regardless of whether one is actually there or not. Would this individual's belief that a house exists at the end of the street be *justified* given what he knows about the chip? Without further information about the scientist's intentions and moral character it is hard to see how this belief could be justified.

Now let's suppose that this individual rightly distrusted his new-found belief in the house at the end of the street and decided to look for independent evidence of the house's existence. Is there independent evidence he can consult in order to find out whether the house truly exists that would *justify* any updated beliefs? Quite easily. He can look out the window to see if any cars pull up to the house (under the assumption that the chip *only* makes him see a house, not inhabitants or people who aren't there). He can consult local maps, ask around, or even walk up to the house and touch it. The point of this is straightforward: if we find ourselves with a pre-programmed belief in X, our belief in X isn't justified unless we know further details about how we came to possess this belief and further relevant information.

This thought experiment is very similar to Richard Joyce's "belief pill" thought experiment:

> Pretend there were such things as belief pills, such that taking one would inevitably lead to the forming of a certain particular belief (while at the same time invoking amnesia about the taking of the pill and, to be on the safe side, amnesia about the existence of such pills in general). Suppose that there were a pill that makes you believe that Napoleon won Waterloo, and another one that makes you believe that he lost. Suppose also that there were an antidote that can be taken for either pill. Now imagine that you are proceeding through life happily believing that Napoleon lost Waterloo (as, indeed, you are), and then you discover that at some point in your past someone slipped you a "Napoleon Lost Waterloo" belief pill. It is not a matter of your learning of the existence of such pills and having no way of knowing whether you have ever taken one; rather, we are imagining that

> you somehow discover beyond any shred of doubt that your belief is the product of such a pill. Should this undermine your faith in your belief that Napoleon lost Waterloo? Of course it should (Joyce, 2006, p. 179).

Of course no one thinks evolution selects for *particular* beliefs; rather, it selects for *dispositions* to hold beliefs. Joyce's claim is that "there is a specialized innate mechanism (or series of mechanisms) whose function is to enable this type of acquisition. This mechanism, I hypothesize, comes prepared to categorize the world in morally normative terms; moral *concepts* may be innate even if moral beliefs are not" (Joyce, 2006, p. 181). Joyce alters the original thought experiment to accommodate for this by supposing the following:

> …the imaginary belief pills do not generate particular propositional beliefs but rather dispose you to form beliefs involving a particular concept—a concept that otherwise wouldn't figure in your beliefs. Thus, rather than a pill that makes you believe that Napoleon lost Waterloo, it's just a 'Napoleon pill' that makes you form beliefs about Napoleon in general. Without this pill you would never have formed any beliefs about Napoleon at all. We needn't worry too much about what other factors determine the precise content of these Napoleon beliefs; perhaps it is determined randomly, or perhaps there are certain environmental triggers you discover beyond any doubt that you were slipped one of these pills a few years ago. Does this undermine all the beliefs you have concerning Napoleon? Of course it does. A belief is undermined if one of the concepts figuring in it is undermined (Joyce, 2006, p. 180).

Of course this wouldn't entail that the beliefs in question are *false*, you simply don't know what their truth values are. A reasonable person therefore would suspend judgment and take the antidote to the belief pills and from there on set out to investigate the world and whether the beliefs in question are true. If we concede that individuals in such a scenario would have unjustified beliefs and should suspend judgment, then we are primed to accept Joyce's analogy since our moral judgments are products of a similar sort of process. He writes:

> 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. They *could* be true, but we have no reason for thinking so (Joyce, 2006, p. 210).

This skeptical conclusion would entail agnosticism about the existence of *anything at all* that is morally right or wrong. It is, for Joyce, a matter of "accepting the possibility that describing the world in moral terms is in the same ballpark as taking horoscopes seriously or believing that ancestral spirits move invisibly among us" (Joyce, 2006, p. 182).

