6-2016

Darwinian Debunking Reconsidered

Amanda J. Favia
Graduate Center, City University of New York

DARWINIAN DEBUNKING RECONSIDERED

by

AMANDA J. FAVIA

A dissertation submitted to the Graduate Faculty in Philosophy in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2016

DARWINIAN DEBUNKING RECONSIDERED

by

AMANDA J. FAVIA

This dissertation has been read and accepted for the
Graduate Faculty in Philosophy to satisfy the dissertation
requirement for the degree of Doctor in Philosophy.

	Rosamond Rhodes
_____	_____
Date	Chair of Examining Committee
	Iakovos Vasiliou
_____	_____
Date	Executive Officer

Supervisory Committee:

Bernard S. Baumrin, Advisor

Rosamond Rhodes

Gary Ostertag

Virginia Held

Jesse Prinz

THE CITY UNIVERSITY OF NEW YORK

Abstract

DARWINIAN DEBUNKING RECONSIDERED

by

Amanda J. Favia

Adviser: Bernard S. Baumrin

What can evolutionary theory tell us about morality? From descriptive claims that explain morality as an evolved trait to normative (or prescriptive) claims that rely on evolution to describe how humans ought to behave, philosophers have debated whether or not evolutionary theory can or should inform moral theory. Most recently, the debate about evolutionary ethics has shifted to metaethics. In this case, philosophers have sought after evolutionary explanations in the hopes of resolving long-standing debates between moral realists and moral antirealists.

These metaethical debates have centered on what are called *Darwinian debunking arguments*. Proponents of the Darwinian debunking argument, such as Michael Ruse, Sharon Street, and Richard Joyce, seek to use evolutionary explanations to undermine moral realism. Opponents of Darwinian debunking arguments, such as David Enoch and Erik J. Wielenberg, attempt to defend moral realism by establishing a brand of moral realism consistent with the same evolutionary explanations that debunkers use to undermine moral realism. While those engaged in the debate generally agree that morality is an evolved trait and that evolutionary explanations can help resolve the metaethical controversy, they disagree over which metaethical position—moral realism, moral antirealism, or moral skepticism—those evolutionary explanations support.

This essay provides an analysis of Darwinian debunking arguments, the debate that surrounds it and, more generally, the use of evolutionary explanations to resolve questions about

the nature and justification of moral claims. The debunking debate ends in a deadlock, however, since, as I argue, evolutionary premises fail to yield any substantive metaethical conclusions. This occurs for several reasons. First, evolutionary explanations turn out to be not only inherently speculative or hypothetical, but also historical and irregular in such a way that renders them unsuitable for resolving this metaethical debate. This leads to the second problem, that of "inferential opaqueness": the precise connection between the evolutionary explanations (the premises) and the rejection (or acceptance) of moral realism (the conclusion) at stake in the debunking literature remains unclear. The third problem rests with the theoretical assumptions—epistemological, ontological, moral and scientific—that are embedded in both sides of the debunking debate. Such assumptions are essential to how evolutionary explanations are interpreted and thus how they are meant to support various metaethical conclusions. As a result, evolutionary considerations play a far less significant role in establishing metaethical claims than these views suggest. Finally, the challenge provided by the genetic arguments utilized on either side of the debunking debate is much more limited then their adherents seem to realize. Here I draw on and develop recent work by Kevin C. Klement and Katia Vavova.

Acknowledgments

This dissertation would not have been possible without the love, support, and encouragement of so many. I owe a great deal to my adviser, Stefan Baumrin, whose initial enthusiasm about this project gave me the confidence to believe it was a worthy endeavor—thank you for the steady support, insightful questions, and suggestions throughout the writing of this dissertation. I am also extremely grateful for the help and advice provided by Rosamond Rhodes. My development as a philosopher and an ethicist have been greatly influenced by all of the opportunities you have given me as an Ethics Fellow at the Icahn School of Medicine at Mount Sinai. This dissertation has also greatly benefited from the guidance of Gary Ostertag—thank you for taking the time to engage in countless conversations and for encouraging me to strive for excellence. I also greatly appreciate the time and effort of my readers, Virginia Held and Jesse Prinz.

To all of my former classmates at the CUNY Graduate Center who have helped me develop both intellectually and personally, I am forever grateful. In particular, I'd like to thank Chris Sula, Kamili Posey, Jason Altilio, Lily Frank, Paul Cummins, Ornaith O'Dowd, Steven Birnbaum, and Leonard Finkleman. Many conversations with my colleagues at Mount Sinai, including Nada Gligorov, Katherine Mendis, Phoebe Friesen, and Carolyn Plunkett, as well as many members of the Oxford-Mount Sinai Consortium, have also helped shape this work. I appreciate all of the support my colleagues at Nassau Community College have given me as well. Specifically, I want to thank Frank Pupa for reading, revising, and always listening.

There has also been endless emotional support and encouragement from so many wonderful friends over the years. Very special thanks go to my dear friends Lisa Barkley, Meredith Grausso, Heather Hallam, Tara Power, Joselyn Allen, Laura Salvaggio, Christine

Heller Murphy, Caitlin Healy, Weishun Cheung, Sean Noyce, Katya Usvitsky, Heidi Ryti, and Josephine Atria. I owe the biggest debt of gratitude, however, to Kamili Posey, Chris Sula, and Abbi Leman—this dissertation wouldn't have been possible without you.

Finally, I am extremely grateful for my family. Specifically, I want to thank my brother-in-law Chris Grzymala, my uncle Brad Broadhead, and my cousins Rich and Mary Beth Broadhead for their emotional and economic support over the years. A very special thank you to my "uncle", Dr. Dexter Jefferies, who has provided intellectual guidance and inspiration since I was a little kid. And to my sister Emma Grzymala—thank you for the love and laughter that motivates me. Last, but certainly not least, to my parents, Frank and Pamela Favia. Thank you for providing a home that fostered the intellectual and creative freedom to let me be "just the way I am".

Table of Contents

Introduction

At first glance, evolutionary ethics is a very attractive view. It promises to provide an explanation of morality compatible with naturalism while at the same time abandoning the theological and ontological baggage of prior theories. Historically, evolutionary ethics has focused on normative, prescriptive theory, relying on evolutionary explanations to help establish the "good". This naturalistic notion of the "good" was then employed to provide prescriptions about how humans ought to behave. More recently, however, such projects have turned their focus directly to the foundations of ethics, putting aside prescriptive questions and instead, tackling fundamental questions in metaethics such as whether moral facts exist or how epistemic access to moral facts is possible. The current debate considers whether or not evolutionary explanations support or undermine moral realism, moral antirealism or moral skepticism. But is using evolutionary explanations a useful strategy for resolving these debates in metaethics? My dissertation attempts to answer this question with a negative thesis—that evolutionary explanations cannot resolve metaethical issues.

Chapter One identifies the fundamental question that drives evolutionary ethics in all of its various forms—*where does morality come from?* Although this type of inquiry is not unique to evolutionary ethics, the manner in which evolutionary ethics seeks to answer it is. Closer examination of this question reveals some of the fundamental features of an evolutionary approach to ethics. To begin, an account of the origins of morality using evolutionary explanations is considered not only relevant to morality for the evolutionary ethicist, but also a necessary feature of any adequate account of morality. This leads us to the underlying structure and methods employed in evolutionary ethics—evolutionary explanations are used to justify various aspects of moral theory. For example, some views argue that an evolved innate moral

sense justifies moral realism. The project is, thus, not merely descriptive, but also has a normative dimension.

The first part of this chapter, then, provides an in-depth discussion of the role that explanation and justification play in relation to inquiry. Whether or not an explanation is successful depends on the initial inquiry; an explanation is only satisfactory if it answers the question being asked. Explanations are also distinct from justifications; the former attempt to help clarify or provide understanding, whereas the latter seek to prove what is right or correct. I argue that attention to this basic distinction is absent from much of the literature on evolutionary ethics, with the result that evolutionary explanations are often used to justify various aspects of morality.

The focus then shifts to moral philosophy more broadly, in particular the two general approaches taken to answering the question of where morality comes from—the *genealogical approach* and the *theoretical approach*. The genealogical approach understands the initial question to be concerned with the origins or the etiology of morals, whereas the theoretical approach interprets the question to be concerned with seeking the foundations of morality. Evolutionary ethics, however, merges these approaches and attempts to use the evolutionary origins of morality to either support or undermine the foundations of morality. In these cases, evolutionary explanations are used to justify various metaethical positions: moral realism, moral antirealism, and moral skepticism. The question then arises whether or not this combined approach can be successful. The remainder of the essay attempts to answer this question.

Through a historical overview, Chapter Two provides an analysis of evolutionary ethics as a methodology that merges the genealogical and theoretical approaches. It begins by focusing on Darwinian Moral Nativism, a set of views that each claim that human beings possess an

evolved innate moral sense (i.e., an adaptive capacity to make moral judgments). Both Charles Darwin and Herbert Spencer are representative of this position and laid the groundwork for many contemporary views in evolutionary ethics. Their views are discussed and compared with a particular class of moral theories that were popular before and during their time, moral sentimentalism and utilitarianism. This historical overview also considers the traditional objections to evolutionary ethics, namely, the naturalistic fallacy and Hume's thesis. Analysis of these traditional objections reveal a set of contemporary arguments that appear to avoid these criticisms—these views are called *evolutionary debunking arguments*. Evolutionary debunking arguments seek to undermine moral realism on the basis of morality's evolutionary origins. According to these arguments, since humans posses an evolved innate moral sense, moral realism is either untrue, supporting antirealism, or moral realism is extremely unlikely, supporting moral skepticism. Three prominent, and representative, debunking arguments are presented by Michael Ruse, Sharon Street, and Richard Joyce. Although these arguments appear to escape the traditional objections to evolutionary ethics—the naturalistic fallacy and Hume's thesis—I will show that they face their own set of insurmountable issues.

Chapter Three examines the debunking arguments of Ruse, Street, and Joyce. These are separated into two categories based on the claims they seek to undermine—*ontological debunking*, which targets the existence of moral facts, and *epistemological debunking,* which targets the epistemic status of moral beliefs. Both Ruse and Street employ ontological debunking, while Joyce uses epistemological debunking. A critical discussion of all three debunkers follows. The first (of several) significant problems that all debunkers face is particularly serious—namely, that of global epistemological skepticism, an issue that Henry

Sidgwick raised shortly after Darwin and Spencer published their views. The contemporary debunkers, however, fail to adequately address Sidgwick's worries about this problem.

Ruse, Street, and Joyce also have a set of significant opponents—moral realists who wish to maintain that humans possess an evolved innate moral sense, but who deny that this "fact" undermines moral realism. Chapter Four, therefore, presents some of these Darwinian Moral Realist views that oppose the debunking arguments. There are two categories of Darwinian Moral Realist responses: (1) third factor explanations, which concern realist epistemological claims and aim to establish additional explanations compatible with evolutionary moral realism, and (2) realist normative theories, which concern realist normative claims where particular normative moral theories are introduced that are consistent with evolutionary explanations.

The remainder of this chapter focuses on third factor explanations, since both categories face similar problems. In particular, it examines the third factor explanations presented by David Enoch and Erik J. Wielenberg, revealing why this strategy is unsuccessful. More importantly, this discussion exposes the problems that both sides of the debate face. First, both sides rely on theoretical assumptions that are independent from the evolutionary explanations in these arguments. This shows that the evolutionary explanations are not, in fact, doing much (if anything at all) in helping resolve the metaethical debate. Second, both sides rely on inference to the best explanation when employing evolutionary explanations to support their metaethical view. This is problematic since the issue being debated is normative, and what counts as "best" is a normative question. Third, the problem of "inferential opaqueness" arises; that is, the precise connection between the evolutionary explanations (the premises) and the rejection (or acceptance) of moral realism (the conclusion) at stake in the debunking literature remains unclear. This opacity, I argue, is intrinsic to evolutionary explanations and cannot be rectified.

Finally, Chapter Five addresses the genetic fallacy. Discussion of this fallacy is of fundamental importance, since debunking arguments, and their realist counterparts, are genetic arguments relying on claims about the origins of moral beliefs to either undermine or support those beliefs. Using Kevin C. Klement's analysis of genetic arguments, a discussion of the three primary debunking arguments of Ruse, Street, and Joyce follows. The resulting analysis reveals that Joyce's modest form of debunking is the only form that avoids the genetic fallacy. However, further scrutiny of Joyce's modest debunking shows that other problems remain. Building on the recent work of Katia Vavova, I show that the skeptical challenge Joyce claims realists must face would in fact apply equally to antirealist views, but only if evolutionary explanations actually play a role in the debate. It turns out, however, that evolutionary explanations are not essential to debunking arguments. Here, I consider both the dispensability of evolutionary explanations and the indeterminacy of moral concepts. In the debunking debate, evolutionary explanations not only end up being dispensable but these explanations also underdetermine the moral concepts essential to metaethical discussion. As a result, evolutionary explanations do not help resolve the dispute between evolutionary debunkers and Darwinian Moral Realists.

At the end of Chapter Five, I conclude that the use of evolutionary explanations to either undermine or support metaethical positions is a fruitless endeavor. Instead of providing an explanation of morality compatible with naturalism and free from any theological or ontological baggage, it produces both an unsatisfactory answer to the question of where morality comes from and fails to justify any metaethical positions, positive or negative. Evolutionary metaethics, therefore, does not move the debate between moral realists and moral antirealists any closer to a resolution.

Chapter One: The Origins of Morality

Evolutionary ethics is a naturalistic approach to ethics that connects the natural sciences to moral theory through evolutionary biology and evolutionary psychology. Surveying the history of evolutionary ethics, from Darwin to the mid-1990s, Paul Farber describes evolutionary ethics as a series of failed "attempts to understand human nature from an evolutionary perspective [and to] use evolution as a foundation of ethics" (1994: 6). Although this historical overview reveals a variety of arguments launched against evolutionary ethics, a central problem emerges common to all attempts. Evolutionary ethics is often guilty of reading the social values of the time into nature rather than deriving those values from nature. This problem is exacerbated by the conjectural histories associated with evolution, which are then used to establish particular ethical views.[1] Despite this unsuccessful past, evolutionary ethics remains fairly popular in contemporary moral philosophy. Farber sees little promise in these new theories since "most contemporary arguments are simply recycled versions of earlier ones" (1994: 2).

Although many elements of contemporary evolutionary ethical approaches mirror the past, I suggest that there are some arguments within these views that appear novel and thus deserve special attention. Unlike most previous evolutionary ethical views that seek to establish prescriptive normative conclusions, the current debate in evolutionary ethics chiefly focuses on attacking the metaethical foundations of ethics. It is these debates, regarding the metaethical implications of evolution for ethical theory, that have not yet been fully explored.

---

[1] The term "conjectural history" was coined by Dugald Stewart to identify a type of history that developed to replace conventional history when the conventional method was lacking or insufficient. He says, "in this want of direct evidence, we are under a necessity of supplying the place of fact by conjecture; and when we are unable to ascertain how men have actually conducted themselves upon particular occasions, of considering in what manner they are likely to have proceeded, from the principles of their nature, and the circumstances of their external situation." See Stewart (1863: 18).

In this essay, I will focus on the Neo-Darwinian views that claim that humans have evolved an innate moral sense. I call such views Darwinian Moral Nativism. I will discuss the nature of the moral sense and its basic features as presented by these views, paying special attention to the various metaethical claims these positions imply. Next, I will argue that Darwinian Moral Nativism fails to establish these metaethical claims since they rely on a questionable method of justification—namely the genealogy of morals, which is an account of the origins and historical development of morals—to establish the foundations of morality. Although some aspects of Darwinian Moral Nativism may prove to be philosophically useful, I conclude that the overall approach is problematic. Lastly, I will argue that the limitations of Darwinian Moral Nativism reflect the limitations inherent in evolutionary ethics more broadly construed.

I start by looking at the philosophical methods employed in evolutionary ethics as well as the views that influenced these contemporary views. Such discussion helps explain what the views are, how they are similar to past projects in evolutionary ethics, but also how they are different and thus require new analysis. This analysis begins by looking at explanation and justification more broadly, the role that explanation and justification play in these theories of evolutionary ethics, and the underlying question that drives this debate: Where does morality come from?

A. The Role of Explanation and Justification

An explanation is a set of statements, the *explanans*, that provide clarification of some phenomenon or thing, the *explanandum*.[2] The *explanans* provides the "why" and "how" story

[2] Here, I am not concerned with any particular theory of explanation (as in the debate in philosophy of science) or types of explanation (causal, cognitive, linguistic, etc.) that may be up for debate. Rather, I want to focus on the role that explanations play in any given theory, or story,

that underlies the *explanandum*, thereby elucidating the thing or phenomenon in question. The goal of an explanation then is simply to help understand the *explanandum* more clearly. The following are some examples of explanations:

(1) Emma likes strawberries because they taste sweet.
(2) Strawberries are sweet because they contain natural sugars.

In both cases the statement includes an *explanandum*—"Emma likes strawberries" and "Strawberries are sweet"—that is further elucidated by the *explanans*—"because they are sweet" and "because they contain natural sugars". From these two explanations (statements) we can understand *why* Emma likes strawberries and *why* or *how* strawberries are sweet.

Whether or not an explanation is satisfactory depends on a number of things, most of which goes well beyond the scope of this rudimentary discussion.[3] However, the success of any explanation will often depend on the initial inquiry. That is, what were you trying to understand about the phenomenon and does the explanation provide that understanding? In example (2) above, if the initial inquiry were simply "Why are strawberries sweet?" then the explanation was likely satisfactory. However, if the inquiry were different, such as "What, at the molecular level, makes a strawberry sweet?" then the explanation would not be satisfactory. This example illustrates several important features of explanations. First, there are various levels of explanation for the phenomenon. Second, the success of an explanation depends on the level of explanation being sought. Third, the level of explanation being sought depends on the initial inquiry and the

---

or as an answer to an inquiry more generally, as well as the distinction between explanation and justification.

[3] This is a broad issue within the philosophy of science that goes beyond the scope of this discussion. For discussions on the criteria for satisfactory explanations see: Pitt (1988), Salmon (1989), and Ruben (1993).

goal of that inquiry. Therefore, although the general goal of an explanation is to understand the *explanandum*, there are also levels of explanation based on the varied goals of an explanation.

Perhaps all of this seems obvious; however, things get tricky when you consider the relationship between explanation and justification. Explanations and justifications are different because they have distinct purposes. Unlike explanations, justifications provide the "should" or "ought" story behind a phenomenon rather than the "why" and "how". Moreover, the goal of a justification differs from the goal of an explanation. Justifications aim to show why something is right or correct, whereas explanations aim to provide clarification and understanding. As such, a justification is also a defense and implies an argument while an explanation does not.[4] The following in an example of a justification:

(3) Emma should eat strawberries because natural sugars are healthy.

The difference between (1), (2), and (3) is that the statements in (1) and (2) merely aim to shed light on some phenomena, *why* Emma likes strawberries and *why* strawberries are sweet, while the statement in (3) provides an argument about why it is *right* (or that there are good reasons) for Emma to eat those strawberries. There are also criteria for determining whether or not a justification is satisfactory, although this goes beyond the scope of this discussion.[5]

Although explanations and justifications are clearly different because they serve distinct purposes, there are ways in which they can overlap—sometimes an explanation can be part of a justification. Consider the previous statements again:

(1) Emma likes strawberries because they taste sweet.
(2) Strawberries are sweet because they contain natural sugars.

[4] That is, of course, excluding arguments about which explanation is correct; my point is that the explanation itself is still not an argument per se.
[5] Here, I am concerned with justification more broadly and not simply as an issue of epistemic justification, as in the debate between foundationalism, coherentism, and reliabilism. For a discussion on these issues see: BonJour and Sosa (2003).

(3) Emma should eat strawberries because natural sugars are healthy.

The explanatory statements in (1) and (2) now appear to lend support to the justificatory statement in (3); however, this still depends on the goal of the initial inquiry. If the initial inquiry is only explanatory, then the statements (1) and (2) will simply remain explanations. If the initial inquiry is justificatory, then the explanatory statements become part of a justification.

Even though the statements (1) and (2) can now work together as part of a justification in this context (of a justificatory oriented inquiry) they are still only explanatory statements when taken alone. Explanation requires additional work—changing the statements themselves, the context of these statements, as well as other features—to become a justification. How, exactly, this transition occurs depends on what is being justified and whether or not the explanatory statements count as good reasons for that which is being justified (since what counts as a good reason also varies based on context). Regardless of what counts as a good reason, there is no automatic transition between an explanatory statement and a justificatory statement; you cannot automatically conflate the two types of statements since extra work needs to be done to turn an explanation (or explanatory reason) into a justification (or justificatory reason).

First, the explanation need to become part of an argument. Next, additional information is needed to show why these explanatory statements lend support to the justificatory statement. Without the extra work, this conflation of explanation to justification is merely circular reasoning. Take, for example, the Cartesian circle.[6] Descartes' argument runs something like this:

> (1) If God is benevolent (non-deceiving), then when I perceive *x* clearly and distinctly, then *x* is true.

[6] Descartes (1641/1996): *Meditation 3*. I'm putting aside any debate over whether or not this is the correct interpretation of Descartes. For a debate on the Cartesian Circle, see: Cottingham (1991), Doney (1955), and Gewirth (1970). For Descartes' response to the charges of circularity, see Descartes (1641/1996): *Fourth Replies*.

(2) I clearly and distinctly perceive a benevolent God exists.

(3) Therefore, a benevolent God exists.

(4) If a benevolent God exists, then when I perceive *x* clearly and distinctly, then *x* is true.

(5) Therefore, if I perceive *x* clearly and distinctly, then *x* is true.

Descartes' reasoning is circular because the explanation of how the meditator knows she is not being deceived—that she perceives something clearly and distinctly—acts simultaneously as the justification for why she is not being deceived (as well as why God exists). In other words, the explanation becomes conflated with the justification. Without some further premises—that is, without extra work—this this remains a viciously circular argument.

In evolutionary ethics, evolutionary explanations explain *why* and *how* humans have become the moral creatures that we are today. These explanations aim to help us understand morality at various levels, such as why humans cooperate with others, how altruism developed, and whether or not we possess an innate moral sense. As previously mentioned, part of the success of any explanation depends on the initial inquiry. In evolutionary ethics, the initial question that drives the evolutionary approach is: where does morality come from?[7] An answer to this question requires an explanation but many theories also use these evolutionary explanations as part of their justification for particular moral views at both the prescriptive and metaethical levels. These evolutionary explanations no longer simply explain *why* and *how* morality evolved, but rather seek to *justify* particular moral theories—which moral theory (both prescriptive and metaethical) we *ought* to follow (and even how we *ought* to behave!). It is this transition from explanation to justification to prescription that requires further analysis—an

[7] The term *morality* is, of course, vague. It could refer to various aspects of morality such as moral concepts, moral rules, or our ability for morality. Positions in evolutionary ethics offer different uses of the term *morality* so I will address these variations in the forthcoming chapters.

analysis that will reveal the illegitimate nature of such a transition. In other words, in evolutionary ethics, the explanations and justifications are conflated without the necessary work. The following analysis of the initial inquiry that drives evolutionary ethics—where does morality come from—helps make this issue clear.

B. Where Does Morality Come From?

Within contemporary evolutionary ethics, evolution typically refers to Neo-Darwinian theories of evolution that combine the Darwinian mechanism of natural selection with modern genetics.[8] Evolutionary ethics, however, is not a unified theory. Just as there are numerous types of issues addressed within the field of ethics and various methods used to handle such questions, evolutionary ethics covers a wide range of positions.[9] What these diverse views share is the claim that evolutionary theory helps explain the origins of morality, which is used as an argument to then establish the foundations of morality. These evolutionary foundations then shape the meta-ethical and/or prescriptive claims within each of these positions.[10] As such, evolutionary ethics attempts to answer the following question: where does morality come from?

The question of where morality comes from is ambiguous and can be understood in two different ways—as an inquiry into the origins of morality or as a search for the foundations of morality. And although these two ways of understanding where morality comes from may appear similar, each present a distinct approach with distinct aims. As an inquiry into the origins of morality, an explanation (or set of explanations) answers the question of where morality comes

[8] For a discussion of the development of Neo-Darwinism see Ruse (2009: 30-45).
[9] For a discussion, see Fitzpatrick (2014), Richards (1986: 265-93), and James (2010).
[10] Note that not all positions make both metaethical and prescriptive normative claims, while other views claim to be purely descriptive. For example, much of the work searching for an evolutionary explanation of altruism is purely descriptive, such as Hamilton (1964), Trivers (1971), and Axelrod (1984). Other philosophers often support their normative (metaethical and prescriptive) views with this type of descriptive work.

from, whereas a search for the foundations of morality requires a justification (or set of justifications).

To help clarify these differences, it is useful to consider other philosophers (or philosophical views) that employ these types of approaches. In the discussion that follows, the aim is not to provide a complete analysis of these philosophers' views or defend any particular view. Instead, it merely provides a rough categorization of these views as a way to illustrate the distinct approaches to answering the initial inquiry about the origins of morality. As such, I am less concerned with defending any particular interpretation of these philosophers and more concerned with how mapping these approaches onto their views helps clarify their approaches to answering the initial question that drives evolutionary ethics.

An inquiry into the origins of morality employs the genealogical method to examine current moral values in terms of their historical development[11] According to this approach, the developmental history or story behind the emergence of a moral belief or moral value provides an understanding of why we have the moral beliefs that we have. How we have come to possess the moral values that we possess then plays a significant role in either supporting those moral beliefs/values or undermining them. In other words, the history of a moral belief/value is thought to be relevant to the legitimacy of that belief/value in this approach.

Many philosophers have used this method of inquiry; for example, Thomas Hobbes and Jean-Jacque Rousseau both consider current moral values in terms of a historical process to help shed light on such values. They both examine man's original state of nature and his subsequent rise, or fall, from such a state to better understand both our moral and practical commitments. Hobbes seeks to use this genealogy of man (albeit a hypothetical, imagined history rather than a

[11] See Elbe (2001: 260-61) and Owen (1997: 163).

real one) to support certain values and political structures, whereas Rousseau aims to undermine such values by revealing their history.[12] This approach, or the genealogical method, is also similar to what Henry Sidgwick referred to as the 'historical method', which he describes broadly as "any study of the past pursued with a view to the explanation of the present" (1886: 211).[13]

The aims of the genealogical method, therefore, vary and have been used to either justify or undermine current values. But using the genealogical method to directly justify a value, institution, or practice typically commits the genetic fallacy, which is the logical mistake of accepting (or rejecting) a claim based solely on its origins.[14] In other words, this version of the genealogical method is fallacious because the genealogical explanations become conflated with justifications to support or reject current moral values.

The most common use of a genealogy is to reflect critically on, critique, and undermine current values, practices, or institutions rather than to legitimize them (Geuss 1994: 286-87). Although this critical use of genealogy is not immune from the genetic fallacy, there are formulations that are *prima facie* immune; that is, they at least appear to be immune. These formulations deserve special attention, especially since this form of the genealogical method is most commonly used in the contemporary evolutionary ethical views that focus on metaethical issues.[15] As such, the genealogical method can be divided into two types: the *deductive genealogical method,* which typically commits the genetic fallacy, and the *critical genealogical method*, which appears to evade the genetic fallacy. Let's turn to the critical genealogy, namely Friedrich Nietzsche's work on morality.

---

[12] See Korsgaard (2010), Hobbes (1651/2009), and Rousseau (1761/1923).
[13] Note that Sidgwick is arguing against this method, but that is something I will discuss below.
[14] I will discuss the genetic fallacy in further detail in Chapter Five.
[15] That is, in conjunction with the *theoretical approach* discussed below.

A central aim of Nietzsche's moral writings was to provide an account of the history (qua origins) of morals in order to critique the moral values of his time.[16] Nietzsche focuses on the particular influence Christianity has had in shaping existing moral values. Here, the critical genealogical method is what Stefan Elbe calls 'episodical'—such genealogies focus on a particular set of events rather than all of history (2001: 262-63). Nietzsche then uses an episodical history to critique existing moral values. For example, when discussing religious beliefs in *Human, All too Human,* Nietzsche points out how the origins of a belief can affect one's belief in that belief:

> …how [beliefs in God] originated can at the present stage of comparative ethnology no longer admit of doubt; and with the insight into this origination the belief falls away. (1878/1986: 73)

In his following work, *On the Genealogy of Morals,* Nietzsche aims to reveal the origins of the current moral values as a way to question the beliefs in those values—if existing moral values have 'ignoble origins', then we should revise those moral beliefs. Nietzsche lays out this moral project in the forward to *On the Genealogy of Morals* by stating that:

> …[W]e need a *critique* of moral values, *the values of these values should itself, for once, be examined*—and so we need to know about the conditions and circumstances under which the values grew up, developed and changed…since we have neither had this knowledge up till now or even desired it. (1994: 8)

Here, Nietzsche explains how he aims to question the moral values based upon the origins of those values. Nietzsche is not questioning whether or not any such values exist but rather whether or not our beliefs in existing values are correct.[17] As such, he is providing a normative

[16] See Nietzsche (1887/1994) and (1882/ 2001). For a relevant discussion on Nietzsche see Prinz (2009: 215-43).

[17] Despite what some may think because of his alleged nihilism, Nietzsche does not deny that people hold the values that they do, and as such these values exist in this sense.

critique of existing moral values by suggesting that once we know *why* we believe such values we may have reasons for rejecting them.

In a discussion about the usefulness of a genealogy of morals, Jesse Prinz identifies Nietzsche to be working under the following thesis:

> If a belief in a moral value emerged as a consequence of ignoble historical events, then that belief should be abandoned. (2009b: 237)

According to this form of the critical genealogical method, the history and circumstances under which such values have formed—the origins of a moral belief—can therefore undermine one's belief in such moral values.[18] It is important to point out that this thesis, as employed by Nietzsche, is not an epistemological or metaphysical thesis—that is, it is not *directly* challenging the warrant of moral beliefs, justification of such beliefs, or the truth (existence) of the existing values. As Raymond Geuss points out, "the purpose and effect of a genealogy can't be to criticize values or valuation directly [but] a genealogy can undermine various beliefs about the origins of different forms of valuation" (1994: 286-87).[19] And if the aim were to undermine these values directly, then the genetic fallacy would come into play. As such, Prinz presents a milder version of Nietzsche's thesis that is less subject to these types of criticisms:

> If a belief in a moral value emerged as a consequence of ignoble historical events, then we should *consider* abandoning that belief. (2009b: 239)

Although this form of the critical genealogical method still serves as a critical tool, it is the safest form of this method in that it does not commit the genetic fallacy. In this version, the genealogical explanation is not conflated with a justification to either support or reject current

[18] Note that this form of the genealogical method also implies that there is some "objective moral standpoint" from which we can judge these histories, and thus identify certain values as deriving from 'ignoble' origins (which is the reason for abandoning them). See Prinz (2009: 237). For a debate on this issue/interpretation see Leiter (2000).
[19] Here, Geuss references Nietzsche's *The Gay Science*, §345 and *The Will to Power*, § 254.

moral values. Rather than rejecting a belief solely on its origins, this thesis merely raises a red flag on such beliefs and requires further investigation to prove that the belief should be abandoned.[20]

To summarize, the critical genealogical method, such as that used by Nietzsche, contains the following elements as described by Elbe: First, it is *historical*; it considers the history of present or existing moral value. Second, it is *critical*; it is used to reflect critically on present or existing moral values. And third, it is *episodical*; it focuses on specific historical events relevant to the present or existing moral values in question as opposed to history as a whole (2001: 260-263). And, in the form applied by Nietzsche, it also operates at the normative level, without *directly* addressing epistemological or metaphysical issues (e.g., warrant, justification, truth, existence).[21]

Going back to the original question, "where does morality come from?" the genealogical method seeks an answer through a historical process that is used to critique those values based on its' origins. I will therefore call this first way of interpreting where morality comes from the *genealogical approach*. And, as noted above, there are two types of genealogical approaches, the *deductive genealogical approach* and the *critical genealogical approach.*[22] Since the first type of genealogical approach is obviously fallacious and the second type is most commonly used in the contemporary evolutionary ethical views that focus on metaethical issues, I will focus my attention on the critical genealogical approach.

---

[20] It is important to note, however, that Nietzsche provides other arguments in addition to the genealogy of morals, to try to undermine such values along with a positive moral view. For a discussion of these arguments, see Leiter (2002).

[21] Perhaps one may argue that these normative claims have metaethical implications, but that would only be indirectly and, as such, are not the immediate aim of such critiques. And, they are often additional projects, built up using a different methodology.

[22] Deductive attempts to directly support or reject a particular view of morality based on where morality comes from.

The second way of understanding where morality comes from considers the metaphysical basis of morality. This approach aims to discover the source of moral obligations by establishing where (e.g., in the mind, in nature, in society, etc.) morality is grounded. Such foundations help to clarify the nature of our actual moral obligations and can be used to legitimize (or undermine) current values (and not just our beliefs about them) directly. Searching for such origins, i.e., for the foundation from which moral obligations are derived, however, is not a historical account. In this second approach, an account of the origins of morality, which is the foundation of morality, *justifies* particular moral obligations (or moral values). This approach does not merely provide an *explanation* as to where such obligations came from or why we have come to believe certain moral values (e.g., a moral genealogy), but rather how (or if) such values are *justified*. As such, searching for the foundations of morals usually aims at justifying a set of moral values rather than seeking to undermine current values; it is typically not a critical method.

Immanuel Kant's moral philosophy provides an example of this type of approach—he aims to establish a foundational principle (the categorical imperative) to ground morality (in pure reason) from which he derives normative theory.[23] For Kant, discovering the foundation of morals cannot be a historical process since it concerns the rational part of the metaphysics of morals (the moral law) rather than the empirical part (a practical anthropology) that would be historical, as described in the previous method. As he points out in the preface to *Grounding for the Metaphysics of Morals,*

> …the ground of obligation here must therefore be sought not in the nature of man nor in the circumstances of the world in which man is placed, but must be sought *a priori* solely in the concepts of pure reason. (1785/1993: 2)

[23] Kant (1785/ 1993) and (1886).

In other words, any justification for moral values must not be subject to empirical considerations that are contingent, but logical (or conceptual) considerations that are necessary. Here, Kant is seeking to justify moral values through the foundations of, or origins in, the moral law.

In seeking a metaphysical foundation of morality, questions related to both epistemological or metaphysical issues are considered—that is, the warrant of moral beliefs, justification of such beliefs, the truth of the existing values as well as their ontological status, are the central focus of this approach. Consequently, this approach handles metaethical questions directly, and since it is an ahistorical, non-empirical approach, I will call this second way of interpreting where morality comes from the *theoretical approach*.

The critical genealogical approach and the theoretical approach not only have different aims—the critical genealogical approach aims to critique existing moral values whereas the theoretical approach aims to justify moral values (although not necessarily existing ones)—but they may also be seen as opposing views about how to do philosophical ethics. Nietzsche, for example, saw his genealogical approach as a rejection of Kantian, *a priori*, theoretical approaches dominant during his time.[24]

Elliot Sober makes a distinction between two types of ethical projects that further illustrate some of the differences between the critical genealogical approach and the theoretical approach. According to Sober, exploring (1) "why people have the ethical thoughts and feelings that they do" and (2) "the problem of deciding what the status of those thoughts and feelings is"

[24] Nietzsche's position is the antithesis of Kant's position—in fact Nietzsche's view is a rejection of metaphysics, and in particular a rejection of Kant. See Nietzsche (1887/1994: 5). For Nietzsche there are no transcendent features of humanity (except perhaps, the will to power, but this is debatable). For a discussion of Nietzsche's rejection of metaphysical philosophy, see Couzens Hoy (1983). Here, Couzens Hoy describes Nietzsche's moral project as non-metaphysical. He says, "…Nietzsche will use genealogy to destroy metaphysics altogether. Genealogy itself becomes a way to do non-metaphysical philosophy" (252-53). For an alternative reading of this opposition, see Bailey (2013).

are distinct tasks that are not necessarily connected (1994: 93-94). The former issue is a matter of explanation while the latter is a matter of justification.[25] Inquiring into the origins of morality from the critical genealogical approach lines up with the explanatory project of (1), while asking the same question from the theoretical approach is similar to the justificatory project in (2). In the critical genealogical approach, the history of morals may *explain* why people have come to hold their current ethical beliefs. This historical explanation may also cause us to question those beliefs, but it does not automatically justify or undermine the status of those beliefs. The theoretical approach, however, can *justify* the status of those moral values by establishing their metaphysical foundations.[26]

As previously discussed, explanations and justifications are different because they serve a distinct purpose—there is no automatic transition between an explanatory statement and a justificatory statement. Despite these differences, in evolutionary ethics these two approaches—the critical genealogical approach and the theoretical approach—collapse into one: evolutionary theory provides a genealogy of morals that determines the metaphysical foundations of morality. In other words, evolutionary explanations serve as justifications for particular normative positions.[27] In his essay "Theory of Evolution and Its Application to Practice", Henry Sidgwick

[25] Sober says that there is no automatic connection between (1) and (2), that is, between explanation and justification.

[26] Prinz makes a similar distinction when discussing Nietzsche's genealogical method in *The Emotional Construction of Morals.* Prinz says: "When we view morals historically, we put aside questions of justification and engage in an auto-anthropology. We do not ask 'Why should I believe p?', but rather, 'Why do I believe P?'" (2009: 242).

[27] I think this is not just a simple fact/value distinction or the collapse of it (in the case of evolutionary ethics). The main issue here is a collapse of the distinction between explanation and justification, and exploring the question of when an explanation becomes a justification. Many may think they are equivalent, but I have already illustrated that such a conflation is a mistake, although one may transform into the other. Thus, this makes it more than just a simple matter of something like Moore's naturalistic fallacy. And although it may raise the more general worry

refers to evolutionary theory as the "final extension of the historical method" (1876: 55); for when applying the 'historical method' to the field of ethics, the aim is no longer merely an explanation of the present, but rather "a method for determining-what, after all, is the most interesting question with regard to any class of human beliefs—*viz.*, how far they are true or false" (1886: 12).

Some theorists use this combined approach to either lend support to morality, justifying moral realism, or to undermine morality, which leads to moral skepticism or antirealism.[28] But can a genealogy of morals, based in evolutionary theory, determine the foundations of morality? In other words, can evolutionary explanations function as justification for a particular metaethical view? Can the genealogical and theoretical approaches legitimately collapse into one? The short answer to these questions is *no*, and the long answer will be the subject of this essay.

---

about how or whether ethics can be naturalized, that is not the target of my argument. I mean only to question the particular case of evolution.

[28] For example, Sober points out that genetic arguments in evolutionary ethics often aim to prove ethical subjectivism from the fact that our ethical beliefs come from an evolutionary process. See Sober (1994: 104).

Chapter Two: Darwinian Moral Nativism

To better understand how the critical genealogical and theoretical approaches collapse into one approach in evolutionary ethics, this chapter will explain Darwinian Moral Nativism, the historical views that influenced these contemporary positions, and the traditional criticisms of such views. This discussion will help set the stage for the next chapter, which will focus more closely on a few of the prominent views of Darwinian Moral Nativism, the debunking arguments of Michael Ruse, Richard Joyce, and Sharon Street, all of whom use the evolution of an innate moral sense to support particular metaethical positions.

A. A Combined Approach: Darwinian Moral Nativism

In evolutionary ethics, the combined approach often appears as a form of moral nativism, which is the position that morality is an innate, rather than an acquired, trait. What it means for morality to be innate varies across views, from folk notions of morality being "inborn" and "unlearned" to essentialist views where morality is a trait found in all normal members of a species.[29] The evolutionary conception of moral nativism, however, explains the innateness of morality primarily in terms of natural selection. According to this Darwinian view, morality is an adaptation—morality is a trait that was selected for its reproductive advantage.[30] As such, I will refer to this view as Darwinian Moral Nativism. As Richard Joyce describes it, the Darwinian Moral Nativist endorses the hypothesis that:

[29] See Griffiths (2002), Mameli & Bateson (2007), and Mameli (2008) cited in Joyce (2013b) for a full discussion of these views. Moral nativism, however, is by no means a new or contemporary view—in fact, the view that morality is innate has a robust Platonic history. See Plato's Theory of Recollection, which can be found in the *Phaedo* 72e-78b. Also see Nichols (2005:1).

[30] Richard Joyce defines an adaptation as "traits that emerge and persist in a population through a process of selection: because they are transmitted from parents to offspring and provide the bearers with reproductive advantage" (2013a: 4).

> …a faculty for making moral judgments is a biological adaptation that emerged because this way of thinking provided our ancestors with some sort of reproductive advantage over their competition. (2013a: 1)

Darwinian Moral Nativist views also vary depending on which trait is 'selected for' since it is debatable what, exactly, a moral trait *is*.[31] Joyce points out that on the one hand, an entire faculty or mechanism, such as a moral sense, could be 'selected for.' According to these views, the trait selected for provides a "toolkit" for morality rather than any particular moral principles or judgments. This "toolkit" may supply general moral concepts (e.g., desert, wrongness, etc.) that appear to be required for making moral judgments. On the other hand, some argue that complete non-particular moral judgments (e.g., "killing an innocent person is wrong") are 'selected for' rather than a mechanism (Joyce 2013b: 8-9).

This characterization of Darwinian Moral Nativist views, however, does not capture most actual views, which tend to fall somewhere between these two extremes.[32] Shaun Nichols provides a classification of moral nativist views that seem to capture these in-between positions, although he is not concerned with evolutionary (adaptation and natural selection) views per se. According to Nichols, there are three major types of moral nativist views: Rule Nativism, Moral Principle Nativism, and Moral Judgment Nativism. Rule Nativism claims there is an innate mechanism that has the capacity for "non hypothetical-imperative rule comprehension", which is needed to make moral judgments (Nichols 2005: 3-4). The work of Denise Cummins (1996:24) and Chandra Sripada and Stephan Stich (2007) represent this type of approach. Moral Principle Nativism argues for a set of innate universal moral principles (such as "do not rape" and "do not kill innocent people"). This view is similar to Noam Chomsky's universal grammar and includes

[31] The mechanism for selection is also debatable, but that is outside the scope of this discussion. See Fodor and Piattelli-Palmarini (2010), which challenges the notion of selection and 'selected for'.
[32] See Joyce (2013b).

the work of Gilbert Harman (1999: 107-15), John Mikhail (2002: 54), and Stephan Stich (1993: 215-28). The last position, Moral Judgment Nativism, claims that an innate appreciation of the unique status of morality and the difference between moral and conventional claims helps guide moral judgments. This approach includes the work of Susan Dwyer (1999: 177-71).

Regardless of how to classify these nativist views, it is obvious that there are many ways to be a Darwinian Moral Nativist.[33] But these variations are not merely scientific in nature. That is, disagreement over the underlying scientific facts and explanations (e.g., debates over selection mechanisms) do not fully account for these differences. Rather, much of the variation is tied to conceptual debates over what morality is, which traits represent morality, and thus, which traits are 'selected for'.[34] Consider, for example, the human eye. The eye is a tangible part of the human body—that is, it is not a psychological or behavioral trait like altruism or reason. As such, it is easier to determine how the eye functions and which traits (e.g., light sensitive skin) were selected for to develop the human eye.[35] Morality, however, is a much trickier case. Unlike the eye, it is not an obviously tangible part of the human body. If there is a biological component to morality, it is likely a psychological or behavioral trait. So what trait (or set of traits), are we looking for with morality? Altruism? Cooperation? Pro-social tendencies? Reason? The list of

[33] For example, one particular area where these views differ is in regards to moral universals. In Jesse Prinz's article "Against Moral Nativism", he describes three different positions on universals that cover most evolutionary conceptions of moral nativism—the minimal view, the modest view, and the immodest view. According to Prinz, the minimal view argues that "we have a general capacity for acquiring moral norms but the content of those norms is pre-ordained". The modest view says that "there are innate moral rules but epigenetic factors exert a nontrivial influence on how these rules operate in different societies". And the immodest view claims that "innate moral norms have specific content (the actual norms that govern our lives are innately fixed and culture exerts little influence" (2009a: 2).

[34] For a detailed discussion, see Joyce (2013b).

[35] This is not to deny that there are still debates over how the human eye developed and that this is also a tricky business. For a discussion regarding complexities involved in understanding the evolution of the human eye, see Lamb (2011).

traits seems not only long but also conceptually ambiguous since it all depends on what morality is. Thus, an analysis of Darwinian Moral Nativism must also consider these conceptual debates, which cannot be resolved directly by evolutionary theory. This point will become crucial to understanding and defining the limitations of evolutionary ethics later.