Joyce's conclusion is simply that, armed with an empirically confirmed evolutionary theory about the source of our moral judgments (whether his own account is true, or some forthcoming one), we should refrain from believing that the moral

judgments we possess are true. This does not entail *nihilism,* if construed as the belief that all moral judgments are false. Rather, it is a type of *moral agnosticism.*

Now, it seems to me that accepting Skarsaune's account puts us in a similar situation. We *can't help but* believe in the badness of pain and goodness of pleasure. It would have been programmed into us *whether true or not.* Now, it should be of no comfort to the realist that we could potentially possess true beliefs about pain *if our possessing these beliefs is a result of cosmic coincidence.* The fact that pain *is* intrinsically bad did not *cause* our believing it to be so, *we would have believed it even if it were not so.* We must remember that Street's argument challenges our claims to *moral knowledge*. Skarsaune's account would give us some *true beliefs*. On the conditional that pain *is* intrinsically bad and pleasure intrinsically good many of our moral judgments would be *true,* but what about *justification*? For what independent evidence, other than our own distaste for it, distaste we have no choice but to have, could we point to that would answer the question of pain's intrinsic badness? We can't justify our beliefs about the badness of pain in terms of other evaluative judgments that have nothing to do with pain or pleasure since we have already seen that *those* would be suspect if we were to accept Skarsaune's account. And we can't justify our beliefs about the badness of pain in terms of other beliefs regarding pain for fear of begging the question. What vantage point then could we use to *justify* our evaluative judgments if we accept Skarsaune's account? It isn't clear. It is up to Skarsaune to alleviate these worries.

Before we proceed with more objections to Street, let us consider one objection to Richard Joyce's argument. It should be noted that only if you're a *moral realist* should you consider the implications of this type of evolutionary debunking to be agnosticism.

Fritz Alhoff argues that there are a variety of meta-ethical positions that remain unscathed even if Joyce's evolutionary argument succeeds. Meta-ethical views like subjectivism, where "X is wrong" merely means "I disapprove of X" (Alhoff, 2009, p. 101), cultural relativism where "X is wrong" means "We disapprove of X" (2009, p. 101), and Divine Command ethics in which "X is wrong" means "God disapproves of X" (2009, p. 101) are all immune to this type of debunking. The first two are immune in virtue of the truth maker for a moral utterance being one's own subjective disapproval or society's overall disapproval, something which is quite easily verifiable. The truth makers for divine command theory would be the actual commands of the specific deity in question. Even if these commands aren't epistemically accessible to us, they wouldn't be so in virtue of any type of evolutionary debunking argument.

Besides these, there are several other types of meta-ethical positions that are immune to Joyce's evolutionary debunking argument. Several types of contractarianism and contractualism are such views. What they have in common is that both are types of *constructivism*. What differentiates constructivism from moral realism is that realism posits a realm of "moral facts" that exist independently of human beings. Realism is a view that

> there are moral truths that obtain independently of any preferred perspective, *in the sense that the moral standards that fix the moral facts are not made true by virtue of their ratification from within any given actual or hypothetical perspective* (Shafer-Landau, 2003, p. 15).

Further,

> the truth of any first-order normative standard is not a function of what anyone happens to think of it. Such standards, if true, are not *made true,* and, in particular, are not correct in virtue of being vindicated by some process of (inter)personal election or approbation (Shafer-Landau, 2003, p. 16).

In contrast, constructivist views "understand normative truth as not merely *uncovered by or coinciding with* the outcome of a certain procedure, but as *constituted* by emergence from that procedure" (Street, 2010, p. 111). The constructivist holds that "there *is* no truth about justice independent of the procedure; the truth of the principles *consists* in the fact that they are the ones that would be selected by the parties in the original position" (Street, 2010, p. 111).

We should now have a clearer outline as to why Joyce's evolutionary debunking argument might not threaten constructivist views. If we consider his belief pill thought experiment, when we find ourselves with a belief about Napoleon, we might rightly wonder if this belief accurately corresponds to the truth about Napoleon. When we find that we've been given a pill that produces beliefs about Napoleon, we would be right in withholding judgment about the truth or falsity of our Napoleon beliefs without further evidence, either evidence about the reliability of the pill and the intentions of the pill-maker, or evidence about the real Napoleon. If we are moral realists of the non-reductionist variety and are committed to naturalism, there is no further evidence that we can appeal to that would show us whether our moral beliefs accurately correspond to the realm of moral facts or not. Joyce's evolutionary debunking argument leaves the realist in a position where agnosticism seems to be rational course of action.