Despite these differences (and difficulties), Darwinian Moral Nativism claims that morality is an evolved capacity that has an evolutionary genealogy. Many of the evolutionary Moral Nativist views use the genealogy of this evolved capacity either to lend support to morality (justifying moral realism) or to undermine morality (supporting moral skepticism or antirealism). In these cases, it becomes clear how the combined approach is supposed to work—the evolutionary theory provides a genealogy of morals (an explanation) that determines the metaphysical foundations of morality (a justification).[36]

Although Darwinian Moral Nativism is a contemporary position, it has deep roots in Charles Darwin's work on evolutionary theory and his writings on morality in *The Descent of Man*. Contemporary Darwinian Moral Nativism is also influenced by Herbert Spencer's work on morality, albeit more indirectly. Both Darwin and Spencer proposed an evolved moral capacity, which they both referred to as a moral sense. And, for both Darwin and Spencer, the evolution of the moral sense was proof that morality, or at least the capacity for morality, was an innate trait. Understanding what Darwin and Spencer thought about the moral sense helps shed some light on these contemporary views of Darwinian Moral Nativism, as well as illustrating how evolutionary ethics brings together the genealogical and theoretical approaches.

[36] For example, Sober points out that genetic arguments in evolutionary ethics often aim to prove either ethical realism, conventionalism or subjectivism from the fact that our ethical beliefs come from an evolutionary process (1994:100-02).

i. Darwin's Moral Sense

In *The Descent of Man,* Darwin explored moral questions from what he thought was a unique perspective—he looked at morality solely in terms of natural history to uncover the genesis of man's moral sense.[37] Since man was on the same natural continuum as other animals, the study of the physical natural world could be used to understand morality. The nature and origins of man's moral sense—a natural evolved mechanism—was key to Darwin's biological views about ethics. Man's moral sense, according to Darwin, "tells us what we ought to do", providing us with a guide to action, while the conscience "reproves us if we disobey it" (1871: 89). Thus, the moral sense can explain why, in certain cases, we have a sense of duty and even why we perform altruistic acts, which Darwin describes as self-sacrificial actions performed for the good of others.[38] This explanation of the evolution of morality in terms of the development of the moral sense was influenced by the sentimentalist tradition of moral philosophy (e.g., Francis Hutcheson, David Hume and Adam Smith) since Darwin was not only familiar with such views but also presents a view of morality that can be classified as a sentimentalist position.[39]

For Darwin, a biological explanation of the distinction between man and other animals required the discovery of the origins of man's moral sense since it was this feature (morality) that truly distinguished man from the "lower" animals.[40] By studying "lower animals", Darwin tried to show that the moral sense arose in the process of evolution from a combination of social

---

[37] In response to Kant's question about where moral obligations originate in *The Metaphysics of Ethics* (1836: 136) Darwin says "this great question has been discussed by many writers of consummate ability; and my sole excuse for touching on it is the impossibility of here passing it over, and because, *as far as I know, no one has approached it exclusively from the side of natural history"* (Darwin 1871: 68) emphasis added.

[38] See Darwin (1871: 67, 159).

[39] See Korsgaard (2010: 7, 11). Also see Darwin, *Metaphysics, Materialism, and the Evolution of the Mind: Early Writings of Charles Darwin,* Barrett and Gruber (1980).

[40] See Darwin (1871: 71-72) and Farber (1994: 49). Darwin thought this was part of the natural development for a social animal, any animal, not just man.

instincts, social feelings (both innate and instinctual), habit (acquired behaviors that appear unconscious but can be modified by will or reason), and developed intellectual capacities (e.g., the ability to reflect on our actions and to use language). Any species with these capacities would develop a moral sense, although it would differ in content between species.[41]

According to Darwin's view, social instincts such as feelings of sympathy, pleasure derived from the company of others, and the desire to aid others of close relation, all influence behavior. Like all instincts, social instincts are innate (unlearned) and do not require reason or judgment but they are also weaker than other instincts. However, social instincts are enduring and, thus, more persistent than other instincts.[42] The ever-persistent nature of social instincts plays a role in guiding action when combined with a developed intellect. Highly developed mental faculties furnish man with the ability to remember and reflect on past actions and motives. Feelings of satisfaction (or dissatisfaction) are then used to further assess whether such actions and motives were in line with the social instincts.

In addition to the ability to reflect on social instincts, man's developed mental faculties also supply the capacity for language, which facilitates the expression of thoughts. According to Darwin, language provides communities with the ability to express how they think its' members ought to act to promote the good of the community.[43] Since man is a social animal that inherited social traits like self-control and the tendency to be obedient to community leaders, the wishes and judgments of the community become an additional guide to action. Darwin, however, recognizes that following public opinion opens up room for human error; in some cases public

[41] See Darwin (1871: 69 and Chapter III).
[42] See Darwin (1871: 67, 68, 69) and (1861: Chapter VII). Some examples of social instincts include the maternal instinct (Darwin 1871: 83), the instinct to aid the helpless, and sympathy (in its original form) (Darwin 1871:162).
[43] Darwin (1871:81-83). Also note that Darwin argues that actions are no less moral if they come from an instinct (1871: 84).

opinion strengthens the social instinct to promote the community good while in others it does not. Even so, Darwin argues that men's sympathetic feelings and self-control lead men to the best actions, especially when social instincts and public opinion are reinforced in the individual through habit.[44] From this complex combination of social instincts, social feelings, developed mental faculties, and the ability to develop habits, the moral sense evolved in mankind. When men followed their moral sense it led to social cohesion from which mankind benefited as a species.[45]

By rooting moral action in a social instinct, Darwin sought to show that such actions were not purely self-interested but rather encompassed concern for the overall welfare of society, which also benefited the species and not necessarily the individual.[46] Here, Darwin was directly influenced by the work of his contemporary John Stuart Mill, considering the greater good of society as morally significant. Mill, however, thought that moral feelings were not innate, while Darwin thought that an acquired (rather than innate) moral sense was "on the general theory of evolution…at least extremely improbable" (1871: 68-69). Since social feelings and instincts are found in lower animals, Darwin saw no reason to assume it was any different for humans.[47] Even though it required a highly complex explanation, Darwin thought that the origins of the moral sense and sympathy could be explained through natural selection. Darwin, for example, accounts for the emergence of sympathy by arguing that: "communities with the greatest

---

[44] See Darwin (1871: 81-83). Darwin says "social instincts give the first impulse to act but his actions are largely determined by the opinion of the community" (1871: 83).
[45] See Allhoff (2003: 86) and Darwin (1871: 93). For Darwin's explanation of how this passes on from generation to generation, see Darwin (1871: 82-83).
[46] See Darwin (1871: 67) and Richards (1987: 218).
[47] See Darwin (1871: 68-69, footnote 5). If sympathy and other moral feelings were merely acquired traits, such as habits that are learned rather than instinctual traits that are innate, then they couldn't be fully explained by natural selection, the natural mechanism by which evolution occurs. Also See *Origin of the Species,* Chapter VII for a discussion of the difference between habits and instincts.

numbers of most sympathetic members would flourish best and rear the greatest number of offspring" (1871:79). Here, we see the preludes to the contemporary views of altruism, now explained via kin selection and reciprocal altruism, which dominate the current arguments found in evolutionary ethics.[48] Altruism plays a significant role in these contemporary naturalistic approaches to ethics since their explanations of the evolution of normative guidance rely on altruism to provide a bridge between the natural (biological and psychological explanations of altruism) and the moral.[49]

Darwin also disagreed with Mill's *Greatest Happiness Principle,* arguing that the general good or welfare of mankind was not happiness. Looking at the development of social instincts in lower animals and the development of the moral sense in humans, Darwin did not see how or why some philosophers (Mill included) would assume the general good to be equivalent to happiness (or to selfishness as other philosophers proposed).[50] Instead, natural history revealed that the general good, for both humans and lower animals, should be understood as "…the means by which the greatest possible numbers of individuals can be reared in full vigor and health, with all their faculties perfect under the conditions to which they are exposed" (Darwin 1871: 94). Thus, goodness for Darwin was an extension (or some version of) evolutionary fitness.

[48] For a discussion of altruism, see Okasha (2013).
[49] See Kitcher (2006). Also, see a discussion on the distinction between biological and psychological altruism and the normative in Fitzpatrick (2014). Here, also think of Ruse (1986) and Richards (1986). Note that these explanations often only explain the persistence of the trait, rather than the full origins of it.
[50] Here, Darwin references Lecky (1869) and a subsequent discussion of Lecky's work in the Westminster Review (1869: 240-44), which considers the Greatest Happiness Principle in relation to John Stuart Mill, Jeremy Bentham, Herbert Spencer, Alexander Bain and others. For a discussion, also see Farber (1994: 17).

a. The British Moralists Influence on Darwin's Moral Sense

It is easy to see how Darwin's description of the moral sense strongly resembles the British Moralists that preceded him. Although Darwin's view of man's moral sense provides an evolutionary explanation for the way it operates, and thus presents a unique approach to morality, his description is also rooted in both the sentimentalist and utilitarian traditions that preceded his work and dominated during that era. Darwin explicitly refers to the works of David Hume, Adam Smith, and John Stuart Mill, among others, when providing his account of the moral sense, and although he disagrees with some aspects of each of their work, he attempts to expand or improve upon them by embedding these moral theories within an evolutionary framework.[51] In fact, Darwin's view appears to merely reflect the popular moral view of his time, and his naturalistic approach only serves to reinforce and develop this view rather than to establish an entirely new view of morality that resulted from the study of a natural history on its own.[52]

Although this influence does not directly undermine Darwin's view of the moral sense, it is a significant issue to consider. If Darwin already held a general sentimentalist view that included an innate moral sense before providing an evolutionary explanation of morality, then there is a legitimate concern that his evolutionary explanation was developed to support what he already believed, rather than being the direct result of a natural history. And, if that is the case, then the evolutionary explanations provided deserve serious scrutiny. This concern ties back to the conceptual issues mentioned earlier—defining morality and which traits count as moral is inherently a non-scientific debate but rather a conceptual one. Again, these problems do not

[51] Darwin recorded the books he read in his notebooks (1838-52), (1838-58), and (1852-60). His lists included John Stuart Mill, Jeremy Bentham, Adam Smith, David Hume, and Alexander Bain. See also Darwin (1871).

[52] For a more general discussion on the influence of Hume on Darwin, see Huntley (1972). This particular claim that Darwin never establishes this position using natural history alone is my claim.

prove that these explanations are incorrect, only that one should proceed with caution before jumping on the Darwinian bandwagon.[53] And, as will become evident, the contemporary Darwinian Moral Nativist views are also greatly influenced by the same sentimentalist tradition.

In general, the sentimentalist position argues that morality is not from reason alone, but instead includes sentiments, feelings, emotions, and /or intuition. As a result, moral deliberation (decisions/judgments) comes from intuition rather than reason. Alexander Bain describes the doctrine of moral sentiment (or moral sense) as that in which "the human mind possesses an intuition or an instinct whereby we feel or discern at once the right from the wrong" (1872: 430). As such, the sentimentalist position opposed the moral rationalists of the same period, who argued that morality is from reason alone.[54] Some sentimentalist views, however, also argued that morality still requires a developed intellect, even if the sentiments play a primary role in morality.[55] Darwin's description of the moral sense includes both sentiments—the social instincts found in social animals, such as sympathy—and well-developed intellectual capacities, such as the ability to use language and reflect on past events.[56] There are many other parallels

[53] Darwin could simply be "cherry-picking" the explanations that fit what he already believed. This sort of problem is ever persistent, especially in moral psychology and with evolutionary explanations more generally and, at least, should raise a bit of suspicion.

[54] Some examples of the sentimentalists include Francis Hutcheson, David Hume, Anthony A.C. Shaftesbury, and Adam Smith. Some examples of moral rationalists include Ralph Cudworth, Samuel Clarke, and John Balguy. For a discussion see Gill (2006).

[55] See Gill (2010:115-17).

[56] Christine Korsgaard (2010) explains the parallel between Darwin's view and that of both Hume and Smith in her recent lecture "Some Reflections on Evolutionary Ethics". As just described, she points out that Darwin argues morality requires both intellectual faculties, such as memory and the awareness of others (what she calls a "theory of mind") as well as social instincts, which include feelings like sympathy and good will towards other men, that are felt by social animals. As such, Darwin is similar to Hume, who argues that man's sense of moral obligation is rooted in the sentiments of approval and disapproval of X, i.e., the sentiments of love and hate, respectively. Such sentiments require a developed intellect in order to be able to take an impartial perspective and to recognize other minds, the subject of approval and

between Darwin's view and those of the major of figures of the sentimentalist tradition such as Hutcheson, Hume, and Smith.[57]

Hutcheson, as well as Hume and Smith who followed him, explains morality in terms of moral approval and disapproval. The moral distinction between virtue and vice is understood as approval and disapproval, respectively, that come from the moral sense rather than reason. [58] For Hutcheson, the moral sense is a power of perception**,** an intuition or instinct that was implanted by God.[59] As Hutcheson describes it:

> The Author of nature has determined us to receive, by our external senses, pleasant or disagreeable ideas of objects, according as they are useful or hurtful to our bodies; and to receive from uniform objects the pleasures of beauty and harmony, to excite us to the pursuit of knowledge, and to reward us for it…in the same manner he has given us a MORAL SENSE, to direct our actions, and to give us still nobler pleasures: so that while we are only intending the good of others, we undesignedly promote our own private good. (1726/1991: 269)

The moral sense is, therefore, also an internal sense similar to the sense of beauty as opposed to an external sense like hearing or seeing.[60] Since the moral sense is an internal sense, moral approbation does not directly concern the truth or being reasonable.[61] The moral sense does not contain any innate ideas or knowledge either; in fact it does not require knowledge at all. Instead, the pleasure felt from virtue that leads to moral approval is what he calls a "reflex act".[62] There are no justifying reasons before the moral sense. Additionally, complex moral ideas, such as

---

disapproval (2010: 7-8). Her interpretation of Hume's view, however, is somewhat controversial, so I will not focus on her analysis in depth.

[57] Although Darwin never refers to Hutcheson's work directly (he only references Hume and Smith in his writings), he was well acquainted with Bain's "Mental and Moral Science", which discussed Hutcheson's views (among others) of the moral sense. See Bain (1872).

[58] See Slote (2010: 28).

[59] See Hutcheson (1738/1991: 269 and 263-65).

[60] Hutcheson develops his view of the moral sense from Shaftesbury's work. See Shaftesbury (1699/1991). For a discussion of internal senses, see Hutcheson (1726/2004: Treatise I: Section VII & Section VIII).

[61] See Hutcheson (1742/1991: 309).

[62] See Hutcheson (1738/1991: 269).

obligation and rights, can also be deduced from the moral sense, as opposed to reason or a set of moral laws.[63]

Hume has a similar explanation of morality. For Hume, moral approval also requires sentiments and moral distinctions between virtues and vice is not derived from reason either. Instead, moral distinctions come from the moral sense, and it is felt rather than judged—moral approbation is not a judgment but a feeling or sentiment.[64] We find virtue to be agreeable, and thus approve of it, whereas we feel vices to be disagreeable and thus disapprove of it. These sentiments are also universal, in that all mankind feels them. Hume argues that,

> The notion of morals implies some sentiment common to all mankind, which recommends the same object to general approbation, and makes every man, or most men, agree in the same opinion or decision concerning it. It also implies some sentiment, so universal and comprehensive as to extend to all mankind, and render the actions and conduct, even of the persons the most remote, an object of applause or censure, according as they agree or disagree with that rule of right which is established. (1751/1991: 78-83)

Morality, therefore, does not concern truth or falsity since reason deals with the discovery of truth or falsity and the sentiments or passion do not. In fact, for Hume "reason is, and ought only to be, the slave of the passions and can never pretend to any other office than to serve and obey them" (1738/1978: 415). The ultimate ends of human actions—virtue—can be explained by sentiments and affections alone, and there is no need for intellectual faculties such as reason. As Hume points out in *The Enquiry*: "It appears evident, that the ultimate ends of human actions, can never, in any case, be accounted for by *reason,* but recommend themselves entirely to the sentiments and affections of mankind, without any dependence on the intellectual faculties" (1751/1991: 89).

[63] See Hutcheson (1738/1991: 292-94).
[64] See Hume (1738/1978: 455-76) and (1751/1991: 87).

The similarities between Darwin's view of the moral sense and those of Hutcheson and Hume are already quite apparent. Like Hutcheson and Hume (and we will see later with Smith as well), Darwin frames morality in terms of approbation and disapprobation stemming from a moral sense. As previously discussed, for Darwin, the moral sense is a faculty that provides the ability to approve or disapprove of certain actions, which are reflected in certain moral sentiments. Thus reason by itself does not provide the basis for moral judgment for Darwin either. Darwin, however, differs from these two views in that he emphasizes the need for well-developed intellectual capacities, such as the ability to use language and reflect on past events. But even though Darwin emphasizes the need for a developed intellect, he does not go so far as to say that reason plays a role in making moral judgments that come from the moral sense.[65]

Both Hutcheson and Hume argue not only that moral good (or moral virtues) consists of sentiments naturally felt by the moral sense, but also that morality concerns the public good or what is beneficial and useful to society. In a discussion of the moral quality of actions, Hutcheson says "[T]hat Action is best, which procures the greatest Happiness for the greatest Numbers; and that, worst, which, in like manner, occasions Misery" (1738/1991: 284). This claim is clearly a precursor to the principle of utility (or the greatest happiness principle) as introduced by Bentham and later Mill. Hume similarly considers the usefulness of actions in achieving the end of moral virtue (particular morally good consequences), although such ends are determined by the moral sentiments.[66] For Hume, artificial virtues, which arise out of the conventions of society rather than nature (although influenced by the natural virtues), are

[65] It is also important to note that Darwin never mentions the issue of truth or falsity in relation to moral judgments; however, most contemporary Darwinian Moral Nativists are concerned with this issue. As we will see, both sentimentalist and Darwinian traditions influence them as well.
[66] See Hume (1751/1991: 84).

directed towards the interest of society and promote obligations to further the good of society.[67] Thus the sentimentalist view of Hutcheson and Hume (and Smith as well) also include some utilitarian elements.[68] And, as previously mentioned, Darwin's view also entails some utilitarian elements (mostly inspired by Mill), and thus the public good plays an important role in Darwin's view as well.

For Hutcheson and Hume the public good is directly linked to the sentiment of benevolence, which benefits mankind. As Hume says,

> Upon the whole, then, it seems undeniable, THAT nothing can bestow more merit on any human creature than the sentiment of benevolence in an eminent degree; and THAT a PART, at least, of its merit arises from its tendency to promote the interests of our species, and bestow happiness on human society. (1751/1991: 60-62)

The sentiment of benevolence is, thus, useful to society and, as such, plays an important role in morality. Similarly, Hutcheson thought man had genuine feelings of benevolence.[69] Hutcheson says that we have a disinterested concern for others; we have benevolent instincts found in our nature, like compassion.[70] According to Hutcheson, self-interest cannot explain benevolence and how we have the feelings of approval and disapproval that are associated with it. This non-reductionist view of benevolence denies the egoist positions regarding human nature, such as

[67] See Hume (1738/1978: 477-84).
[68] As noted in the (editor's) footnote in Hutcheson's "Inquiry Concerning Good and Evil", Hutcheson presents what is likely to be the "first explicit articulation of the utilitarian doctrine" (Hutcheson 1751/1991: 284). For a discussion of Hutcheson and Hume's influence on classical utilitarianism, see Driver (2012: 5-12).
[69] See Hutcheson (1751/1991: 282-83). Hutcheson's description of benevolence may even be considered genuine altruism, since rational agents "study the interest, and desire happiness of other beings" (1751/1991: 264). It is important to note that Hutcheson's discussion of benevolence was an argument against Hobbes' view of psychological egoism.
[70] Hutcheson (1751/1991: 273 and 289).

Hobbes—although we have self-love for survival, we also have benevolence in our nature, for the good of the community that the moral sense works off of.[71]

Hume agrees that self-love or self-interest does not promote the public good, while feelings of benevolence promote the moral good. However, for Hume, such natural sentiments are often felt only towards those we are immediately connected with—to oneself and one's own kin—and we are not naturally motivated to the public good because it is "too remote" (1738/1978: 481). So even though it is a natural virtue to be benevolent and generous, such moral sentiments are fairly limited. The only way that man can have concern for society more generally (beyond self and kin) is through sympathy. Sympathy, according to Hume, "…takes us so far out of ourselves to give us the same pleasure or uneasiness in the character of others as if they had a tendency to our own advantage or loss" (1738/1978: 578-79). And for artificial virtues, it is only the feeling of sympathy that produces moral sentiments.[72]

Again, there is a clear connection with Darwin's work. For Darwin, the public good and benefiting the community is central to morality. What is beneficial to society, however, is defined in terms of evolutionary fitness. Instead of focusing on the sentiment of benevolence, Darwin talks more generally about the social instincts, which give humans feelings of pleasure when in the company of others, a desire to help others, as well as sympathy towards others. Feelings of sympathy assist moral behavior since they promote the desire to help others. But whereas Hume thought natural feelings of sympathy could help extend benevolence beyond one's kin, Darwin argued that since sympathy was closely tied to the social instincts, the desire

[71] Hutcheson (1751/1991: 282-83). Note that sympathy does not play the same role in Hutcheson's work as it does in others (like Hume) because he thought man had genuine feelings of benevolence that could be other-directed without sympathy. Sympathy, however, plays an important role in Darwin's work.

[72] Hume says, "sympathy produces our sentiments of morals in all artificial virtues" (1738/1978: 577).

to be helpful was usually only felt for those of close relation. Instead, Darwin argues that sympathy needs to be cultivated, through habit (and experience), to extend beyond immediate kin and into the broader community.[73] This capacity to develop and direct sympathy is similar to the role that artificial virtues play in Hume's view, which I will discuss shortly.

Darwin's discussion of sympathy focused not on Hume, but on the work of Adam Smith.[74] Like both Hutcheson and Hume before him, Smith roots morality in moral approval and disapproval that is guided by sentiments. However, for Smith, such moral approval and disapproval does not stem from a particular faculty, a moral sense, but rather from a particular sentiment, sympathy. Moral approval is to be "in sympathy", and moral disapproval is to be "out of sympathy". This process of moral deliberation requires the faculty of imagination; to sympathize with another, one must be able to have "fellow-feeling with any passion" of another. In other words, one must be able to see someone else's perspective or put themselves in someone else's shoes.[75] Thus, like Hutcheson and Hume, man is not entirely selfish but rather has some natural tendency towards benevolence. As Smith states in *The Theory of Moral Sentiments*:

> However selfish man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though they derive nothing from it except the pleasure of seeing it. (1790/1991: 201)

Darwin agreed with Smith that sympathy plays an important role in morality, but as we have seen, since sympathy is closely tied to the social instincts for Darwin, it does not work in the same way. And, the most significant difference is that Darwin relies on a moral sense whereas Smith does not.

---

[73] See Darwin (1871: 78-79).
[74] Darwin referred to Bain as well.
[75] See Smith (1790/1991: 200-03). Note that he is critical of Hutcheson. See Smith (1790/1991: 247-54).

Also relevant to Darwin's explanation of morality and the moral sense is the additional distinction between the natural and artificial virtues that Hume makes. For Hume, artificial virtues are based on social convention and "produce pleasure and approbation by means of an artifice or contrivance, which arises from the circumstances and necessity of mankind" (1738/1978: 477). For example, promising and justice are both artificial virtues that promote the interests of society but require social conventions to exist. As a result, the moral sentiments that occur from artificial virtue require one to be a member of some society or social group. Natural virtue, however, "proceeds immediately from original principles, without the intervention of thought or reflection" (Hume 1738/1978: 484). Natural virtues do not require social (or cultural) conventions and, thus, occur in all mankind regardless of one's cultural background or societal membership (that is, one need not be a member of any society to have such moral sentiments). For Hume, many moral virtues such as meekness, benevolence, beneficence, and generosity have a tendency towards the good of society.[76]

Similarly, Darwin's argument for the role of language in the promotion of the public good emphasizes the need for something like Hume's artificial virtues for man to develop sympathy and to fully realize the good of the community. [77] Language provides communities with the ability to express how they think its' members ought to act to promote the good of the community. The wishes and judgments of the community thus become an additional guide to action beyond the immediate sentiments felt through the moral sense. For Darwin, social or

[76] Hume (1738/1978: 578).
[77] Darwin (1871: 81-83).

cultural conventions can help promote the good of the community, which is similar to the role artificial virtues play for Hume.[78]

An area where Darwin diverges from the sentimentalist tradition, and specifically from Hutcheson, is the distinction between the moral and natural good. For Hutcheson, a distinction between natural and moral was necessary to explain the pro-social or altruistic actions associated with the moral good. First, the natural good is only about the perception of pleasure and the desire to possess those objects that are deemed good, whereas moral good is about approbation and perceptions of moral excellence.[79] Next, the natural good may concern things that are advantageous to or good for the individual, but it is not primarily for the good of others or the good of the public. Instead, natural good is a self-interested good, what Hutcheson refers to a "self-love" (1738/1991: 262). However, since we do feel pleasure from the public good and approve it, there must be another type of goodness—this is the moral good, or moral virtue, and it comes from separate sense, the moral sense.[80] When discussing why self-love (egoism) is not enough to support morality, he uses the example of approving of a successful tyrant or traitor. Here Hutcheson says,

> Why don't we then approve both alike? It is plain that we have some secret sense, which determines our approbation without regard to self-interest; otherwise we would always favor the fortunate side without regard to virtue, and suppose ourselves engaged with that party. (1738/1991: 266)

In other words, without a moral sense it is impossible to distinguish between the natural, self-interested good and the moral, altruistic good. For Darwin, however, there is no distinction since

[78] As mentioned, Darwin does not deny that the public opinion can be wrong, or in conflict with our social instincts, and thus detrimental to society at times.
[79] Hutcheson (1738/1991: 264).
[80] Hutcheson (1738/1991: 265-66). Here, he also discusses the difference between natural and moral evil, such as the evil that befalls someone from a storm versus the evil that comes a person cheating you.

what turns out to be morally good is part of the natural good. As previously mentioned, Darwin sought to show that some actions—moral actions—were not purely self-interested but rather encompassed concern for the overall welfare of society, which also benefited the species and not necessarily the individual.[81] And, these actions were rooted in a social instinct that could be explained via natural selection. Goodness, for Darwin, was merely an extension (or some version of) evolutionary fitness—the moral good could be explained via a natural good.

From this comparative analysis, it is clear that Darwin was influenced by the sentimentalist tradition dominant during his time and, as such, raises questions of how much of Darwin's explanation of the moral sense comes from these views and how much is the result of a genuine natural history of man. As we can see from the previous discussion of Darwin's view, he draws on many of the elements found in the dominant sentimentalists views: a reliance on sentiments, in particular the sentiment of sympathy, the moral process of approval and disapproval in relation to the good (virtue), a developed intellect (regardless of the exact role it plays), a utilitarian underpinning that aims for the public good, and some reliance on the role of social convention in morality.[82] These connections between the purely philosophical views and the natural sciences (e.g., evolutionary theory) will emerge again in the discussion of contemporary Darwinian Moral Nativism, which mirror this tradition.

Darwin's evolutionary explanation of the general moral good, however, did not extend much further. Unlike Mill's work in ethics and the other moral philosophers that influenced Darwin's writings, Darwin did not seek to justify any of the "oughts" provided by the moral sense or what he considered to be the "general good". Skeptical that such justifications for

[81] See Darwin (1871: 67) and Richards (1987: 218).
[82] What seems unique is the replacement of "God" as the starting point with "Nature". Others, such as Newton and Locke also made this shift.

morality could be found in natural history, Darwin did not draw any explicit normative (or meta-ethical) conclusions from the existence of an innate moral sense.[83] However, with Darwin we start to see how the genealogical (both deductive and critical) and theoretical approaches merge in evolutionary ethics—bringing together evolutionary explanations (the evolution of an innate moral sense) with the justification for the foundations of morality in the moral sense (the good of man described in evolutionary terms).

Darwin did not take the further justificatory step, perhaps because he had other doubts. In his correspondence with William Graham (1881), Darwin reveals a general skepticism about the reliability of evolved mechanisms like the human mind—he questions whether humans can rely on their own reasoning processes since they are products of evolution. Darwin says,

> But then with me the horrid doubt always arises whether the convictions of man's mind, which has been developed from the mind of lower animals, are any value or at all trustworthy. Would any one trust in the convictions of a monkey's mind, if there are any convictions in such a mind? (1881)[84]

Regardless of these doubts, by establishing the moral good in evolutionary terms, Darwin clearly set the stage for others to draw normative conclusions from an evolved innate moral sense and establishes the groundwork to interpret the moral sense as the foundation of morality. And many did just that—from his contemporaries such as William Kingdon Clifford and Leslie Stephan to present day Darwinian Moral Nativists like Robert Richards (1986).[85] That is not to say that Darwin didn't have any critics; in fact, he had many. Some, such as Thomas Henry Huxley, rejected the idea of evolutionary ethics entirely, while others, such as Herbert Spencer, presented

---

[83] As Allhoff points out, "nowhere in [Darwin's] ethical writings does he offer normative guidance; rather, he is interested in the development of man's moral sense" (2003: 90-91). Others interpret him as making normative claims, see Tim Lewens (2007: 167-76). I think Darwin is pretty close to making normative, prescriptive claims.

[84] Here, Darwin questioned what we could learn from such natural evolved mechanisms.

[85] For a discussion of William Kingdon Clifford and Leslie Stephan, see Farber (1994: Chapter 2).

a competing view of evolution of an innate moral sense. Before turning to the critics who rejected Darwin's view of morality, let's look more closely at Spencer's work since his view clearly merges the genealogical and theoretical approaches.

ii. Spencer's Moral Sense

During the same period, Darwin's contemporary Herbert Spencer also wrote about the connection between evolutionary theory and morality. Like Darwin, Spencer argued for an innate moral sense; however, Spencer quickly moved from merely explaining the origins of man's moral sense to arguing explicitly for a scientific morality with evolutionary theory as its foundation. Spencer's interest in evolution was different than Darwin's—evolution for Spencer provided a framework for his philosophy rather than a natural history or science. As such, Spencer aimed to justify certain social norms in terms of man's evolved moral sense, thereby rendering his normative theory "exclusively naturalistic—that is, evolutionary" (1898: x).[86] Spencer focused less on how the evolution of the moral sense worked and more on how the moral sense, once evolved, fit into a bigger evolutionary picture of man and his broader philosophical framework.

In *Social Statics* Spencer provides an in depth description of the moral sense as the foundation for a systematic morality. According to Spencer, the innate moral sense is an emotional faculty, independent of the intellect, which produces moral feelings (i.e., intuitions) that lead to our perceptions of right and wrong. In fact, for Spencer, intuitions that originate in the moral sense actually contain elements of the moral code. For example, our feelings of duty to behave in the right manner when dealing with others and feelings of gratification from honest

[86] For a discussion, see Farber (1994: 47). Note that the full citation is a reaction to the non-naturalistic elements (e.g., god) in his earlier writings.

transactions originate in the moral sense and lead man to a sense of justice.[87] Spencer also considered these intuitions to be an innate instinct or impulse towards right actions.[88] Following in the tradition of sentimentalists like Shaftesbury, Hutcheson, Reid, and others, Spencer placed intuitions, and therefore emotions rather than reason, at the center of his moral system.[89]

Spencer diverges from the intuitionist tradition, however, by placing limitations on moral intuitions and how we derive the moral code from them. According to Spencer, intuitions are not capable of solving every ethical issue. Instead, intuitions appear to differ between individuals, and especially between cultures. To settle this discrepancy, Spencer proposes that moral axioms be derived from the moral sense. Using reason, we can deduce a systematic morality. In other words, the moral law can be deduced from the moral sense.[90] Such a scientific deduction is possible for Spencer by viewing morality within an evolutionary framework. Here, Spencer is clearing taking the combined approach to the question "where does morality come from?" by using the genealogical explanation of an evolved innate moral sense to justify the foundations of the moral law (i.e., he deduces the moral law from the moral sense). Merging the genealogical and theoretical approaches together, Spencer creates an evolutionary framework for his moral theory.

Spencer's evolutionary framework, however, differs from Darwin's view of evolution. According to Darwin, natural selection was the central mechanism responsible for evolutionary changes that led to adaptations, and it was unlikely that any changes acquired during an

[87] Other examples include freedom, equality, and safety (Spencer 1851: 24).
[88] See Spencer (1851: 18-28).
[89] Spencer refers to these philosophers when describing the moral sense (1851: 27).
[90] Spencer (1879: 26) and (1851: 28-30). Also, see "Letter to Mr. Mill" in Bain (1872: 722).

organism's lifetime could be passed to the next generation.[91] Spencer, on the other hand, subscribed to a version of Lamarck's theory in which acquired traits could be inherited since natural selection failed to fully explain evolution.[92] It is important to note that neither Darwin nor Spencer had an understanding of genetics and that the mechanisms of heredity were not understood—in fact, this is just what Darwin was trying to discover in his work.[93] In addition to their different views on the generational transfer of traits, they also disagreed with the underlying purpose of the mechanism of selection—not just how it selected but what is was selecting for and why. For Darwin, there was no *telos* or goal that selection was directed towards, and as a result, the moral sense was simply the natural consequence of selection, a contingent feature of man that developed as a result of evolution, and not the result of some "master plan". As Farber points out, Darwin thought that "the evolution of life and the evolution of culture were each the unanticipated result of a blind process of natural selection" (1994: 16). The evolutionary process for Spencer, however, was part of a larger, goal-directed endeavor with a particular *telos* in mind. Farber describes it as "not only broader [than Darwin's view] but goal-directed, self-corrective, and optimistic in that it assumed progress over time" (1994: 44).[94] As such, Spencer argued that the moral sense developed in man in order to help achieve a particular goal—that is, the moral goal of equal freedom which is the "liberty of each, limited by the like liberty of all" (1851:

[91] Darwin (1861: Chapter 4). Darwin talks about habit changing things during one's lifetime and then the possibility of passing them on.
[92] See Farber (1994: 49) and Allhoff (2003: 90). Also See Lamarck (1809/1984).
[93] The term *genetics* was not coined until 1905. Although Mendel was working on these issues around the same time as Darwin and Spencer, there is no evidence that they knew about his work. Even today, with advances in modern genetics and the human genome project, exactly how heredity works still remains somewhat mysterious and many unresolved problems linger. For example, recent debates in epigenetics about the ability of proteins to regulate gene expression by turning certain genes "on" and "off due to different environmental factors raise questions about how genes function. For a discussion see Krimsky (2013).
[94] See Spencer (1851: 85-89, 103, 131).

page). This goal oriented, teleological approach to morality via evolution becomes prominent in some of the Darwinian Moral Nativist views, such as that of Richards (1986) that seek to justify moral realism with evolutionary theory.[95]

Spencer proposed three evolutionary ends of human conduct that furthered this ultimate goal of equal freedom: (1) self-preservation/individual fitness; (2) fitness of one's offspring; and (3) group welfare.[96] Since these evolutionary ends were equivalent to moral ends for Spencer, he also believed that the evolutionary law was the moral law.[97] Following the laws of evolution, man would eventually evolve, or progress, into the ultimate moral man—a man "…whose private requirements coincide with public ones…who, in spontaneously fulfilling his own nature, incidentally performs the functions of a social unit; and yet is only enabled so to fulfill his own nature by all others doing the like" (Spencer 1851: 417).

Spencer's understanding of equal freedom and the ultimate moral man was influenced by Mill's utilitarianism. Reflecting on Mill's utilitarianism, Spencer argued that the moral end for man of equal freedom included happiness, which he defined in evolutionary terms rather than as a hedonistic value.[98] For Spencer,

> [H]appiness was found in justice and the individuals' acquisition of the maximal freedoms consistent similar freedoms for all…[because]…maximal freedom allows for the greatest variability on which evolution could act and that those variability, in turn, would provide the necessary means to bring about improved states of affairs in the world. (Thomson 1999: 476)

[95] Some other examples include, David Copp (2008), David Enoch (2010), and Erik J. Wielenberg (2010). Although these views try to deny a teleological view of evolution, we will see that they end up implying it anyway.
[96] Richards (1987: 304) and Spencer (1879: 58-60, 167).
[97] See Allhoff (1994: 95) and Spencer (1898: 58-60).
[98] See Spencer (1851: 227, 233 and 336) and Richards (1987: 300). Mill was in direct correspondence with Spencer about these views.

Thus, man ought to strive to follow the natural direction of evolution towards greater diversity, which would eventually lead to a utopian society. Spencer describes this utopian society as an "industrial society in which mutual aid replaced competition as the motive social force and in which the greatest individual freedom prevailed" (Farber 1994: 44, 49).

Spencer also saw himself as a "rational" utilitarian, rather than an empirical utilitarian like Bentham because he saw it as the next step in the evolution of utilitarianism.[99] As he says in a letter to Mill (against empirical utilitarianism):

> The view for which I contend is, that morality properly so-called—the science of right conduct—has for its object to determine how and why certain modes of conduct are detrimental, and certain other modes beneficial. These good and bad results cannot be accidental, but must be necessary consequences of the constitution of things; and I conceive it to be the business of moral science to deduce, from the laws of life and the conditions of existence, what kinds of action necessarily tend to produce happiness, and what kinds to produce unhappiness. Having done this, its deductions are to be recognized as laws of conduct; and are to be conformed to irrespective of a direct estimation of happiness or misery…And the objection which I have to the current utilitarianism is, that it recognizes no more developed form of morality—does not see that it has reached but the initial stage of moral science. (Spencer 1879: 64)[100]

For Spencer, what is good and bad for humans can be deduced from fundamental principles grounded in the moral sense; therefore, utilitarian calculations should stem from those principles as well. On this view, moral intuitions alone are insufficient for making a proper utilitarian calculation. Instead, intuitions about good and bad in conjunction with the moral (evolutionary) principles, such as equal freedom, lead to the best actions. And although moral intuitions are

[99] Spencer says, "at present, utilitarians pay no attention to this distinction. Even when it is pointed out, they disregard the fact that empirical utilitarianism is but a transitional form to be passed through on the way to rational utilitarianism" (1879: 21). In *Herbert Spencer: A Century Later*, Tibor R. Machan describes the rational utilitarian as holding the view that "actions should reap benefits, but they can do so only if rationally guided, if they are principled" (1978: Sec. "Spencer's Principles of Ethics", para. 5).

[100] Here, Spencer also refers to Mill as a rational utilitarian like himself.

limited, these utilitarian calculations are intuition-laden for Spencer because they are grounded in fundamental moral principles that are derived from the moral sense. According to Spencer,

> [C]onduct in its highest form will take as guides, innate perceptions of right duly enlightened and made precise by an analytic intelligence; while conscious that these guides are proximately supreme solely because they lead to the ultimately supreme end, happiness special and general. (1879:173)

As such, Spencer ties his utilitarianism into an evolutionary framework because he believes there is a science to morality—there is a natural mechanism, an evolved innate moral sense, that provides the requisite moral intuitions from which we can derive the moral rules that underpin proper utilitarian calculations. Thus, Spencer gives an evolutionary explanation not only of the origins of morality, but also what he considers the correct moral theory.

But Spencer's evolutionary explanations were not without issue. Like Darwin (and contemporary evolutionary ethicists), Spencer also recognized the problem of altruism, which is often thought to be central to any explanation of morality. According to Spencer, altruism should be understood broadly as promoting the welfare of others or as "acts by which offspring are preserved and the species maintained" (1879: 201). Altruistic acts also include both conscious and unconscious acts, those "without mental representation of the welfare—acts of automatic altruism" and thus altruism does not require intentional acts of selflessness (1879: 201). Instead, altruism evolves from the lowest, unconscious form of other-regarding acts to the highest, conscious form. Spencer also describes the evolution of altruism as moving from the individual and their kin to the broader society—developing "from unconscious parental altruism to conscious parental altruism of the highest kind… from the altruism of the family to social altruism" (1879: 204).

Since altruism involves selflessness (self-sacrifice), or as Spencer says the "expenditure of individual life to the end of increasing life in other individuals", it is difficult to explain and

understand how it could evolve as he describes (1879: 201). Spencer recognizes this problem and attempts to explain the evolution of altruism by tying it closely to the development of egoism. Spencer argues that:

> From the dawn of life, then, egoism has been dependent upon altruism as altruism has been dependent upon egoism; and in the course of evolution the reciprocal services of the two have been increasing. (1879: 201)[101]

According to Spencer, altruism and egoism are "co-essential"—pure egoism and pure altruism do not exist since following either on its own is ultimately destructive.[102] Instead, egoism and altruism work together to promote both individual and social welfare, and one's individual welfare (egoistic satisfaction) depends on social welfare (altruistic satisfaction of others). Robert Richards summarizes Spencer's explanations of the evolution of altruism in terms of several selection mechanisms: family altruism (kin selection), societal altruism (group selection and acquired habits), and adaptation to the social state (you need altruists to survive, too many egoists and you won't survive—survival of the fittest).[103] As such, Spencer claims "the well-being of each rises and falls with the well-being of all" (1879: 208).[104]

As with Darwin, Smith's view of sympathy also influenced Spencer's explanation of altruism.[105] According to Spencer, "the root of all altruistic sentiments is sympathy" (Spencer,

---

[101] Here, Spencer also notes that altruism is only recent in mankind's history.

[102] As Spencer says, "The need for a compromise between egoism and altruism is thus made conspicuous. We are forced to recognize the claims, which his own well-being has on the attention of each by noting how, in some directions we come to a deadlock, in others to contradictions, and in others to disastrous results, if they are ignored. Conversely, it is undeniable that disregard of others by each, carried to a great extent is fatal to society, and carried to a still greater extent is fatal to the family, and eventually to the race. Egoism and altruism are therefore co-essential" (1879: 237).

[103] For a discussion, see Richards (1987: 259-60, 308-11).

[104] Here, Spencer provides some examples of how this might occur: breaches of contract and direct aggression will cause the well-being of society to fall whereas sacrificing for social welfare will cause the well-being of society to rise.

[105] See Spencer (1851: 96-103).

1885: 32). Thus the development of altruism requires the moral sentiment of sympathy. Here, Spencer draws from Smith's work since he thought Smith explained the role of sympathy correctly as "giving rise to these superior controlling emotions" needed for altruism (Spencer 1885: 32). However, he thought that Smith failed to explain how sympathy evolves to become part of human nature and how sympathy evolves into a sense of justice. Spencer argued that the development of sympathy is tied with the development of the intellect, which includes the ability to recognize and understand language and the feelings they reflect. A developed intellect also allows us to create representations of those feelings and reflect on them.[106] Mankind still requires further development to create a society where sympathy is dominant, but this is (as described above) the direction of man's evolution, as Spencer believed.

Obviously, Spencer had greater expectations regarding the potential of evolutionary theory to resolve moral questions than Darwin, who stuck primarily (although not exclusively) to the biological and natural history of man. Thus, unlike Darwin, Spencer quickly moved from merely explaining the origins of man's moral sense to arguing explicitly for a scientific morality based in evolution with normative implications. For Spencer, the development of an innate evolved moral sense grounds a fundamental moral principle, and, thus, the foundations of morality, in the moral sense. Consequently, Spencer's view fully merges the genealogical and theoretical approaches by using an inquiry into the origins of morality—an evolutionary explanation of how morality develops in humans—to establish the fundamental moral principle

[106] See Spencer (1885) and (1879). Regarding sympathy Spencer says, "Through sympathy both these forms of feeling are excited…the average development of sympathy must, therefore, be regulated by the average manifestations of pleasure and pain in others…if the social state is such that manifestations of pleasure predominate, sympathy will increase; since sympathetic pleasures, adding to the totality of pleasures enhancing vitality, conduce to the physical prosperity of the most sympathetic, and since the pleasures of sympathy exceeding its pains in all, lead to an exercise of it which strengthens it" (1879: 244-45).

rooted in man's moral sense. It is important to point out that here Spencer is providing not only an evolutionary explanation of where morality comes from, but also a justification of a particular morality—rational utilitarianism. And that, unlike Darwin's case, is what makes this an example of a fully combined approach, an example of evolutionary ethics. Spencer's evolutionary ethics brings us back to the initial question: can two disparate approaches, with different methods and aims, be used as one? In other words, can the evolutionary history of morality be used to ground moral obligations in man's moral sense?[107]

Many of the original arguments against both Darwin's and Spencer's views of the moral sense provide a negative answer to these questions and challenge the combined approach. Examining these arguments will not only shed light on some of the problems the evolutionary ethics (and combined approaches) face, but it will also help illustrate why some of the views in contemporary Darwinian Moral Nativism require a closer look.