On the other hand, moral truth for the constructivist *is constituted by* what would emerge from a process of construction (what that process is varies according to different theorists, but as Street points out, the notion of a procedure is merely a heuristic device), what matters is that

> …standards of correctness in the normative domain are not set by an 'independent order' that holds apart from the attitudes of valuing creatures, but rather are set by, or from within, the point of view of beings who make some normative judgments or other (Street, 2010, p. 117).

The point is this: if you're a moral realist and you're confronted with Joyce's genealogy of the moral sentiments, a genealogy in which truth is never invoked in order to explain why we possess it, you'd immediately wonder if there is any reason to believe that our current moral judgments "map onto" the relevant moral facts. A constructivist on the other hand need not worry about such "mapping" since there are no independent moral facts anyway. A Kantian constructivist for instance believes that "the truth of a normative claim consists in its being entailed from within the practical point of view, where the practical point of view is given a formal characterization" (Street, 2010, p. 118) To give the practical point of view a formal characterization is to "give an account of the standpoint of valuing or normative judgment *as such*, where this involves giving an account of the attitude of valuing that does not itself presuppose any substantive values but rather merely explicates what is involved in valuing anything at all" (Street, 2010, p. 118). But nothing in Joyce's genealogy of morals casts *this* into doubt. We're creatures who value things and act on the basis of reasons. For a Korsgaardian constructivist, those are the "basic ingredients" needed to get the full package of morality. Certain normative

conclusions *follow* from these non-normative facts. We merely need to be in the right sort of epistemic position to see these entailments. And of course nothing in the evolutionary debunking argument provided by Joyce casts doubt on *this* part either.

It seems then that Joyce's evolutionary debunking argument is indeed a threat to morality, but only morality construed in a certain way. If the only two games in town are moral realism and moral agnosticism then agnosticism would be the way to go; but these *aren't* the only two options available to us. Constructivist theories remain unscathed. A moral realist *should* worry about a full genealogy of our moral sentiments. Such a genealogy threatens to relegate sui generis moral facts to the bin of irrelevancy given that they seem to play no explanatory role in any comprehensive story we can tell about our moral sentiments.

Kevin Brosnan

Kevin Brosnan takes Street's argument to be a *general* epistemic argument. Evolution isn't integral to the argument, for it has the following general form: "So long as our moral beliefs are caused *by some process or other,* this process will either track presumptive moral facts or it will not" (Brosnan, 2011, p. 53). Brosnan argues against Street on several fronts. First, he disagrees with Street on the claim that rational reflection cannot help us correct "contaminated" beliefs. According to Brosnan, Street overlooks the fact that rational reflection is not necessarily merely about reflecting *with* some moral beliefs that are false, but can also be a process of reflection *on* those moral beliefs. As an illustration, he gives the following example:

> J1: One ought to help kin, but not members of one's community who aren't kin (Brosnan, 2011, p. 57).

> J2: One ought to help unrelated members of one's community, but not outsiders (Brosnan, 2011, p. 57).

Assuming the falsehood of both of the above, it does not follow that upon reflection we can't extract whatever kernels of truth they may possess. Upon reflection one may find that the reasons we have to help one another have little to do with physical proximity or genealogy but rather have to do with the fact that others would suffer in the absence of help that would reduce their suffering. In so doing "one might come to judge that one ought to help others, irrespective of whether they are kin or outsiders" (Brosnan, 2011, p. 57). Brosnan explains:

> …the falsity of a moral belief does not automatically exclude it from its being used effectively in a reflective process aimed at moral progress. So even if evolutionary influences have a corrosive effect on justification, the judicious use of our reflective capacities might be more than enough to counteract this effect (Brosnan, 2011, p. 57).

Fair enough, though the success of this reflection will be contingent on the amount of contamination present in the first place. As Brosnan grants, no reflection in the world is likely to fix *radically false* beliefs. He is correct in his point that we can correct for beliefs that have gone astray, but we'd have to have a good idea of just how astray they are before we can be confident that our updated, post-reflective beliefs are likely to have steered us back on track. Now, in the absence of argument it would be presumptuous of the anti-realist to merely assume the extreme view that evolution has steered us towards *radically* false beliefs. And yet, in the absence of argument, the realist cannot simply assume the convenient alternative either, so in the absence of a compelling

reason for either alternative we might wish to remain agnostic, and agnosticism about our updated post-reflective beliefs would commit us to an anti-realist view. Brosnan's point is this: *If* our moral judgments aren't thoroughly contaminated, then rational reflection can get us back on track. An anti-realist can grant this conditional but respond that this still says nothing about *our present situation*. We're interested in whether *our* moral judgments are rationally justified. The Darwinian dilemma is aimed *at us*, and so conditionals of this sort aren't helpful if we're concerned about finding out what's the case *in the actual world*.

The second contention of Brosnan's is that natural selection *can* indeed track moral facts. In arguing for this he helps himself to Elliott Sober's selection *of* versus selection *for* distinction. Natural selection would have failed to track for moral facts if both of the following are true: "a) there was no selection *for* a tracking capacity, or b) there was no selection *of* a tracking capacity". (Brosnan, 2011, p. 58) Brosnan then argues that a) does not support the implication that our moral beliefs are products of a process that fail to track moral facts, and that b) does imply this, but there are good reasons to believe b) is false anyway.

The selection *of* versus selection *for* distinction is straightforward. The white color of bones was not selected *for,* but it was selected *of.* Whiteness didn't contribute anything such that natural selection would select *for* it, but bones themselves *were* selected for, and bones are white, so their whiteness "piggy-backed" and was selected *of*. Similarly, Brosnan argues that the mere fact that the psychological mechanisms responsible for our moral beliefs weren't selected *for* their ability to track moral facts, it nevertheless could be the case that this ability was selected *of.* So the fact that natural

selection did not select *for* truth tracking does not necessarily commit one to the idea that our faculties *don't* track the truth. This ability could have been selected *of*.

How would this happen? Here Brosnan adopts a similar strategy to Skarsaune. Assume for the sake of argument that cooperation with others is morally good. If this belief was favored by natural selection because it enhanced our capacity for helping behavior, and helping behavior promotes fitness, then we'd have a case of a true moral judgment being selected *of*. What's being selected *for* are dispositions to believe certain things that contribute to fitness, and along for the ride being selected *of* is the (true) moral judgment that cooperation with others is morally good. More generally, if one assumes that "what's morally good has to do with behaviors that promote rather than hinder wellbeing, then part of what makes cooperation good is that it typically has this effect" (Brosnan, 2011, p. 60).

The similarity between Brosnan's and Skarsaune's approach makes Brosnan's approach susceptible to the same kinds of objections raised against Skarsaune. We're going to believe X is true *whether true or not*. And the *truth* of X does not *cause* our believing it to be so. Even if human well-being *were* intrinsically good this moral fact would not *cause* our belief that it is so. The forces of natural selection have influenced our valuing of our own well-being for obvious reasons that are not related to the *truth* of this moral fact. As with Skarsaune's account, we might be left with some *true* beliefs, but justification goes out the window. Brosnan's defense of the possibility of moral knowledge then fails for the same reasons Skarsaune's defense fails.

Erik Wielenberg

Responding to evolutionary debunking arguments Erik Wielenberg argues that *if* moral barriers or rights exist, then there is also a plausible explanation of how we could have knowledge of these rights. In his approach, Wielenberg assumes the following:

> 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. . .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 (Wielenberg, 2010, p. 449).

Wielenberg's general strategy is this: the existence of rights is generally accepted to be contingent on the possession of certain cognitive faculties. He writes

> In order to form the belief that one has certain rights, one must be able to have some grasp of the concept of rights. 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 faculties that lead such

> beings to believe that they possess moral barriers also entail the presence of those very barriers (Wielenberg, 2010, p. 450-451).