B. Historical Criticisms and their Impact on Contemporary Darwinian Moral Nativism

Neither Darwin nor Spencer escaped without serious criticism—criticisms that still haunt most contemporary views in evolutionary ethics. In *The Theory of Evolution and its Application* (1876), Henry Sidgwick criticized both Darwin and Spencer, arguing that the origins of morality, and more specifically of a moral sense, fail to provide any philosophically useful information. Sidgwick acknowledged that evolution might indeed tell us about the origins of morality in that we evolved some capacity for morality and thus, Darwin's theories of the origins and growth of

[107] Considering Nietzsche's criticism of Paul Ree's evolutionary ethics, he would likely disapprove of Spencer's use of the genealogical method. See Nietzsche (1887/1994: 6, 8-9). As Elbe points out, "it reduced the history of morality to the notion of utility in the present" (Elbe: 2001: 262). Also see Prinz (2009b: 216).

the moral sense could be true. However, such information about the origins of a moral capacity does not help explain or justify which ethical theories or which ethical principles are correct.[108]

i. The Naturalistic Fallacy and Hume's Thesis: Why They Are No Longer Knock-out Arguments

The most prominent philosophical argument against evolutionary ethics is the identification of the naturalistic fallacy, which was first proposed by Sidgwick (although he did not call it such). Broadly speaking, the naturalistic fallacy aims to reveal the inherent problems associated with equating the 'natural' with the 'good'. Along with Sidgwick, G.E. Moore (who coined the term naturalistic fallacy) was one of the first to focus this type of argument directly against evolutionary ethics by directing it at Spencer's work.[109] In addition to, and often closely associated with, the naturalistic fallacy is Hume's thesis: "that *ought-statements* cannot be deduced from exclusively *is-statements*" (Sober 1994: 103).[110] Hume did not use this thesis to criticize evolutionary ethics; however, Sidgwick relied on a similar thesis, in conjunction with the naturalistic fallacy, to further argue against evolutionary ethics and, more broadly, against ethical naturalism. Both the naturalistic fallacy and Hume's thesis remain the most prominent arguments lodged against all forms of evolutionary ethics and thus they deserve special attention.[111]

Analysis of the naturalistic fallacy and Hume's thesis reveals several important considerations. First, we will see how the naturalistic fallacy and Hume's thesis provide

---

[108] See Sidgwick (1876: 59).
[109] See Moore (1903/2004) and Sidgwick (1874/1981) and (1876). Each had a slightly different formulation of the fallacy and arguments to support it.
[110] There is much debate over the naturalistic fallacy, including what fallacy it actually refers to and whether such fallacy can really be equated with the arguments presented by G.E. Moore in *Principia Ethica* or in Hume's *Treatise on Human Nature* Book III, Part I, Section I. For a discussion see Joyce (2006: 143-56).
[111] For example, see Harman (2012), Woolcock (1999), and Barrett (1991). Of course, these two were not the only arguments against evolutionary ethics. For a discussion see Farber (1994).

additional reasons—controversial as they may be—to reject the collapse of the deductive genealogical approach with the theoretical approach that appears in most forms evolutionary ethics.[112] More importantly, close examination of these criticisms also expose a set of positions that these seminal arguments do not directly impact. These elusive arguments are a form of Darwinian Moral Nativism that merges the critical genealogical approach with the theoretical approach. Rather than attempting to use evolutionary explanations to justify any normative (prescriptive) claims or to support any particular metaethical foundation, these arguments attempt to undermine the very foundation of ethics itself. As such, these arguments are generally classified as *evolutionary debunking arguments*. Although I will provide a more detailed explanation of evolutionary debunking arguments in the next chapter, broadly defined "evolutionary debunking arguments are arguments that appeal to the evolutionary origins of certain evaluative beliefs to undermine their justification" (Kahane 2011:103). Many of these Darwinian Moral Nativist debunking arguments explicitly argue that they are immune from the wrath of the naturalistic fallacy and Hume's thesis. As a result, these arguments—the Darwinian Moral Nativist debunking arguments—will require further exploration and will become the focus of the remainder of this dissertation.

In *The Methods of Ethics*, Sidgwick presents several arguments against evolutionary ethics as well as any attempt to ground ethics is nature. The two central issues facing evolutionary ethics and ethical naturalism for Sidgwick are the naturalistic fallacy and Hume's thesis, although he does not refer to them as such. The first problem—the naturalistic fallacy—stems from attempting to equate what is natural with the good. Sidgwick provides the following

[112] Controversial as these criticisms may be, most proponents of evolutionary ethics take time to address both the naturalistic fallacy and Hume's thesis. Some simply deny that these are valid criticisms while others take time present counterarguments.

argument to support this claim. Any position that attempts to equate what is natural with what is good can be reduced to something like: It is good because it is natural because what is natural is good.[113] As such, equating the good with the natural begs the question unless more can be said about what is natural. This is not going to be possible, however, since according to Sidgwick the term 'natural' is indefinable—the term natural "is derived from simple, unanalyzable impressions" (1874/1981: 83). In other words, there is no further explanation that can be provided for what the term 'natural' means, and thus deeming what is good as natural is a mistake, or at least completely uninformative.

In "Theory of Evolution and Its Practice", Sidgwick further argues that even if there is some truth to what Darwin claims—that the moral sense helps us strive for self preservation—evolutionary fitness should not to be equated with the 'good'.[114] There is a sharp and profound difference between the concepts of 'being', which concerns existence itself or living, and 'well-being', which concerns living well. Neither Darwin nor Spencer (or views like them) accounted for this difference.[115]

Although the fallacy that equating the natural with the good originates with Sidgwick, it was his student G.E. Moore who coined the term the naturalistic fallacy in his work *Principa Ethica* in 1903. As Sidgwick already argued, Moore points out that the problem occurs when any moral theory attempts to define the term 'good' because 'good' is a simple, unanalyzable, and

[113] See Sidgwick (1874/1981: 81).
[114] Recall that Darwin claims that the good is "the means by which the greatest possible numbers of individuals can be reared in full vigor and health, with all their faculties perfect under the conditions to which they are exposed" (Darwin 1871: 94).
[115] Note that Sidgwick also makes an argument against Spencer's view of utilitarianism. See Sidgwick (1876: 59-60).

indefinable term.[116] In other words, the term 'good' (and others such a 'pleasure' and 'yellow') cannot be further explained by equating it with any other properties (except itself). This confusion, however, is quite common in moral theories. As Moore says,

> …if he confuses 'good' which is not in the same sense a natural object with any natural object whatever, then there is a reason for calling it a naturalistic fallacy. (1903/2004: 13)

What Moore adds to this thesis is not only a defense of the claim that these terms are simple and unanalyzable, but also the additional claim that even non-naturalistic moral theories are susceptible to this problem.[117] To defend this view that 'good' is a simple, indefinable term, Moore presents what has now become known as the *Open Question Argument*. Simply put, no matter how one defines good or goodness, it is still a reasonable and logical question to ask what makes that explanation (or definition) good. As such, Moore argues that this proves that 'good' is a simple and indefinable term.[118]

Moore specifically mentions Herbert Spencer when discussing the naturalistic fallacy. Moore says,

> It is absolutely useless, so far as Ethics is concerned, to prove, as Mr. Spencer tries to do, that increase in pleasure coincide with increase of life, unless good means something different from either life or pleasure. He might as well try to prove that an orange is yellow by showing that it is always wrapped up in paper. (1903/2004: 15)

Many contemporary views are also guilty of defining goodness as an extension of evolutionary fitness. For example, Richards (1986, 1987), a strong proponent of evolutionary ethics, claims

---

[116] Note that this something Moore takes from Sidgwick and whom he references as well. See Moore (1903/2004: 17-21).

[117] According to Moore, the naturalistic fallacy is a fallacy that applies not only to naturalistic moral theories but metaphysical theories as well. Since the issue is one of equating 'good' with any other property or object, that property or object need not be natural—it could be the supersensible objects provided by metaphysical views such as Kant, Spinoza and others. Thus the naturalistic fallacy is a problem not simply for those who attempt to equate 'good' with 'natural' in an attempt to define 'goodness' but for any attempt to equate 'good' with anything whatsoever. See Moore (1903/2004: Chapter 2 and 3).

[118] See Moore (1903/2004: 15-17).

that community welfare is the highest moral good because it is grounded in an evolved moral sense.[119] Such contemporary views, however, do not see themselves as committing the naturalistic fallacy. They either deny the fallacy entirely or they attempt to argue that their view is not in violation of this fallacy.[120] The details of these responses to the naturalistic fallacy are not immediately relevant. For regardless of the strength or validity of the naturalistic fallacy, it still aims to be an argument against evolutionary ethics (and, more broadly, ethical naturalism). Thus, any attempt to define what is good using evolutionary theory will be guilty of committing the naturalistic fallacy—the only question that remains is whether Moore and Sidgwick are correct about there being such a fallacy.[121]

More often, however, contemporary views are actually focused on Hume's thesis and not the naturalistic fallacy, as they are often confused with one another.[122] Hume's Thesis, which can be found in *Treatise of Human Nature* is derived from the following claim:

> I am surprised to find, that instead of the usual copulations of propositions, is, and is not, I meet with no proposition that is not connected with an ought, or an ought not. This change is imperceptible; but is, however, of the last consequence. For as this ought, or ought not, expresses some new relation or affirmation, it is necessary that it should be observed and explained; and at the same time that a reason should be given, for what seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it. But as authors do not commonly use this precaution, I shall presume to recommend it to the readers; and am persuaded, that this small

---

[119] See Richards (1987: 603) and Richards (1986: 286).

[120] For example, Richards argues that all ethical systems must make some empirical assumptions to get off the ground. The naturalistic fallacy therefore does not undermine the empirical assumptions he makes regarding the evolutionary impact on human nature and the ethical implications that follow (Richards 1986: 281-83). Also see Richards (1987: 612-20), Ruse (1986), Joyce (2006: 146-56), Walter (2006), Wilson, Dietrich and Clark (2003), and Darwall, Gibbard and Railton (1992). For a more general discussion of why the naturalistic fallacy fails, see Baumrin (1968).

[121] The naturalistic fallacy has been the subject of countless debates, one that I do not aim to resolve here.

[122] See Joyce (2006: 155). I imagine this confusion likely stems from what Sidgwick says in *The Method of Ethics*, "Every attempt to derive what 'out to be' from what 'is' palpably fails, the moment it is freed from the fundamental confusion of thought" (Sidgwick 1874/1981: 81).

> attention would subvert all the vulgar systems of morality, and let us see, that the distinction of vice and virtue is not founded merely on the relations of objects, nor is perceived by reason. (1738/1978: 469-70)

From this passage, many interpret Hume as making the claim that you cannot derive an 'ought' from an 'is'. In other words, you cannot infer or deduce prescriptive, evaluative claims (conclusions) from descriptive, non-evaluative claims (premises). To make such an attempt is to make a logical leap and thus commit a logical fallacy. As applied to evolutionary ethics, the problem would arise when deriving moral oughts from evolutionary explanations (e.g., a moral sense, altruism, etc.).

Sidgwick makes a similar argument in *The Methods of Ethics*. According to Sidgwick, evolutionary ethics is an example of an empirical ethics that uses the deductive method. It claims to start from empirical evidence—certain facts about the evolution of man—and deduces from it moral rules or principles. Herbert Spencer is a prime example of this type of approach. Sidgwick, however, makes an interesting observation.[123] He points out that Spencer, and those similar to Spencer who attempt to employ this type of empirical deductive method to ethics, don't actually use empirical data. Instead, they create an idealized version of man living in an ideal society (e.g., Spencer's absolute ethics).[124] In other words, although they claim to successfully derive an ought from an is, they don't actually do so. Thus, no actual deductions take place in these theories. Of course, this does not prove that one cannot derive an ought from an is, but it does raise one's suspicions. Many contemporary views also present questionable empirical data (e.g., evolutionary explanations of evaluative judgments) that raise similar issues (which I will explore in more depth later).

---

[123] Spencer tried to argue that he was not guilty of equating the natural with the good, but he fully supported the idea that one could derive the moral law from science (empirical facts).
[124] See Sidgwick (1874/1981:177-78).

Again here, as with Moore's naturalistic fallacy, there have been many interpretations and arguments over not only whether Hume was correct about this thesis but also what Hume really meant when he wrote this passage.[125] Even so, Hume's thesis has been used as an argument against evolutionary ethics, and moral naturalism more generally, because such views clearly attempt to derive what ought to be (prescriptive claims) from what is. In the case of evolutionary ethics, the 'is' includes evolutionary explanations of evolved mechanisms, such as the moral sense, and other aspects of our moral psychology.

In terms of the genealogical and theoretical approaches that I introduced earlier, if both of these traditional counterarguments—the naturalistic fallacy and Hume's thesis—turn out to be legitimate, then they provide additional reasons to reject the collapse of these two approaches. Clearly, any attempt to use an evolutionary genealogy (which includes, of course, empirical data) to justify any particular theoretical position in ethics will be guilty of violating Hume's thesis. And, a certain subset will also be guilty of equating the 'natural' with the 'good', thereby transgressing the naturalistic fallacy. But, the views that will be subject to the criticism only appear to include those that collapse the deductive genealogical approach with the theoretical approach. For the critical genealogical approaches that don't aim to justify directly any theoretical positions, these traditional criticisms won't directly undermine them.

C. A New Critique of Evolutionary Ethics

At this point, some might argue that the collapse I have presented is merely another version of Hume's thesis—that the distinction between the two approaches, which is ultimately a distinction between explanation and justification, is just another way of framing the is-ought problem. If explanation reduces to an "is" and justification reduces to an "ought", then the

[125] For such debates, see Pigden (2010) and Hudson (1969).

distinction I provide is just the same problematic gap found in Hume's thesis. Reductionist arguments can be tempting—they promise sufficient explanations without superfluous complexities, giving us those Quinean desert landscapes that are so aesthetically pleasing. But such a reductionist objection would be unfounded.

An explanation is not simply an "is" and a justification is not simply an "ought" because there are varying levels of explanation and justification that are determined by the initial inquiry. For any particular phenomenon, there will be several explanations that satisfy different inquiries about that phenomenon. As a result, there is no one "is" for an explanation to map onto; there cannot be a one-to-one correlation between an explanation and an "is". In this case, "is" would be better understood as the phenomenon that requires explanation rather than the explanation itself. The same holds for justification. There is no one "ought" for justification to map onto since the relationships vary based on the aims of inquiry as well. As such, there is no way automatically to eliminate or reduce one to the other.

The argument I present, therefore, introduces a new approach to dealing with evolutionary ethics that does not rely on these traditional arguments. Framing evolutionary ethics in this way reveals a group of positions that appear to escape these traditional arguments—those that merge the critical genealogical approach with the theoretical approach. In evolutionary ethics, these approaches are typically found in the form of the Darwinian Moral Nativist debunking arguments such as those prominent views of Michael Ruse, Richard Joyce, and Sharon Street. All of these views use the evolution of an innate moral sense to support particular metaethical positions.

Broadly speaking, these Darwinian Moral Nativist debunking views argue that the manner in which our moral sense evolved—the mechanism that influences our moral beliefs and

evaluative judgments—gives us reason to distrust those moral beliefs and evaluative judgments. These views rely on evolutionary explanations to undermine moral realism in various ways. As such, they do not directly attempt to derive an "ought" from an "is"—they do not use evolutionary ethics to support a particular moral claim directly—nor do they attempt to equate what is natural with the good. Rather than justifying any moral "oughts" from an evolutionary "is", they aim to get rid of the "oughts" altogether. It is these arguments then that require further consideration. And although these views appear to escape these traditional arguments, they will not emerge from further analysis unscathed.

Chapter Three: The Debunking of Morality

In this chapter, I will begin by clarifying the general nature of debunking arguments. I will then look at the connection between Darwinian Moral Nativism and debunking arguments that seek to undermine moral realism. Next, I will provide a brief overview of the various types of evolutionary debunking arguments that have been made, followed by a more in depth analysis of a few examples of these types of debunking arguments. I will categorize these arguments based on the metaethical claims they attempt to debunk. In so doing, I will provide a brief analysis of each of the views. In my analysis, I will discuss the logic behind (or the methodology that underlies) the debunking argument. Then, I will analyze each of the views more closely. Here, I will show that although these debunking arguments take aim solely at beliefs about morality (that is, at moral realism), they can't help but include beliefs about mathematics and science. Drawing a distinction between beliefs about morality and beliefs about mathematics and science is arbitrary, therefore these views end up entailing more than they set out for.

A. Darwinian Moral Nativism and Debunking Arguments

Some contemporary moral nativist views seek to undermine morality—that is, they argue that evolutionary theory provides the best explanation for morality and that this explanation undermines, or debunks, morality.[126] Evolutionary debunking arguments start with some form of moral nativism, i.e., the claim that humans possess an evolved innate moral sense.[127] They then argue that the existence of this evolved innate moral sense provides evidence against various aspects of morality, and in most cases this evolutionary "data" is used to undermine moral

[126] See Joyce (2013a: 17).
[127] In Kahane (2011: 112), he points out that all debunking arguments of this type (evolutionary and moral) must start with the assumption that there is an innate moral sense.

realism.[128] Moral realism can generally be understood as the metaethical position that there are moral facts that are independent of our minds, and when we make moral judgments, they are either true or false.[129] Thus, moral realism proposes both ontological claims (there are mind-independent moral facts) and epistemological claims (these claims can be true or false and are knowable) about morality. Of course, how a particular brand of moral realism establishes these claims varies and, thus, so do the different debunking arguments that target them.

In general, debunking arguments aim to show how the origin of a belief—i.e., the genealogy of a belief—undermines the truth or plausibility of that belief. As such, debunking arguments are a form of the genealogical approach, some demonstrative while others are critical. For most debunking arguments, the origin of the belief is thought to act as, what Guy Kahane refers to, an "undermining defeater", undermining the justification of the belief (although not the truth of that belief in itself) (2011: 106).[130] There are several types of debunking arguments that focus on different types of origin stories (or genealogical explanations).[131] For example, historical debunking arguments focus on the historical origins of a belief to undermine that belief.

---

[128] In some cases, debunking arguments target specific types of moral beliefs. For example, Peter Singer (2005) and Joshua Greene (2008) both employ debunking arguments that target deontological intuitions. For a discussion, see Kahane (2011: 110-14). Joyce also argues that Singer and Greene's arguments target moral theories and thus they fall under the category of debunking arguments that claim "certain moral theories (e.g., Kantianism)—but not all such theories—should be rejected" (2016: 2-3).

[129] Geoffrey Sayre-McCord's definition of moral realism provides a decent broad definition: "[m]oral realists hold that there are moral facts, that it is in light of these facts that peoples' moral judgments are true or false, and that the facts being what they are (and so the judgments being true, when they are) is not merely a reflection of our *thinking* the facts are one way or another" (1-2).

[130] To clarify, this is a purely epistemic point. The believer is not justified in holding the belief, but that does not prove or disprove the truth (or falsity) of the belief itself.

[131] Note that some genealogies are not used to question the truth or justification of a belief. For example, some, such as Hobbes (1651/2009), use causal explanations that are neutral about the these issues regarding the truth of a belief.

Marx and Nietzsche often employ this methodology in their work.[132] For both Marx and Nietzsche, however, certain historical origins act not merely as an undermining defeater, simply undermining the justification of the belief, but also as a means to undermine the truth of that belief itself.[133] In other words, there is an added inference that if the justification of the belief is undermined by the historical origins, that belief is not true. Another example of debunking is psychological debunking in which various psychological influences create biases in the formation of a belief that thereby undermine the credibility of that belief. Depression, for example, exerts such a distorting influence over an individual's beliefs such that many of the depressives' beliefs, say about reasons to live or how well one's life is going, become unjustified.[134] In the case of evolutionary debunking arguments, it is the evolutionary origins (or evolutionary explanations) of certain beliefs that determine the truth or plausibility of those beliefs and ultimately undermine those beliefs.[135] In some cases of evolutionary debunking, evolutionary origins act only as an undermining defeater, undermining the justification of the belief, while in others the evolutionary origins also serve as a means to undermine the truth of that belief itself.

---

[132] For a discussion, see Kahane (2011: 105-08) and Leiter (2004: 74-105). Also see my earlier discussion of Nietzsche in Chapter One.

[133] See Nietzsche, *On the Genealogy of Morals* (1887/1994) and Marx, *The German Ideology: Including Thesis on Feuerbach and a Critique of Political Economy* (1845-88/1998). For a discussion on this point, also see my Chapter One.

[134] For a more in-depth discussion of these two types of debunking, see Kahane (2011:105-9).

[135] As Wielenberg says, "the evolutionary explanation of moral beliefs imply that the beliefs in question can be explained without appealing to the truth of those beliefs" (2010: 452). See Griffiths and Wilkins (2013:133-46) for a discussion of the different domains of debunking and arguments about religion and science. In their discussion of evolutionary debunking, Griffiths and Wilkins point out that evolutionary debunking arguments often take aim at three main categories (or domains) of beliefs: religion, morality and science. However, I will only focus on evolutionary debunking arguments about morality since it is the implications of an evolved innate moral sense that are in question.

All debunking arguments are, therefore, a form of genetic argument; they are the type of argument that utilizes the origin of a belief to reach a conclusion about the truth or plausibility of that belief.[136] Genetic arguments are controversial since they often commit the genetic fallacy—a genetic argument is fallacious when it aims either to support or discredit a claim based on its origins when its origins are irrelevant to its truth. As noted in Chapter One, some genetic arguments, such as those employing the demonstrative genealogical approach, are subject to this common criticism while others, such as those employing the critical genealogical, appear to escape it. For now, however, I simply want to discuss what debunking arguments claim to achieve, particularly in the case of evolutionary ethics.

In evolutionary ethics, there is an initial divide between two basic types of debunking arguments: *theory debunking* and *metaethical debunking*.[137] Cases of theory debunking use evolution to debunk one normative theory while vindicating others. For example, both Peter Singer (2005) and Joshua Greene (2003) use evolution to undermine deontology and to support utilitarianism. Metaethical debunking arguments take aim primarily at the particular elements that make up moral realism, a metaethical view, rather than any particular normative claims (e.g., we ought to do x) or normative theory (e.g., utilitarianism). Granted, there may be certain normative implications that stem from these metaethical underpinnings; nevertheless the focus of these debunking arguments does not rest on any particular normative content *per se*. In fact, it is the metaethical nature of this type of debunking that separates contemporary evolutionary ethics from previous attempts at evolutionary ethics discussed in Chapter Two, which focused solely on prescriptive normative claims and on normative theory (e.g., Utilitarianism, Kantian Deontology,

[136] See Sober (1994: 104).
[137] Although I am not using his categories, Joyce (2013c: 3) classifies evolutionary debunking arguments into three categories: truth debunking, theory debunking and justification debunking. In using the term "theory debunking" I refer only to normative ethical theory.