Wielenberg argues that there is a *logical* connection between the presence of rights and the possession of cognitive faculties responsible for the *belief* that there are such rights such that if one *believes* one has such rights, then ipso facto one *does* have such rights. Now, this does assume that evolutionary processes are *generally* reliable, but such an assumption is reasonable as the previous discussion on Plantinga illustrated and as several other authors have argued (P. Carruthers, 1992). *If* we grant this, then Wielenberg's model entails that "at least some moral beliefs are produced by reliable processes" (2010, p. 452) Thus Wielenberg's model entails the following:

> 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 (Wielenberg, 2010, p. 452).

This model shows how a belief about the existence of a moral barrier that applies to oneself can be justified. Extending his model to cover justification for beliefs about the existence of moral barriers for *others* Wielenberg argues that creatures like us tend to reason in accordance with principles like the likeness principle according to which

"entities that are alike with respect to their known properties are alike with respect to their unknown properties" (2010, 453). Given the laws of nature that hold in our universe, Wielenberg thinks this principle has a high degree of conditional reliability:

> That is, when it takes true beliefs as inputs, it tends to produce true beliefs as outputs. 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 (Wielenberg, 2010, p. 453).

And so, if beliefs produced by highly reliable processes operating on true beliefs have a degree of warrant sufficient for knowledge, then this model implies that at least some of our moral beliefs constitute *knowledge.*

Wielenberg gives us an evolutionary story as to why individuals would come to believe they are bearers of rights. This story does not invoke the *truth* of the matter in order to account for why we believe we have such rights. Wielenberg's basic account is that we are disposed to view ourselves "as surrounded by a kind of moral barrier that it is wrong, unjust, evil, or somehow morally inappropriate or illegitimate for others to cross" (Wielenberg, 2010, p. 445) because it is fitness-enhancing to believe so. The disposition to hold such beliefs can be fitness enhancing given that at least part of the explanation for why we believe such things is that they function as conversation stoppers.

> 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 (Wielenberg, 2010, p. 445).

This account doesn't presuppose the truth of these beliefs in order to account for why we might be disposed to hold them. It does not answer the question "why do we believe X?" by saying "Because it is self-evident that X, and we hold on to beliefs that are self-evident." This account explains our propensity to believe in the existence of moral boundaries because it was fitness-enhancing to do so. In other words, it is likely that we would have believed X *even if X were not true.* So a question naturally arises. If some creature believes in the existence of a moral barrier which others ought not transgress, what *justifies* this creature's belief that X? Wielenberg's answer is that X is simply properly basic. Wielenberg contends that the specific way in which we come to believe X as true isn't relevant. One such way that a belief in X can be justified is that we simply have an intellectual appearance (intuition) that X is true, and that this intuition is enough to justify the moral belief in question. But this is too quick. If X were *not* true then we would presumably *still* have an intuition that X is true. It would seem to us like a moral barrier exists *whether one exists or not*. Wielenberg objects to this by claiming that such a scenario only arises in an impossible world:

> In section 4, 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." (Wielenberg, 2010, 459)

But is this true? *Can* we simply assume that if I lacked these rights I would similarly lack the *belief* that I have such rights? Given the amount of disagreement about the existence and coherence of the notion of rights there should be much room for caution in granting this to Wielenberg (M. Ignatieff, 2003; J. Waldron, 1987). If it would seem to us that a moral barrier exists regardless of the *actual* existence of one, then an *intuition* that one does in fact exist *can't* confer warrant on the belief in question. Only if we had good reason to believe such an intuition would not be present if the belief were false would we have good reason to trust the reliability of the intuition. Wielenberg thinks the absence of the right would entail the absence of the cognitive faculties necessary for being a bearer of rights. Whether or not this is true would require an independent debate about the existence and coherence of rights, something I cannot dive into, much less hope to settle, here. Whether one thinks Wielenberg's strategy works might then depend on what views one has about the nature of human rights and their coherence. I myself am unsure about these matters and will therefore leave the question open here.

Street's argument forces moral realists to sketch out an account of moral knowledge that is compatible with our evolutionary origins. I have surveyed the relevant literature and critically evaluated the accounts offered thus far. While some of these accounts are promising, more work needs to be done to spell out just how it is that we can

be confident in possessing moral knowledge if this knowledge is taken to be knowledge of mind independent, sui generis, moral facts. There is no reason to believe such an account is not possible. Given the current state of the literature however, the Darwinian Dilemma for Realist Theories of Value still confronts moral realists with difficult choices.

Conclusion

As we have seen, our shared evolutionary history is far from philosophically benign. The unguided evolutionary development of our cognitive faculties presents a challenge to creatures who must rely on these faculties to form beliefs about the world around them and make decisions about how to act based on these beliefs. Some philosophers, such as Alvin Plantinga, have argued that the implications of unguided Darwinian processes have far reaching consequences that render a creature's entire web of beliefs unreliable. Without someone overseeing the evolutionary process and guiding it towards the development of cognitive faculties that are aimed at truth, we simply have no reason to trust the mechanisms of Darwinian evolution coupled with naturalism to produce reliable faculties. Evolutionary considerations therefore lead to a thoroughgoing skepticism. Other authors have similarly argued for skeptical hypotheses based on evolutionary considerations, but have restricted their conclusions to the moral domain.

I have examined the arguments of four authors, Alvin Plantinga, Michael Ruse, Sharon Street, and Richard Joyce, as well as the arguments of their critics, and have aimed to show that of these four, only two of these arguments are convincing.

Plantinga's critics have been correct in pointing out how improbable it would be for a creature to possess on average mostly false beliefs and yet nevertheless behave in ways that prove adaptive from an evolutionary point of view.

While I agree with Michael Ruse's general hunch that evolutionary considerations threaten to undermine our most deeply held moral convictions, his arguments to *support* this conclusion prove not to be sound. His further contention that even though evolution reveals the illusory nature of our moral beliefs we can still be confident in their

vindication is similarly unsound. As a general strategy, Ruse has the right idea, but fails to deliver the conclusions he seeks.

Sharon Street and Richard Joyce refine their evolutionary debunking arguments and restrict their domains of application to fashion what I believe are two powerful challenges to moral realism. I have argued that their critics have so far failed to adequately address their arguments and, as a consequence, these arguments present moral realism with its strongest epistemic challenge yet.

As we have seen, however, evolutionary considerations do not threaten moral constructivism as defended by philosophers like Christine Korsgaard and Sharon Street. Evolutionary epistemic challenges only threaten realist *conceptions* of morality. The reader who has found the arguments in defense of these epistemic challenges compelling might therefore want to look for alternative meta-ethical conceptions of morality that are not threatened by the evolutionary considerations explored in this thesis. What view one ultimately adopts will depend on the relative strengths and weaknesses of the offered theories on hand. The reader who finds cognitivist theories attractive might therefore want to explore moral realism's main competitor in this area: constructivism. Of course nothing I have argued has given the reader any reason to prefer constructivism to the moral agnosticism prescribed by Richard Joyce, so I leave it to the reader to further explore these topics and come to a reasoned decision as to what meta-ethical theories seem most plausible once we have rejected moral realism as a viable candidate.

Works Cited

Alhoff, F. (2009). The evolution of the moral sentiments and the metaphysics of morals. *Ethical Theory and Moral Practice*, *12*, 97-114.

Beilby, J. (Ed.), (2002), *Naturalism Defeated?* (1 ed.). Ithaca & London: Cornell University Press.

Brandt, R. (1950). The emotive theory of ethics. *Philosophical Review 59* (3):305-318.

Brink, D. O. (1989). *Moral realism and the foundations of ethics*. (pp. 182-197). Cambridge University Press.

Brosnan, K. Dissolving A Dilemma: Why Darwinian Considerations Don't Confront Moral Realism With Hard Choices. Unpublished Manuscript.

Brosnan, K. (2011). Do the evolutionary origins of our moral beliefs undermine moral knowledge? *Biology and Philosophy*, *26*, 51-64.

Burgess, C. A., Kirsch, I., Shane, H., Niederauer, K., Graham, S., & Bacon, A. (1998). Facilitated communication as an ideomotor response. *Psychological Science*, *9*(1), 71-74. Retrieved from http://www.jstor.org/stable/40063250

Carruthers, P. (1992). Human Knowledge and Human Nature: A New Introduction to an Ancient Debate. Oxford University Press.