Virtue Ethics, etc.). I will focus on the metaethical debunking arguments—that is, the arguments that claim that an evolved innate moral sense undermines moral realism in every form.

i. Metaethical Debunking Arguments

To understand how the existence of an evolved innate moral sense might undermine moral realism, there are three prominent evolutionary debunking views of moral realism in the literature that require further analysis. These are arguments put forth by Michael Ruse, Richard Joyce, and Sharon Street. These three analyses all share the primary goal of debunking morality at the metaethical level and not (necessarily) to vindicate any other particular normative theory (or claims) in morality. David Copp describes these evolutionary debunking arguments as all holding a "Darwinian Hypothesis", which is the view that:

> Darwinian forces have strongly influenced the phenomena of moral psychology with the result that the nature of the phenomena, including the content of our moral beliefs can be explained, at least in part, on the basis of evolutionary theory. (2008: 186)[138]

The Darwinian Hypothesis is the basic evolutionary explanation debunkers claim causes trouble for morality and ultimately undermines moral realism, or at least leads to moral skepticism. Although all debunkers share this initial hypothesis, each of their arguments aims to debunk morality using evolutionary theory in a different way.[139]

These three approaches to evolutionary debunking (Ruse, Street, and Joyce) are usually understood by further classifying them in terms of the metaethical claims they *seek* to challenge.[140] In Hallvard Lillehammer's (2003) discussion of evolutionary naturalism and moral error theory, he marks a distinction between epistemological debunking arguments, those that

---

[138] Copp refers to others who hold this general hypothesis, including Blackburn (2000), Gibbard (1982) and (1990), Kitcher (1993) and (2006), and Sober and D.S. Wilson (1998).
[139] In his most recent works, Joyce (2016a) and (2016b) doesn't classify himself as debunking moral realism directly. I will address this below.
[140] Whether these metaethical claims are the only thing these arguments challenge is questionable and, thus, is something I will discuss below.

consider the justification of moral judgments, and metaphysical debunking arguments, those that consider the existence of certain moral claims used to explain moral judgments.[141] Erik J. Wielenberg further clarifies this distinction by explaining that,

> [t]he [metaphysical debunking arguments] seek to [debunk] by showing that no moral belief that can be given an evolutionary explanation is true [while] [t]he [epistemological debunking arguments] make the case that the existence of an evolutionary explanation for a given moral belief implies that even if the belief is true, it is not knowledge. (2010: 442)[142]

In other words, the metaphysical debunking arguments seek to challenge the ontological status of moral beliefs while the epistemological debunking arguments seek to challenge the epistemological standing of moral beliefs. This divide between types of debunking arguments is reminiscent of a standard divide in metaethical theory between metaphysical moral theories that "tell us about the ontological status of the facts that render our moral beliefs true" and epistemological metaethical theories, in which "moral beliefs can be shown to be justified irrespective of the ontological status of the facts that make them true" (Joyce 2006: 211).[143]

In order to help clarify the various forms of debunking, it is better to understand metaphysical debunking as ontological debunking since it deals directly with the existence of moral facts rather than other metaphysical issues more generally. So, in terms of what these

---

[141] Lillehammer provides R.D. Alexander (1987) as his example of an epistemological debunking argument. Michael Ruse (1986b) and (1995) and Richard Wright (1994) are his examples of metaphysical debunking arguments. Additionally, Lillehammer defines moral error theory as the view that "moral judgments in general are false or incoherent" (2004: 95). The version of moral error theory he supports claims that "moral judgments are made relative to a framework of assumptions with no claim to reasonable convergence on their behalf" (2004: 95).

[142] Wielenberg sees the distinction to be similar to Alvin Plantinga's de facto and de re objections to Christian beliefs (Wielenberg 2010: 442, footnote 6). Also note that he refers to Ruse as making an epistemological debunking argument but later he admits, in a footnote, that it is a case of metaphysical debunking (just that it is outside the scope of his argument) (Wielenberg 2010: 454, footnote 48).

[143] For an in depth discussion, see Joyce (2006: Chapter 2). His examples of metaphysical theories include moral naturalism, moral non-naturalism, and supernaturalism. Examples of epistemological theories include reliabilism, conservatism, coherentism, and foundationalism.

debunking arguments *seek* to challenge, they are traditionally divided into two categories: *ontological debunking* and *epistemological debunking*. Ontological debunking arguments target the existence of moral facts whereas epistemological debunking arguments target epistemic status of moral beliefs. When considering the debunking arguments of Ruse, Joyce, and Street, it is typical to classify Ruse's argument as a form of ontological debunking, while Joyce and Street aim for forms of epistemological debunking.[144] However, this standard way of classifying the debunking arguments of Ruse, Joyce, and Street is slightly misleading. I will argue that Ruse and Street are both forms of ontological debunking, while Joyce's is the only epistemological debunking argument. Despite what these views seek to challenge, they often involve more than what they aim for. As a result, all three views will ultimately be shown to be unsuccessful.

It is also important to keep in mind that each of these arguments also includes various theoretical assumptions that play a significant role in their debunking arguments—from scientific claims about evolution to particular epistemological views, including claims about justification, truth, and knowledge. For example, all three debunking arguments appear to accept something like a correspondence theory of truth, in which the truth corresponds to facts in the world, which in turn, they claim, poses an unique problem for the moral realist in the face of evolutionary considerations. Thus, for each of these debunking views, I will aim to reveal the underlying scientific and theoretical assumptions that play a crucial role in the debunking argument, which will help shed light on some of their weaknesses. I will start the following section with Ruse's ontological debunking argument, since he attempts to establish the strongest conclusion of the three.

[144] For example, this classification can be found in Lillehammer (2003), Wielenberg (2010), Kahane (2011), and Joyce (2013c).

a. Ontological Debunking

The next two debunking arguments, by Ruse and Street, are examples of ontological debunking arguments that target the ontological status of moral beliefs. Although Street is typically considered to present an epistemological debunking argument, targeting the justification for moral beliefs, I argue that she includes enough elements of ontological debunking to classify her otherwise. I will start with Michael Ruse's view since he presents the most straightforward example of ontological debunking. Next, I will consider Sharon Street's debunking argument, which is a more complex argument than Ruse's. Some aspects of Street's argument, however, bear a close resemblance to Ruse's view, which explains why Street is an ontological debunker.

1. Michael Ruse's Debunking Argument

The first debunking view to consider is that of Michael Ruse, which targets ontological elements of moral realism. In his article "Evolutionary Ethics: A Phoenix Arisen", Ruse presents his main ontological debunking argument and begins his argument by providing evidence for an evolved innate moral sense.[145] Drawing on Darwin's *On the Origin of the Species* (1859*)* and *The Descent of Man and Selection in Relation to Sex* (1871), the work of other sociobiologists such as Robert Trivers (1971), E.O. Wilson (1975), Richard Dawkins (1976), and Charles Lumsden (1981), as well as his own work, Ruse aims to establish an empirical case for an evolved innate moral sense. Ruse describes an evolved moral sense as a set of innate mental dispositions (epigenetic rules) that incline humans to cooperate.[146] This type of cooperation

[145] Note that in addition to this debunking argument, Ruse argues elsewhere for both a normative and metaethical evolutionary ethics. See Ruse (1986b) and (2006).
[146] According to Ruse, epigentic rules are "genetically based processes of development that predispose the individual to adopt one or a few forms of behaviors as opposed to others. The rules are rooted in the physiological processes leading from genes to thought and action" (Wilson

occurs as a form of moral altruism, i.e., we help others because we think it is the morally right thing to do. Moral altruism, however, only serves the interests of biological altruism—when humans cooperate with their groups they benefit biologically. In evolutionary terms, to benefit biologically means to survive and reproduce; cooperation increases offspring. Ruse describes the distinction between moral and biological altruism as a distinction between *metaphorical* and *literal* altruism. Metaphorical (biological) altruism occurs in many different species and stems from causal mechanisms, such as kin selection and reciprocal altruism, which result in the kind of cooperation that appears similar to moral behavior—like protecting one's young—but turns out to be merely unreflective behavior.[147] Literal (moral) altruism, however, is a genuine form of altruism that requires the capacity for conscious reflection. Such capacity provides humans (or similar type beings, if there are any) with the ability to think that there exist certain objective moral obligations, which influences their behavior. For example, a human protects her young because she thinks (believes) it is an actual obligation. Cooperating for personal evolutionary gain, that is biological (metaphorical) altruism, however, explains why the moral sense evolved.[148]

According to what Ruse (1986) calls the "modern Darwinian biologist's evolutionary ethics", the moral sense is an adaptation—it was selected for because it increased fitness. As such, he makes the bold claim that evolution literally furnished our minds with thoughts about right and wrong. As Ruse says,

---

and Ruse 1986: 180). Such claims are still up for debate in the scientific literature; for a discussion, see Krimsky and Gruber (2013).

[147] For an in-depth discussion of Ruse's view of altruism see Ruse (2009c: 497-509). For an in-depth discussion of evolutionary altruism, see Ruse (2009b: Part IV, 411-548).

[148] See Wilson and Ruse (1986:179-80) and Ruse (1986: 97-100). Even though moral altruism can be explained in terms of biological altruism (through mechanisms of kin selection and reciprocal altruism), Ruse does not think these explanations eliminate the possibility of genuine feelings of altruism, that is, we are not egoists (1986: 97-100).

> 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…Thus we have developed innate metal dispositions…inclining us to cooperate, in the name of this thing we call morality. (1986: 99)

But even though moral altruism is innate, and thus we are inclined to behave morally, we are not pre-determined. Ruse states that "we are not blindly locked into our course of action like robots" (1986: 99). So, we are hard-wired to a particular morality, although we are free to choose whether or not to follow that morality. These are obviously controversial scientific claims, and thus the scientific validity of this type of argument for an innate moral sense is questionable. However, let us take these empirical claims at face value for now and focus on what is claimed to be the implications for morality.

Ruse argues that this evolutionary explanation of the moral sense leads to some major metaethical implications; that is, it leads to his argument for the ontological debunking of moral realism. According to Ruse, the basis of morality is a collective illusion of the human race (a set of shared beliefs) caused by evolutionary forces (e.g., natural selection, adaptations for reproduction). Ruse further claims that,

> Because of this, a causal analysis of the type offered by the evolutionist is appropriate and adequate, whereas a justification of moral claims in terms of reasoned foundations is neither needed nor appropriate. (1986: 103)

Consequently, there is no "foundation" for morality—morality, moral altruism, and our inclinations to act this way are all due to an evolved moral sense that is simply a mechanism to make us act in ways that promote biological altruism, which promotes individual fitness. Morality, therefore, consists only of subjective feelings and sentiments that we merely *think* are objective. The moral sense persuades us to think that morality is objective, that it has a normative force, even though it is only based in our subjective feelings. As Ruse states, "it is precisely because we think that morality is more than subjective desires that we are led to obey it" (1986:

103). This normative force also has an evolutionary explanation; if we did not believe that morality was real (objective), then we would not follow it and individual fitness would not be promoted.[149]

Already, we can start to see some of the weaknesses in Ruse's view. It is very odd to argue that even though we all share the same feelings and sentiments as a species, they are entirely subjective. If we all have the *same* subjective feelings, then how subjective are they? Additionally, if the normative force of such subjective feelings is also universal, then there does seem to be something objective about morality. Ruse, however, appears to be accepting some version of a moral-error theory. Broadly speaking, moral error theory argues that moral statements (or assertions) are systematically false because there are no moral properties or moral facts to make them true.[150] In Ruse's case, he is arguing that we think our moral claims are objectively true because we've evolved to believe this about the world and ourselves; however, there are no objective facts about morality to support those beliefs or make them true.

And this is how Ruse attempts to "debunk" moral realism: if moral feelings are found only in our subjective psychology, then morality does not have an objective foundation—there are no mind independent moral facts. In Richard Joyce's discussion of debunking views, he describes Ruse's conclusion as making the following claims:

> (1) All moral judgments are false.
> (2) All moral judgments are false insofar as they involve a claim to objectivity. (2016: 2)

Joyce goes on to point out that Ruse seems to argue that,

> If morality is necessarily objective, then establishing that there are no objective moral truths—as (2) claims—entails that there are no moral truths *tout court*—as (1) claims. (2016: 3)

[149] See Ruse (1986: 102-03).
[150] Probably the most prominent account of moral-error theory is by J.L. Mackie (1977).

So evolutionary influences have distorted our beliefs such that we think our moral judgments are true and objective, but when we examine them closely enough we see that all of our moral judgments false.[151] According to Ruse, "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" (1986: 103).

Ruse's primary argument against an objective foundation of morality (that is, for claims (1) and (2)) is that such an explanation of our moral beliefs is superfluous—"the objective foundation has to be judged redundant" (1986: 108)[152] In other words, the evolutionary explanation of the origins of our moral beliefs proves that there is no need for an objective morality. There may be causal explanations for why we have these moral feelings, but there is no justification (in terms of an objective foundation) for them.[153] Here, Ruse is applying the principle of parsimony to eliminate moral objectivity. As Joyce explains, Ruse is working under the premise that if we have a complete explanation of moral beliefs via moral nativism and there is no need to posit any extra ontology in the form of moral facts to explain morality, then we can just get rid of them—Ruse employs Ockham's razor to cut out moral facts.[154] And, as Ruse himself points out "[a]ll 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" (1986: 102). In other words, Ruse claims that the evolutionary explanation of morality is sufficient.

Ruse's argument is, in some ways, a version of Gilbert Harman's (1977) challenge to moral realism. Harman's challenge is an explanatory challenge, which asks the moral realist to

---

[151] Note that "all moral judgments is false" is usually taken to amount to an error theorist position. Here, Joyce refers to Mackie (1977) as an example of such a view.
[152] See also Ruse and Wilson (1986: 186-87).
[153] See Ruse (1986: 103). Examples of causal explanations of ethics include Murphy (1982), Mackie (1977), and Ruse and E.O. Wilson (1986).
[154] See Joyce (2006: 188) and Joyce (2013c: 531).

provide an account of how moral facts play an explanatory role in our best explanation of the phenomena in question. For example, what role do moral facts about "justice" play in explaining why it is "wrong" to torture an innocent child? On Harman's view, an adequate explanation of why torturing an innocent child is "unjust" only seems to appeal to the non-moral facts of the phenomena—e.g., the child is being tortured and the child is screaming and crying, asking the torturer to stop, etc. It is possible that a reductive account of moral (non-natural) facts to non-moral (natural) facts could show how the moral facts relate to the natural facts, but Harman doubts that any such satisfactory account has been given. If the best explanation for why it is wrong to torture an innocent child does not require any moral facts, then they are superfluous because they have no explanatory power.[155] Joyce frames Harman's challenge to moral realism as the following conditional statement:

> [I]f there is no reductive account available explaining how moral facts relate to naturalistic facts, then moral claims cannot be tested, moral theories cannot be confirmed or disconfirmed, and we have no evidence for the existence of moral facts. (2006:184-85)

Joyce considers Ruse's debunking argument in the context of Harman's challenge; for Ruse, the evolutionary explanation makes moral facts non-reductive and thus redundant; therefore, there are no moral facts. Joyce, however, questions Ruse's conclusion since he does not seriously consider the possibility of a reduction of moral facts to natural facts, although both Harman and Joyce doubt that such an explanation is forthcoming.[156]

When Ruse's argument is understood as a version of Harman's challenge, it becomes clear why Ruse's debunking argument is an example of ontological debunking. The evolutionary explanation of moral beliefs undermines moral objectivity. As Wielenberg explains, Ruse is essentially arguing that "if we do not need to posit moral facts to explain our moral sentiments,

[155] See Harman (1977: 3-26).
[156] See Joyce (2006: 188-89).

we should conclude that there are no moral facts", and as such it is a form of ontological debunking (2010: 454, footnote 48). This is also one of the boldest and most direct attacks on moral realism—there are no moral facts (all moral judgments are false) rather than the slightly weaker claim that (2) All moral judgments are false insofar as they involve a claim to objectivity, or the even milder epistemic claim that we are simply not justified in believing moral judgments.

It is important to point out that Harman's challenge, and Ruse's redundancy argument, face a significant problem when we examine the role of explanation more closely. Both Harman and Ruse presume that there can be *a* best explanation for any given phenomenon. But as I have argued in Chapter One, the success of an explanation is tied with the initial inquiry. Therefore, there may be various explanations of the same phenomenon depending on what the initial inquiry was. As a result, what we consider the "best" explanation will vary as well—there is not a single best explanation for any single phenomenon. This causes a problem for both Harman and Ruse because they both rely on the notion that there is a single best explanation available to eliminate moral facts, or more broadly moral objectivity.

Now, it may be argued that for each inquiry there is a best possible explanation for that inquiry; however, this will not help either Joyce or Ruse in the elimination of moral facts. Take the earlier example of torturing an innocent child. When the initial inquiry is "why are these non-moral facts unjust", then simply appealing to the non-moral facts of the phenomena—e.g., the child is being tortured and the child is screaming and crying, asking the torturer to stop, etc.—will not suffice. And to claim that this proves there are no moral facts (or at least no need for them) does not logically follow either. Instead, Moore's open question argument rears its ugly head: it can still be asked of the phenomenon, why is that unjust? Instead of discovering a best explanation, Harman and Ruse merely have an unsatisfactory explanation.

To add to the redundancy argument, Ruse provides a counterfactual (possible worlds) argument to bolster his conclusion that there are no moral facts. Ruse emphasizes the idea that since moral altruism (which is really biological altruism) is an adaptation, it is possible that mechanisms other than moral altruism could have evolved. As a result, "morality is a function of our actual human nature and…it cannot be divorced from the contingencies of our evolution" (Ruse 1986: 110). According to Ruse, if morality is contingent in this manner, then it cannot have any connection to objective moral facts. As he says, "morality, as we know it, cannot have the necessity or objectivity sought by the Kantian or Rawlsian" (1986: 110).[157] Since the moral sense could have developed differently, morality could have been otherwise, and so there is no objective foundation. Surprisingly, Ruse also claims that morality is not relative since we are all under the same illusion, with the same biological function underlying it.[158]

It is not clear, however, why the existence of an objective morality would be tied to the contingencies of evolution. Ruse's argument presumes that the objectivity of morality requires that humans have access to it. But, if a realist objective morality exists, it would exist independently of man (or the evolution of man) since that is what it means to have a realist objective morality—that there are moral facts that exist independently of humans and our minds. There is no requirement to objectivity that we have epistemic access to that objective reality (although it would be nice). In other words, morality could be objective without human awareness of it. Therefore, Ruse's possible worlds arguments fails to support his conclusion.

---

[157] Although Kant and Rawls are not realists, Ruse interprets Kantian and Rawlsian objectivity as a form of moral constructivism in which moral truths are a non-contingent feature of what it means to be a rational agent.

[158] As such, he claims to be a subjectivist about ethics but not a relativist (since we all share in the same moral 'game'"). Ruse also wants morality to "stay the same"—that is, he thinks we can keep living as if morality were true. In fact, we cannot help but to continue to believe that morality is objective because it is part of our biology to do just that. See Ruse (1986: 111).

Ruse's debunking argument clearly faces several problems, and I shall revisit these later. Let us now turn to the second ontological debunking argument by Sharon Street.

2. Sharon Street's Darwinian Dilemma: The Euthyphro Dilemma for Moral Realism

Sharon Street presents a challenge to moral realism, the 'Darwinian Dilemma', which targets epistemological issues yet concludes with the stronger metaphysical claim that evolutionary considerations support moral antirealism. Most of Street's critics classify her argument as epistemological because the primary metaethical argument that her view appears to address concerns moral knowledge.[159] However, Street reaches conclusions that extend beyond moral epistemology. She makes broader assertions about the metaphysical standing of morality by endorsing a particular metaethical position—Street attempts not only to debunk moral realism, but also to vindicate a version of moral antirealism that these critics overlook.

In "A Darwinian Dilemma for Realist Theories of Value", Street defines moral realists as those who believe that there are evaluative facts or truths that are stance independent. Street divides the moral realists into two camps: naturalist moral realism and non-naturalist moral realism. The naturalist claims that moral facts are identical with natural facts and as such, they play a role in causal explanations. Some examples of this position are Sturgeon (1985), Railton (1986), Boyd (1988), and Brink (1989). The non-naturalists, however, claims that moral facts are non-reductive (irreducible) and do not play a role in causal explanations. Proponents of this view include Nagel (1986), Dworkin (1996), and Shafer-Landau (2003).[160] This distinction between

[159] In addition to Lillehammer (2003) and Wielenberg (2010), this distinction can be found in Kahane (2011), Enoch (2010), and Skarsaune (2011).

[160] Street sets aside quasi-realist views entirely. Also note that she uses the term "value" instead of "moral". Although Street does not refer to other moral realists, perhaps another example of a non-naturalist moral realist is G.E. Moore, while Aristotelian Essentialism is an example of naturalistic moral realism.

versions of moral realism is important since Street points out that the non-naturalist view is the direct target of the Darwinian Dilemma.[161]

Street recognizes that evolutionary explanations are speculations, that is, they should be understood in conditional form as: "if the evolutionary facts are roughly as I speculate, here is what can be said philosophically" (2006: 112).[162] Although the conditional form of her argument strikes me as suspect (and thus an issue I will take up shortly), she is fairly confident in these explanations (broadly construed) and attempts to bolster this evolutionary explanation with empirical evidence. As she states, "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 the content" (2006: 117). For example, Street argues that evaluative judgments such as "the fact that something would promote one's survival is a reason to favor of it" and "the fact that something would promote the interests of a family member is a reason to do it" (2006: 114) are not only pervasive across time and cultures, but can also be easily explained by evolutionary biology—such judgments tended to promote survival and reproduction better than other evaluative judgments.

---

[161] See Street (2006: 111-12). Whether naturalist views fit in is a much more complicated matter, especially since she also argues that such views are impossible (Street 2006: 135-41). Note that this is very interesting, considering how using evolutionary ethics would seem like a naturalist argument (that is, some version of a pro-naturalist realism). In addition to Street's distinction between naturalist and non-naturalist views of moral realism, she includes another distinction between realism about practical reasons and realism about epistemic reasons. Practical reasons consider reasons for action and epistemic reasons consider reasons for belief. Although Street thinks her debunking argument could take care of both types of reasons, the Darwinian Dilemma focuses on realism about practical reasons. She also understands 'reasons' here as, "normative reason in the sense of a consideration that counts in favor or justifies some action" (Street 2006: 156, footnote 2). This means that it's not just an epistemic issue: it's not about being justified in believing x, but being justified in doing x; the moral judgment is about doing x: I ought (ought not) do x. It is an additional question whether I am justified in believing that I ought (ought not) do x. For further discussion on this point, see Street (2009).

[162] Here, she refers to Gibbard (1990: 30).

Street argues that basic evaluative tendencies in humans could have been a genetically inheritable trait and, as such, developed into a full-fledged capacity for evaluative judgment—that is, a capacity with the ability to use language and reflection to make judgments.[163] Through the mechanisms of kin selection and reciprocal altruism, Street explains how the selective pressures on the content of our evaluative tendencies (the proto-version of evaluative judgments) pushed man in the direction of making those evaluative judgments, which tended towards reproductive success and against those that decreased success.[164] According to Street,

> The details of how survival and reproduction were promoted will vary depending on the evaluative tendency in question. In the case of [some] judgment[s]…the rough explanation 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. The explanation of [other] evaluative tendencies…will be somewhat more complicated, drawing on the evolutionary theory of kin selection…and the biological theory of reciprocal altruism. (2006: 115-16)

So according to Street's explanation, there was a direct evolutionary influence on the content of the evaluative tendencies, and these tendencies had a major influence on the content of evaluative judgments. Although this makes the evolutionary influence on moral judgments indirect, Street maintains that our moral judgments are so "thoroughly saturated with evolutionary influence" that the content of our evaluative judgments would have been different if our evaluative tendencies had evolved differently (2006: 114).

Although Street does not use the terminology 'moral sense' in this evolutionary explanation of morality, I take her claims about the evolution of moral judgments to constitute such a view; that is, humans have an evolved innate mechanism that influences beliefs about

---

[163] See Street (2006: 118).

[164] For kin selection, Street (2006) refers to the work of Hamilton (1963) and (1964), Sober and Wilson (1998: Chapter 2), and Buss (1999: Chapters 7 and 8). For reciprocal altruism, Street refers to the work of Trivers (1971), Axelrod (1984), Sober and Wilson (1998: Chapter 2), and Buss (1999: Chapters 7 and 8).

morality. As Guy Kahane points out, "evolutionary debunking arguments commit us to claims about innateness" (2011: 112), so Street must presuppose an innate mechanism (the content of our moral judgments) or a 'moral sense', for her argument to get off the ground. In Street's discussion of the adaptive link account, she refers to evaluative judgments as an adaptive response produced by a "mechanism" of evaluative judgment—a "conscious mental state, subject to reflection and possible revision in light of reflection" rather than a "brute hard-wired physical mechanism" (2006: 128). But just as a brute hard-wired physical mechanism, like the reflex mechanism gets us to respond to the environment, the evaluative judgment mechanism (what I refer to as the moral sense in this debate) gets us to respond to the environment as well—it creates the experience of normativity, where we experience something as being called for or demanded by the circumstances.

The evolutionary claims Street relies on are much more controversial than she admits.[165] But even putting such controversy aside, it remains unclear what it means to "observe patterns in the actual content of human evaluative judgments". Street assumes that such observations are not only easily and accurately made, but also that such observations could be made about the content of past human beings' (or pre-humans') evaluative judgments to provide an evolutionary explanation. The issue of access—not only to the content of our own judgments but also the content of other people's judgments—arises. Artifacts and fossils alone are insufficient to account for the thoughts—for the *actual* content of evaluative judgments or tendencies—of past

[165] There are several types of arguments against Street's evolutionary explanation. Primarily, there are those who take the anti-adaptation stance—that morality (or evaluative judgments) is not an adaptation. These include, Gould (1997), Fodor (2000), Ayala (2010), Prinz (2008), and Machery and Mallon (2010).

human and pre-humans.[166] Street's detailed explanation of how our judgments became "thoroughly saturated with evolutionary influence" does not address this problem. Assuming that evaluative tendencies are inheritable traits that were 'selected for', it remains up for debate as to what the actual content of these tendencies, or judgments produced by those tendencies, would be. At this point, Street might argue that this problem is either too broad—that it is an issue for all of evolutionary psychology, which she is merely accepting, and thus not her problem to solve—or, that by accepting evolutionary psychology she is denying that this is a problem entirely. The evolutionary claims she presents, however, are the crux of her argument, and therefore, it is essential that she can explain how we have reliable access to the content of these judgments, especially to the judgments of past humans or pre-humans. It is likely that Street would stress the conditional form of these explanations, but then it is difficult to take the argument that follows from these evolutionary speculations seriously.[167]

By presenting her argument in the conditional form, Street intentionally tries to sidestep such objections to proceed with her argument against moral realism. So, assuming Street's evolutionary explanation is correct, what are the implications for morality? Street's answer is that it leads to an irreconcilable dilemma—the Darwinian Dilemma.

---

[166] And, if you take an anti-behaviorist line, you could even question our ability to use observations of present behavior as proof of the content of such judgments!

[167] In Justin Clarke-Doane's article "Morality and Mathematics: The Evolutionary Challenge", he argues the following: "the challenge does not depend on the genealogical speculation that our moral beliefs actually are the products of evolutionary forces. This speculation is redundant given the premise that if our moral beliefs were the products of such forces, then those forces would be 'non-truth-tacking'" (2012: 339). As he sees it, the conditional nature of the argument makes it one of mere counterfactuals and thus no actual genealogy is required (2012: 323-26). However, despite the conditional form in which Street (and others) present their arguments, they do in fact rely heavily on the genealogical speculations purposed and do not present their arguments in the weaker version that Clarke-Doane proposes. Additionally, I find it difficult to see what the purpose of an entirely counterfactual argument about morality could be since it merely acts as a mechanism to test our intuitions but does not lead to concrete conclusions on its own.

The Darwinian Dilemma rests on the idea that evolution shapes the content of our evaluative judgments. Since evolutionary forces have greatly influenced human evaluative attitudes, albeit indirectly, the realist must explain the relation between evolutionary forces and evaluative facts, between "the selective forces that have influenced our evaluative judgments on the one hand, and the independent evaluative truths that realism posits, on the other" (Street 2006: 121). Thus, like Ruse, Street's position centers on an explanatory issue. As we will see, Street's Darwinian Dilemma is an explanatory dilemma—one Street thinks arises in light of evolutionary considerations of morality (evaluative judgments/facts) and that leads to the debunking of moral realism.[168]

According to Street, when realists attempt to explain this relation they face an irreconcilable dilemma. Realists have two options: they can either (a) deny the relation between evolutionary forces and evaluative attitudes or (b) accept the relation between evolutionary forces and evaluative attitudes. [169] If realists chose the first option (a), then they are faced with two unacceptable positions: either a skeptical conclusion that most of our evaluative judgments are "off-track" (our evaluative judgments don't actually "track" independent evaluative truths), or they must accept the implausible claim that it is a complete coincidence that our evaluative judgments are in-line with independent evaluative truths. If realists chose the second option (b), however, they must provide an evolutionary account explaining how natural selection led to this truth-tracking ability in humans. According to Street, such an account would be unscientific, and thus, the realists again face an unacceptable conclusion.

[168] The explanatory nature of Street's argument appears to resemble Harman's challenge as well.
[169] Street defines evaluative attitudes as "states such as desires, attitudes of approval or disapproval…and consciously or unconsciously held evaluative judgments…" (2006: 110).

On the first horn of the dilemma (a), in which the realist denies the relation between evaluative judgments and independent evaluative truths, Street claims that the realist is led to an implausible skeptical conclusion because the selective forces that shaped evaluative judgments acted as distorting influences, pushing such judgments away from any independent evaluative truth. Street recognizes the possibility that evaluative judgments and evaluative truths could still match up; however, such correspondences would only be a matter of chance, a "happy coincidence" and a little too convenient for the realist. As such, she argues that the more reasonable assumption to make is that such connections would be unlikely, and thus, most of our evaluative judgments turn out to be "off-track." But since humans are clearly not hard-wired automata—according to Street's view, the mechanism of evaluative judgment is a "conscious mental state, subject to reflection and possible revision in light of reflection" (2006: 128)—then the process of rational reflection appears to offer some hope to the realist. Perhaps even these "off-track" judgments can be directed towards the evaluative truth through the process of reflective equilibrium. Street rules out this possibility, however, arguing that since the initial starting point of the reflective process is so far off-track (those evaluative judgments used to get the reflective process going, that are used to assess other evaluative judgments) that we can never get to the truth. As Street puts it, "the tools of rational reflection are equally contaminated by selective forces" (2006: 124). As such, there is another dilemma embedded in the first horn of the Darwinian Dilemma: if the realist denies the relation, then either (a)(i) our evaluative judgments are hopelessly off-track and we are faced with skepticism, or (a)(ii) our evaluative judgments are on-track via an unexplained miracle. If the challenge to the realist is an

explanatory issue, then neither of these options provides the realist with a sufficient explanation.[170]

Before tackling the second horn of the dilemma, it is important to point out some underlying philosophical assumptions about how we get or "track" the truth. By employing the concept of "truth-tracking", Street relies on Robert Nozick's view from his book *Philosophical Explanations* that knowledge is tracking the truth.[171] According to this theory, we "get" the truth when we stand in the right relationship to the world. According to Nozick, "knowledge is a particular way of being connected to the world, having a specific real factual connection to the world: tracking it" (1981: 178). Thus to know something about the world is to say that if *p* is true then you not only believe *p* but if it were false you would not believe it and if it were true you would believe it.[172] Like any philosophical view, there are of course objections to this take on knowledge.[173] I do not want to address these arguments here but merely point out that if the truth-tracking view of knowledge fails to withstand scrutiny, then Street's argument would start to fall apart as well. Street, however, accepts the truth-tracking hypothesis at face value.

On the second horn of the dilemma (b), the realist asserts the relation between evaluative judgments and evaluative truths, but since this is an explanatory dilemma, the realist must also explain (or provide an account of) this relation they are asserting. Street argues that if the realist takes this position, then the realist must accept a "tracking account" to explain the relation, a

---

[170] Of course, there is the glaring question of how anyone could know any of this and that Street's position possesses the only scientifically valid position when this is all mere speculation. It is also unclear how Street knows that the starting point would be so off-track that reflective equilibrium would be impossible. I will return to these issues in Chapter Four and Five.
[171] Street refers to Nozick (1981: 337) in her footnotes (2006: 159). Nozick's view is also typically considered an early form of reliabilism.
[172] See Nozick (1981: 178).
[173] Note that there are, of course, some issues with Nozick's truth-tracking view. For discussions of the various problems, see Vogel (2000), Sosa (1991 and 1999), and Williamson (2000).

scientific explanation of the relation in which the ability to grasp evaluative truths promoted fitness (survival and reproduction)—evaluative judgments promote fitness because they are true. According to Street, this scientific explanation, however, is inadequate when compared with a rival scientific explanation, the "adaptive link account". According to the adaptive link account:

> [T]endencies to make certain evaluative judgments rather than others contributed to our ancestors success not because they constituted perceptions of independent evaluative truths but rather because they forged an adaptive link 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: 127)[174]

So, according to the adaptive link account, evaluative judgments promote fitness because they get us to respond to the environment in ways that lead to reproductive success.

Street provides an analysis of the scientific adequacy of these two accounts, examining their explanatory power in terms of (1) parsimony, (2) clarity, and (3) ability to shed light on the explanandum in question, that is, why humans tend to make certain evaluative judgments. This debate over which evolutionary explanation best captures the phenomena raises an important question about non-evaluative judgments (or facts), and the implication of evolutionary explanations on realism about things other than morality (e.g., mathematics, science, etc.).[175] For now, I will focus on Street's arguments for evaluative judgments, but I will come back to this issue in the last section of this chapter. On all three accounts of scientific adequacy (explanatory power), Street argues that the adaptive link account fairs much better than the tracking account. Briefly, the tracking account is less parsimonious because it adds "truth" into the explanation, it

[174] Note that this claim is based on Blackburn (1993).
[175] It also related to an issue mentioned earlier in this chapter regarding "best" explanations and from Chapter One regarding the idea that there is *one* "best" explanation. I will come back to this issue in Section B of this chapter.

is less clear since exactly how "truth" promotes fitness is extremely vague, and it fails to shed light on why humans tend to make certain kinds of evaluative judgments.[176] Street claims that,

> The power of the adaptive link account is that it exposes much of this seeming unrelatedness as an illusion; it illuminates as 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. (2006: 134)

As such, Street concludes that the adaptive link account is the best explanation of evaluative judgments, which means that the tracking account is unscientific.

If Street is correct, then the realist is, yet again, faced with an impossible dilemma: if the realist asserts the relation then either (b)(i) they accept the tracking account, which includes evaluative truths but is unscientific, or (b)(ii) they accept the adaptive link account, which is scientific but fails to include evaluative truths. If the challenge to the realist is an explanatory issue, then neither of these options provides the realist with a sufficient explanation.

In some ways, Street's Darwinian Dilemma is reminiscent of Plato's Euthyphro problem.[177] Very briefly, the Euthyphro problem can be presented as the following dilemma. Some claim that an action is good (or right) because of God; I will call this position the "God-Realist". To explain why an action is good because of God, the God-Realist can claim that either (1) an action is good because God says it is good, or (2) God says the action is good because the action is good. On the one hand, if (1) is the correct explanation, that is, good is merely what God says is good, then such an explanation is uninformative because it is a tautological explanation: "good is what God says is good because God says it is good". On the other hand, if

[176] Note that (1) is similar to Ruse and his redundancy argument; and (2) is like Harman's challenge. I will come back to these issues in the last section of this chapter.
[177] Although I developed this application of the Euthyphro problem onto Street's Darwinian Dilemma independently, you can also find an in-depth discussion of this type of reading of Street's Darwinian Dilemma in Rabinowitz (2012: 3-21).

(2) is the correct explanation, then there is an independent standard that explains what is good and that explanation does not require the existence of God.[178] Consequently, the God-Realist faces an insurmountable dilemma, since both options (1) and (2) fail to explain why an action is good because of God—option (1) is uninformative and option (2) provides an explanation that doesn't require the existence of God, which is no explanation at all.

Street's Darwinian Dilemma attempts to pose a similar problem for the moral realist. If the moral realist accepts (a) and denies the relation between evolutionary forces and evaluative attitudes, then the only way to explain the connection between the two is a miracle or random fluke, which is uninformative. If the moral realist accepts (b) and asserts the relation between evolutionary forces and evaluative attitudes, they are left with a scientific explanation (the adaptive link account) that doesn't require the existence of evaluative facts. As a result, the moral realist is in the same position as the God-Realist, facing the option of an uninformative explanation or no explanation at all.

While Plato leaves it at that, a dilemma (Socrates doesn't tell Euthyphro to stop believing in God but to think about it), Street takes a different course. Confronted with this dilemma, Street argues boldly that the only reasonable position is to give up moral realism entirely and instead adopt an antirealist position that avoids this problem. According to Street, the antirealist "sidesteps" the Darwinian Dilemma. Street defines evaluative (moral) antirealism broadly as the position that evaluative facts (or truths) are a function of our evaluative attitudes (2006: 152). Although there are different antirealist views, Street focuses on an antirealist constructivist position as her example. The antirealist constructivist holds the view that evaluative truths are merely a function of our evaluative attitudes, which have been strongly influenced by Darwinian

[178] Plato talks of "piety" rather than goodness. However, it is generally agreed upon in the literature that piety is something like moral goodness.

selective forces (a causal explanation). Unlike the realist who takes evaluative truths to exist prior to evolution (independently of it), the antirealist takes evolutionary causes to exist prior to evaluative attitudes on which evaluative truths are based. As Street explains,

> Both accounts offer an explanation of why it is no coincidence that there is a significant overlap between evaluative truth and the kinds of evaluative judgments that natural selection would have pushed is in the direction of. The difference is that the antirealist account of the overlap is consistent with science. (2006: 154)

As such, the antirealist avoids the dilemma because the antirealist can assert the relation between evaluative judgments and evaluative truths while explaining the relation with the adaptive link account (or whatever scientifically viable account is available)—the antirealist can accept the second horn of the dilemma without conflict because the evaluative truth is dependent on the evaluative attitudes that have an evolutionary explanation.[179]

Street's antirealist constructivist view is reminiscent of Ruse's view that morality is merely a collective illusion. As previously discussed, Ruse argues that since morality is a function of our evolved human nature, and thus subject to the contingencies of evolution, it cannot be objective—in other words, moral realism must be false. However, since human beings share in the same evolutionary history, we also share the same illusion of morality—the moral judgments we make are based in our evolved human nature, which makes them feel objectively true when they are merely true to us. Street reaches a similar conclusion. The antirealist takes evolutionary causes to exist prior to evaluative attitudes from which evaluative truths are based—in other words, the moral judgments humans make are based on our shared evolved human nature or moral sense. For Street, moral judgments are not objectively true, but they are not entirely relative either—their truth is dependent on us. Although Street and Ruse appear to present different debunking arguments, they certainly reach very similar conclusions.

[179] See Street (2006: 153-54).

As Street sees it, evolutionary considerations force the realist into an irreconcilable dilemma—the Darwinian Dilemma—while the antirealist can easily accept such evolutionary considerations without issue. The evolutionary origins of evaluative judgments debunk (undermine) moral realism by posing an explanatory challenge that Street argues the realist cannot reasonably overcome but that the antirealist can. Going back to Harman's challenge, we can see that Street's debunking argument carries with it a similar explanatory challenge, and like Ruse, she takes the additional step of concluding that such a challenge is impossible for the realist to overcome, leaving antirealism as the only suitable option. Although Street and Ruse are similar in this respect, their debunking arguments still have different aims. In Joyce's discussion of debunking arguments, he describes Street's Darwinian Dilemma advocating for the thesis that "all moral judgments are false insofar as they involve a claim to objectivity" (Joyce 2016: 2). This thesis undermines moral realism in the sense that it "undermines any metaethical theory according to which there are objectively true moral claims" (Joyce 2016: 3). Here, Street is not like Ruse, who includes an additional thesis that all moral judgments are false.[180]

Under the "standard" reading of debunking arguments, Street is classified as an epistemological debunker—a debunking argument that seeks to challenge the epistemological standing of moral beliefs—because she takes aim at the moral sense's capacity to grasp the truth, thereby questioning the ability to be justified in believing "moral facts". But, it is clear after this analysis that she takes a further step by denying that there are any objectively true moral claims, thus making her view closer to a case of ontological debunking like Ruse's, challenging the ontological status of moral beliefs and not merely their epistemic justification.

[180] See Joyce (2016: 2-3).

There is one last additional difference between Street and Ruse that must be noted. Street concludes that the Darwinian Dilemma represents not just an evolutionary argument against moral realism, but also a more general causal argument against the realist. As such, she argues that you could substitute any other scientifically viable causal explanation for the Darwinian hypothesis and a similar problem will arise for the realist. My focus is on the evolutionary genealogies, but Street's view that there is nothing special about evolution that plays a role in debunking evaluative realism is important. For if she is correct, she would not only have to accept a much broader antirealist view than I think she would be willing to accept, but she would also be forced to embrace a radical epistemological skepticism—a conclusion that I will address after discussing the final debunking view presented by Joyce.

b. Epistemological Debunking: Richard Joyce's Epistemological Challenge

The last debunking argument to consider is Joyce's epistemological debunking argument. Joyce's debunking argument makes the weakest claims against moral realism. He argues for an innate moral sense that developed from the process of evolution via kin selection. The evolution of an innate moral sense then undermines morality, leading to moral skepticism. Joyce's debunking argument is a form of epistemological debunking because it targets the epistemological justification of moral beliefs. Joyce first presented his evolutionary debunking argument in his book *The Evolution of Morality*.

Just like Street, Joyce offers up a "caveat" regarding the evolutionary explanations that he uses to establish an innate moral sense, recognizing that the scientific research that supports such explanations is still in its formative stages.[181] Although Joyce is confident that the evolutionary

[181] As he says, "this result will be provisional and to a degree speculative, since the present evidence does not warrant answering the question ["is human morality innate?"] in either a positive or negative way with any confidence" (Joyce, 2006: 2).

picture he paints is fairly accurate (or at least on the right track), he presents his argument in the following conditional form: if morality is innate, does the evolutionary genealogy of morality vindicate morality, thereby avoiding moral skepticism and supporting some form of moral realism, or does it debunk morality, undermining the authority of morality?[182] By presenting his argument in this form, Joyce, like Street, attempts to avoid any scientific objections to his position, pushing aside the contentious debate over the innateness of morality. I will address the ramifications of such a move later in this chapter, but first let's look at how Joyce describes an innate moral sense and how he argues that it debunks, rather than vindicates morality.

According to Joyce's view presented in *The Evolution of Morality*, pro-social helping behaviors and the neurological mechanisms that furnished early humans with altruistic emotions and motivations developed into the ability to make moral judgments. Thus, Joyce defines the moral sense as an innate faculty for making moral judgments. Joyce describes moral judgments as speech acts (as a linguistic convention) that express two types of mental states (beliefs and attitudes), which reflect agents' approval or disapproval. Here, Joyce attempts to reject both pure non-cognitvism and pure cognitivism and instead proposes a "quasi-cognitvism" by claiming that a person holds two mental states when she makes a moral judgment—a belief state and a connotative non-belief state.[183]

Moral judgments, however, are not merely an expression of an agent's personal feelings and attitudes. Rather, moral judgments purport to possess moral authority—they carry a practical clout that others ought to consider regardless of their own views or interests. The primary subject matter of these judgments is interpersonal relations (e.g., values of reciprocity, refraining from harming others and avoiding too much selfishness) that promote society and thus govern our

[182] See Joyce (2006: 2). Note here he quotes Ruse as a debunker.
[183] See Joyce (2006: 51-57).

behaviors as such. Additionally, the related notions of desert and justice also appear as a requirement for making moral judgments, i.e., a system of punishment and reward. Emotions of guilt or a "moral conscience" also play a role in regulating our behaviors.[184] Joyce recognizes that these features of moral judgments are by no means a set of the necessary and sufficient conditions but rather provide a general outline of the nature of moral judgments:

> It is doubtful that our concept of a *moral judgment* is sufficiently determinate to allow such exposition. The sensibly cautious claim to make is that so long as a kind of value system satisfies *enough* of the above, then it counts as a moral system. (2006: 71)

With this basic understanding of moral judgments (and morality more broadly construed), Joyce then proceeds to complete his description of the moral sense.