Chalmers, D. J. (2009). The matrix as metaphysics. In *Science fiction and philosophy* (pp. 33-52). United States: Blackwell Publishing.

Churchland, P. J. (1987). Epistemology in the age of neuroscience. *Journal of Philosophy*, *84*, 546-553.

Darwin, C. Letter to william graham, down. In *The life and letters of Charles Darwin including an autobiographical chapter* (pp. 315-316).

Fales, E. (2002). Darwin's doubt, calvin's calvary. In J. Beilby (Ed.), *Naturalism Defeated?* (1 ed., pp. 43-58). Ithaca & London: Cornell University Press.

Fodor, J. (2002). Is science biologically possible? . In J. Beilby (Ed.), *Naturalism Defeated?* (1 ed., pp. 30-42). Ithaca & London: Cornell University Press.

Graber, A. (2012). Medusa's gaze reflected: A darwinian dilemma for anti-realist theories of value. *Ethical Theory and Moral Practice*, *15*, 586-601.

Griffiths, Paul E. & Wilkins, John S. (forthcoming). When do evolutionary explanations of belief debunk belief? In *Darwin in the 21st Century*.

Ignatieff, M., 2003, Human Rights, Sovereignty, and Intervention. In Owen, N. (ed.) *Human Rights and Human Wrongs: The Oxford Amnesty Lectures 2001*, Oxford: Oxford University Press.

Joyce, R. (2013). The evolutionary debunking of morality. In *Reason and responsibility: Readings in some basic problems of philosophy.* Cengage Learning.

Joyce, R. (2006). *The evolution of morality*. (1st ed.). Cambridge & London: The MIT Press.

Kahane, G. (2011). Evolutionary debunking arguments. *Noûs*, *45*, 103-125.

Korsgaard, C. M. (1996). *The sources of normativity.* Cambridge University Press.

Law, S. (2011). Plantinga's belief-cum-desire argument refuted. *Religious Studies, 47*, 245-256.

Plantinga, A. (2002). The evolutionary argument against naturalism: An initial statement of the argument. In J. Beilby (Ed.), *Naturalism Defeated?* (1 ed., pp. 1-12). Ithaca & London: Cornell University Press.

Popper, K. & Eccles, J. (1977) *The Self and Its Brain*, New York: Springer-Verlag.

Rachels, J. (2010). *The elements of moral philosophy.* New York: McGraw Hill.

Ramsey, W. (2002). Naturalism defended. In J. Beilby (Ed.), *Naturalism Defeated?* (1 ed., pp. 15-29). Ithaca & London: Cornell University Press.

Rawls, J. (1971). *A theory of justice.* Cambridge, Mass: Harvard University Press.

Ruse, M. (1986). Evolutionary ethics: A phoenix arisen. *Zygon*, *21*, 95-112.

Shafer-Landau, R. (2003). *Moral realism: A defence.* Oxford University Press.

Skarsaune, K. O. (2011). Darwin and moral realism: Survival of the iffiest. *Philosophical Studies*, *152*, 229-243.

Street, S. (2006). A darwinian dilemma for realist theories of value. *Philosophical Studies*, *127*, 109-166.

Street, S. (2010). What is constructivism in ethics and metaethics? *Philosophical Studies*, *127*, 363-384.

Taylor, R. (1963) *Metaphysics*, Englewood Cliffs, NJ: Prentice Hall.

Van Inwagen, P. (1990). *Material Beings*. Cornell University Press.

Waldron, J., (ed.), 1987a, *Nonsense Upon Stilts: Bentham, Burke, and Marx on the Rights of Man,* London: Methuen.

Wielenberg, E. J. (2010). On the evolutionary debunking of morality. *Ethics*, *120*, 441-464.

Young, L. (2013). Moral realism as moral motivation: The impact of meta-ethics on every day decision-making. *Journal of Experimental Psychology*, *49*(2), 302-306.

Thesis and Dissertation Services