Joyce argues that human's innate moral sense is a combination of various psychological and neurological mechanisms that evolved through the process of natural biological selection.[185] Here, Joyce relies on the work of many evolutionary psychologists and sociobiologists to develop this view of the moral sense. Some of these include (but are not limited to) Hamilton (1964), Trivers (1971 and 1985), de Waal and Luttrell (1988), Alexander (1987), and Sober and Wilson, D.S. (1998). Joyce is adamant that the proposed hypothesis for the moral sense is a serious scientific hypotheses:

> …that the human moral sense is the product of biological selection—is something to be taken seriously; it is more than a mere just-so story. (2006: 73)

Furthermore, from this description of the moral sense, it is clear that he is a Darwinian Moral Nativist. Joyce's moral sense contains many elements of Darwin's view of the moral sense as well as the sentimentalist tradition that influenced it.

---

[184] See Joyce (2006: 64-70).
[185] That is, rather than being discoverable in any particular gene or one portion of the brain.

According to Joyce, the environment and culture influences the content of moral thinking, although it is likely that an innate moral sense puts some limitations on such content. Joyce thinks this helps explain the universal character of moral thinking; however, for Joyce the existence of an evolved innate moral sense supports moral skepticism, a moral agnosticism in which there are no epistemic justifications for or against any particular morality (or moral judgments).

Joyce uses a thought experiment and an argument from analogy to explain why the existence of an evolved innate moral sense (as he has described) leads to our moral judgments being unjustified—that is, why an evolutionary explanation of morality "debunks" it. Joyce starts with the claim that if every belief has a causal etiology, then in some cases the "*knowledge* of the beliefs origins can undermine it" (2006: 179).[186] Regardless of whether the belief is true or false, the knowledge of a belief's origin may reveal that you are not *justified* in holding that belief.[187] To illustrate this point, Joyce asks us to imagine a case in which you find out that your belief about X was caused by a "belief pill" that makes you form certain beliefs and makes you forget that you took this pill, or that there are such pills. Knowledge of this would raise doubts about whether you were actually *justified* in believing X anymore—it doesn't show that X is false, but it should make you doubt that you should believe X. Unless you find some other reasons (or

[186] Joyce argues, "it would be a ridiculous theory that implied that knowledge of a belief's etiology automatically undermines the confidence one should have in that belief" (2006: 179). Therefore, only some beliefs will be in question.

[187] Furthermore, he claims "knowledge of a belief's genealogy could show the belief to be false only if the belief implies a contrary genealogical story. For example, the belief that no beliefs are innate would certainly be shown to be false if we were to discover that it was an innate belief" (Joyce 2006: 179-80). But these cases are rare, so the issue concerns the justification for holding a belief rather than the truth (or falsity) of the belief itself. In a footnote, Joyce points out that the reason this is important is because it shows that the genetic fallacy is wrong (or better, not being committed here) (Joyce 2006: 242). This is something I will discuss in more detail in Chapter Five.

evidence) to believe X, then you should stop believing in X.[188] As mentioned earlier, this is what Kahane (2010) refers to as an "undermining defeater", whereby the belief undermines the justification of the belief (although not the truth of that belief itself).

Joyce then turns to the case of evolution (natural selection) and moral beliefs. If moral beliefs—our concepts such as our obligations, virtue, fairness, etc.—are the result of natural selection, if they come from an evolved innate mechanism (the moral sense) that "comes prepared to categorize the world in morally normative terms", then those beliefs are not justified (2006: 181). Since there is no evidence that the moral sense is likely to have produced true beliefs, it is similar to the belief pill in producing unjustified beliefs. Just as in the case of the belief pill, once we know about the genealogical story of our moral beliefs, we are no longer justified in believing them; we should hold off believing them until we have some other evidence for those moral beliefs. Joyce sums up his argument as follows:

> We have an empirically confirmed theory about where moral judgments come from (we are supposing)...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. (2006: 210)

The existence of an evolved innate moral sense debunks morality—for Joyce, the evolutionary genealogy of moral beliefs (an evolved innate moral sense) represents a belief-formation mechanism (or process) that produces unjustified beliefs.[189] Joyce's argument, if successful, undermines the justification of moral judgments, and therefore, it is an example of epistemological debunking.

[188] See Joyce (2006: 179-80). Note that in the case of morality, there is no other reason offered to believe X. In fact, that is the challenge he thinks his view presents to the realist.
[189] See Joyce (2006: 212).

Joyce considers several objections to his analogy. One of the most serious objections is that the belief pill analogy fails because his proposed view of an evolved moral sense does not include innate beliefs about morality, but rather a belief forming mechanism. Similar to Street's view about the evolved capacity to make evaluative judgments, Joyce responds by claiming that since moral concepts (e.g., morally normative terms such as fairness, obligation, virtue, property and desert) originate from the moral sense and the moral sense is an evolved mechanism, the moral concepts produced by it have an evolutionary etiology. These moral concepts then influence beliefs and, thereby, produce moral beliefs that we would otherwise not have without those concepts. Thus, like Street's claim that evaluative judgments are "thoroughly saturated with evolutionary influence", Joyce argues that even if the moral belief itself was not 'selected for', it still possesses an evolutionary history via the evolved moral sense. As such, he thinks the belief pill analogy can be easily modified to handle his case.[190]

Another significant objection Joyce considers is that the evolutionary process produces true beliefs, and therefore, unlike the belief pill example, we need not worry that our beliefs are wrong. Again, Joyce's response is also similar to Street when facing this kind of objection. Joyce argues that in the case of beliefs about mathematics and science, we have good reason to believe that those beliefs evolved to be true. However, we have no good reason to believe that beliefs about morality evolved to be true. So while we are safe to maintain our beliefs about mathematics and science, we are not when it comes to beliefs about morality.[191]

I will address both of these responses in more detail below. For now, however, the central issue that emerges is how we determine which processes of belief formation are justified and which processes are not. Here, Joyce is clearly working on the intuition that his case of the belief

[190] See Joyce (2006: 180-82).
[191] See Joyce (2006: 182-84).

pill (and thus any similar case) is not a justified means of achieving a belief. But why not assume that the belief pill is a reliable process? For Joyce, this would be like assuming that the "belief-pill pusher" was benign so the believer would need good grounds for thinking this. Furthermore, he argues that in most cases, such evidence would not be forthcoming since the believer needs to be *certain* that the "belief-pill pusher" is benign—thinking they *might be* is not a strong enough reason to accept that they are benign.[192] Joyce's claims are reminiscent of Cartesian epistemological skepticism—if beliefs are caused by an untrustworthy process (e.g., a malicious demon), then we cannot be certain that they truly represent the world and we should set them aside. But why should the default be a Cartesian malicious demon? It would seem that in most cases, we can and do trust what we believe since we are fairly good at recognizing cases in which we cannot trust beliefs (e.g., cases of optical illusions or faulty sense perceptions).[193] And, why should the default for justification be certainty? Joyce simply refers to Harman's view and states that "a mere 'might' won't cut the epistemological mustard" (2016b: 12). This answer is insufficient, however, since Joyce never fully explains what counts as the required standards for justification. So, it is unclear why it cannot be the other way around. That is, why wouldn't the believer need ample evidence to support the idea that the belief-pill pusher was, in fact, *not* benign but perhaps malicious or unconcerned about the truth? Alternatively, under certain circumstances, the believer may be justified in holding beliefs that *might* come from a reliable belief-forming process or a belief forming-process that only produces *somewhat* certain beliefs.

Here, Joyce would likely accuse me of endorsing some unlimited form of epistemological conservatism, granting *prima facie* justification to all extant beliefs—that is, the fact that one

[192] See Joyce (2016b: 12-13).
[193] Even the meditator in Descartes's *Meditations* is hard pressed to find a reason for radical skepticism and recognizes that the epistemological challenge created by the malicious demon is only a tool used in a thought experiment and not real.

believes something provides a reason or some justification to believe it. I agree, however, with Joyce that epistemological conservatism is only acceptable up to a certain point, and that such *prima facie* justification should only be endorsed "if it is reasonable to think that we are pretty good believers—only if, in other words, it is reasonable to think our beliefs are likely to be reliable indicators of the facts" (Joyce 2016: 14). The problem is that in most cases we are pretty good believers, and thus we should only question the justification of a belief when we have good reason to believe that we are *not* pretty good believers—when it is reasonable to think that our beliefs are *not* reliable indicators of facts. Joyce's belief pill analogy does not present a case where it is reasonable (or at least that is obvious) to think that our beliefs are not reliable indicators of the facts. We need additional reasons to think the belief-pill pusher is *not* benign (i.e., like Descartes' malicious demon) before rejecting the beliefs it produced. Although Joyce denies this, some of the burden of proof lies with him (or any debunker) to show why we have good reasons for doubting any beliefs and, more specifically, moral beliefs. Joyce thinks that the burden of proof lies with the moral realist to provide reasons for thinking that moral beliefs stemming from an evolved moral sense are true; however, Joyce also needs to provide a clearer standard for what would count as good reasons for what would make beliefs, and in particular moral beliefs, true and/or justified. So, at best his analogy is incomplete, and therefore, the comparison between beliefs produced by the pill and the beliefs produced by an evolved moral sense starts to fall apart.

The epistemic problem presented by the belief pill analogy also assumes (a) that justification is external (rather than internal) and (b) that we are required to stop believing X if we are unjustified. Both of these claims are debatable, and even Joyce himself eventually denies

(b) when he argues, like Ruse, that we do not need to revise our everyday morality, even though it would be based on a set of unjustified beliefs.[194]

Although Joyce, like Street, also presents his argument in conditional form, the epistemological problem he presents relies on a particular set of evolutionary explanations. Just as in the case of Street, the evolutionary explanations presented are still up for debate. And yet, the center of Joyce's argument—that if the reasons for belief (or the process of belief formation) are unreliable, then the beliefs we hold are unjustified—depends on the accuracy of these evolutionary explanations. Joyce attempts to circumvent this problem by claiming, "I am not putting this hypothesis forward as *true*. It is a hypothesis that is plausible, coherent, and testable—and its truth remains to be established" (Joyce 2006: 134).[195] To say that the empirical evidence is forthcoming begs the question. Despite these issues, Joyce proceeds with confidence.

Joyce's debunking argument takes direct aim at the epistemological metaethical theories that attempt to show that moral beliefs are justifiable irrespective of the ontological status of the facts that make them so. To clarify his debunking argument, Joyce considers various epistemological theories that he classifies as reliablism, conservatism, coherentism, foundationalism and evidentialism, each of which provides a different kind of justification. For example, if one adheres to evidentialism, then the debunking can be understood as the epistemological challenge that "until a plausible justifying account is presented, moral judgments must be considered unjustified" (Joyce 2016: 13). But if one adheres to reliablism, then the epistemological challenge is "until plausible evidence of a reliable process is presented, we do

[194] See Joyce (2006: 223-30).

[195] Recently, Joyce, like Street, claims that the epistemological challenge to moral realism is not dependent of evolutionary explanations—instead "the evolutionary perspective is, strictly, dispensable" (Joyce, 2016b: 2). As in Street's case, my focus is on the evolutionary arguments, so I am not going to consider the other cases. However, it is interesting that both Joyce and Street are willing to make this concession.

not know whether moral judgments are justified" (Joyce 2016: 13).[196] For Joyce, however, regardless of which view of epistemic justification one holds (here, I think he has the moral realist in mind), moral beliefs/judgments are not justified because our evolved innate moral sense strongly influences these beliefs about morality in such a way that we don't know whether or not we are justified in holding moral beliefs no matter what form of justification is employed. This leads Joyce to moral skepticism, "meaning that the evidence favors neither a proposition nor its negation, and thus one should choose (or is required) to reserve judgment on the matter, at least until new evidence shows up" (2006: 223).

In *The Evolution of Morality*, this epistemological debunking leads Joyce to consider the possibility that his view of "agnostic" moral skepticism just described could be understood as a form of Kalderon-style fictionalist error theory in which no moral judgments are epistemically justified.[197] Very briefly, Kalderon argues "the fictionalist maintains that [moral] sentences…express propositions that represent the putative subject matter; but he maintains, as well, that their acceptance is not belief" (Kalderon 2005: 112). The acceptance of a moral sentence is not a belief, but rather a non-cognitivist attitude towards the content of the moral sentence. In his article "Moral Fictionalism", Joyce describes his own fictionalist error theory in which "fictions can produce real emotions which have motivational force. Even if such beliefs are false, they are valuable for prudential reasons—one 'ought' to act in accordance with alleged moral requirements for fear of punishment, the benefits of cooperation, etc." (Joyce 2005: 289). On this view, we have no justification for believing in any moral propositions since we cannot

[196] In later work Joyce totally dismisses coherentism as a possibility because he thinks it is an untenable view. See Joyce (2016: 5).
[197]According to J.L. Mackie's view of error theory "moral utterances are typically assertions…but they are systematically untrue, since there are no moral properties to make them true" (Joyce: 2015). Joyce calls this position an atheistic moral skeptic rather than the agnostic moral skeptic he wants to be (2006: 223).

know whether or not moral propositions are true or untrue; however, practically speaking, our everyday "moral" lives need not change.[198] It is not clear, however, that if we accept Joyce's conclusion of moral skepticism or a version of fictionalism, that morality, as we know it, can stay intact. As Joyce pointed out, if the justification for a belief is undermined, then we should doubt that belief unless we can find other evidence for believing it. And, if we do not find any other evidence for that belief, then we should stop believing it. Just as Ruse fails to provide an adequate account of his anti-relativism, Joyce does not address this issue fully, and it is something I will discuss below.

In his more recent work, Joyce attempts to clarify his debunking argument. He argues that his epistemological debunking does not undermine moral realism per se (that is, it does not support antirealism), but rather that he is only arguing "all moral judgments lack justification" (2016: 2). Joyce argues that moral realism (that moral facts exist objectively) is compatible with his "skepticism" because even if all moral judgments lack justification it is still possible that there are objective moral facts—it's just that we do not know one way or the other if they do or do not exist (or which is which). As such, Joyce tries to retract his earlier error theoretic proposals:

> I did once suggest that the label "error theory" might be expanded so as to denote the disjunction of [all moral judgments are false] and [all moral judgments lack justification] (Joyce 2006: 223); and I then proceeded to attempt to establish [all moral judgments lack justification] by an EDA and called it an "error theoretic" conclusion. I now think that this suggested expansion of the label was injudicious. (2016a: 3)

[198] Joyce (2006: 227-28). There seems to be something like Pascal's Wager going on here but with moral beliefs rather than religious beliefs. That is, Joyce is suggesting that we believe in morality even though we cannot know if it is real because it will have the best outcome. If so, then it would surely face the same criticisms as Pascal's Wager. For a discussion of such criticism, see Blackburn (2009).

Joyce insists that his debunking argument and the conclusions he reaches are much more modest and more vague than his critics characterize him as presenting.[199]

Even if Joyce's argument does not directly undermine moral realism (that is, it is not a form of ontological debunking), he does make a strong metaethical claim that affects epistemological metaethical theories that attempt to provide a "realist" type account of morality—that is, those that claim we are justified in holding certain moral beliefs. Furthermore, Joyce is not convinced that we can ever find such justification:

> There seems no ground for optimism that the evolutionary process that produced the moral faculty will turn out to be the benign sort that happens to track truths (even if it lacks this as its adaptive function). (2016a: 15)

At one point he even compares moral beliefs to beliefs about fairytales and ghosts:

> It is a matter of maintaining an open mind about whether there exists anything that is morally right or wrong, 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 (as John Mackie argued in this case). (2006: 181-82)

He concludes that we need to come up with alternative ways for determining whether we should believe or disbelieve moral propositions and that these reasons for belief (or disbelief) need to be more reliable.[200]

As a result, Joyce claims that his position does not amount to a *direct* attack on moral realism.[201] The kind of skepticism he proposes, however, clearly leans towards antirealism—his view has antirealist sympathies and is more antithetical to the realist than he wants to admit. Joyce's agnosticism is akin to a religious agnosticism, which theists often argue still amounts to

---

[199] Also see Joyce (2016b). Here, Joyce identifies several philosophers he claims to have mischaracterized his view, including Kelby Mason (2010), Scott James (2011), and E.J. Wielenberg (2010).
[200] See Joyce (2006: 210).
[201] I assume he is trying to avoid the problems associated with making stronger ontological claims like Street or Ruse do.

a form of atheism. Most religious agnostics view God's existence as an epistemic problem, claiming that since it is impossible to know whether God does or does not exist, withholding belief about God is the only rational position to take. However, to withhold a belief about God means that one does not hold the belief that God exists. For the theist, not holding the belief that God exists is atheism despite the reasons why one does not hold the belief. Similarly, the moral realist, like the theist, might argue that Joyce's moral agnosticism is tantamount to moral antirealism. Withholding the belief in the existence of moral properties (or whatever moral realism proposes to exist) means that one does not hold the belief that moral properties exist. Like the theist, the moral realist would define this position as antirealism since not holding the belief that moral properties exist is antirealism despite the reasons why one does not hold the belief. [202]

Joyce tries to argue that his moral skepticism is just a "challenge" the realist must face, but at the same time he does not think the realists can ever meet it. And as such, he argues that the burden of proof lies with the realists. "A Darwinian analysis of moral belief should have the same effect of arousing suspicion about morality, to the extent that the burden lies on the shoulders of the anti-skeptic to articulate a persuasive defense" (2016a: 15).

As a result, Joyce ends up in a similar position as Street by presenting what he sees as an irreconcilable dilemma for the realist. What neither Joyce nor Street are willing to accept, however, is that if they are correct about the impact of evolution on belief formation and our truth-tracking abilities, there is a much broader epistemological dilemma *they* must face—*skepticism and antirealism would not be limited to morality.*

[202] Joyce faces another pragmatic issue. How can we be motivated to act on these moral beliefs if we are skeptics about them and it's recommended that we should not endorse them? This contradicts Joyce's' own view that we need not revise morality.

B. Global Skepticism and Debunking Non-evaluative Beliefs

In "The Theory of Evolution in its Application to Practice", Sidgwick describes how the evolutionary development of an innate moral sense may lead some towards a general moral skepticism, which is what cotemporary debunkers like Ruse, Street, and Joyce have all concluded. Sidgwick, however, argues that not only is moral skepticism unfounded, but that such skepticism would ultimately include too much. According to Sidgwick, moral skepticism is unfounded because there is no obvious reason to assume that the evolutionary process leads to an untrustworthy or unreliable moral faculty. According to Sidgwick,

> It is obviously absurd to make the validity or invalidity of judgments depend on the particular stage in the process of development at which the class of judgments first made their appearance; especially since it is an essential point of the Evolution-theory to conceive this process as fundamentally similar in all its parts. (1876: 54)

But worse, if the process of evolutionary development did lead to an untrustworthy and unreliable faculty such as the moral sense, then it is equally possible that all such evolved faculties of the mind are also untrustworthy and unreliable. He argues the following:

> But surely there can be no reason why we should single out for distrust the enunciations of the moral faculty, merely because it is the outcome of a long process of development. Such a line of argument would leave us no faculty stable and trustworthy: and would therefore end by destroying its own premises. (1876: 54)

As Sidgwick points out, this type of global skepticism—which would include mathematical and scientific beliefs—is incoherent. If we accept that all evolved faculties of the mind are untrustworthy and unreliable, then the faculty by which we know that evolved faculties are untrustworthy and unreliable is also unreliable—we cannot know, with any certainty, that we cannot know. This, of course, is simply the problem that any global (universal) epistemological skepticism will face. Sidgwick merely aims (in this discussion) to show how the moral skeptic

who uses evolutionary explanations to undermine morality is faced with a much bigger epistemological problem—one he thinks no one would readily accept.[203]

Sidgwick was correct that the moral skeptics generally do not embrace global epistemological skepticism. In the contemporary debunking arguments, Street and Joyce attempt to maintain a form of moral skepticism while rejecting the global epistemological skepticism that Sidgwick alerted us to.[204] Joyce argues that epistemological skepticism about moral (or evaluative) beliefs is warranted whereas epistemological skepticism about mathematical and scientific beliefs is not warranted even though both types of beliefs are products of an evolutionary process. Similarly, Street argues that epistemological skepticism about common sense objects—"facts about fires, predictors, and cliffs" (Street 2006: 130)—are not warranted even though these types of beliefs are products of an evolutionary process. According to Street and Joyce, the evolutionary explanation of mathematical, scientific, and common sense object beliefs require truth (or truth-tracking) in order for such beliefs to have promoted fitness, unlike the case of moral beliefs where the truth of the belief was irrelevant to the promotion of fitness. Although Street and Joyce (as well as others) think they are able to avoid global skepticism by making these distinction they are ultimately unsuccessful, and Sidgwick's original criticisms still apply.[205]

Joyce argues that the moral debunking argument does not apply to other beliefs that may also be innate and have an evolutionary explanation. For these types of beliefs, which include mathematical and scientific beliefs, Joyce argues that there is no evolutionary explanation that is

[203] Sidgwick deals with skepticism in other places. See Sdigwick (1874/1981: 509).
[204] I do not know of any place that Ruse addresses this issue, but I assume he would take the same line of argument as Street and Joyce.
[205] Others who have made similar arguments include Sinnott-Armstrong (2006) and Crisp (2006).

independent of the truth. In other words, the truth of such beliefs is necessary for such beliefs to promote fitness. Putting aside the question of whether such beliefs (or mechanism for those beliefs) were actually selected for, and thus an adaptation rather than a by-product or spandrel, Joyce's claim is that mathematical and scientific beliefs must be true in order to be useful and thus promote fitness, whereas moral beliefs need not be true in order to promote fitness.

Joyce uses the example of simple arithmetic, e.g., 1+1=2. According to Joyce, if there is an innate faculty for simple arithmetic, that is, if humans are "hard-wired" to believe that 1+1=2, the genealogical explanation of this belief (and others similar to it) does not undermine mathematical realism in the way that the genealogical explanation of moral beliefs are undermined by their genealogy (the mathematical beliefs are still justified despite their origin). In the case of simple arithmetic, Joyce claims that there is no way to provide an evolutionary explanation—how the belief enhanced fitness and thus was 'selected for'—without the background assumption that such beliefs are true. Joyce uses the example of being chased by three lions to illustrate his point:

> Suppose you are being chased by three lions, you observe two quit the chase, and you conclude that it is safe to slow down. The truth of "1+1= 2" is a background assumption to any reasonable hypothesis of how this belief might have come to be innate. (2006: 182)

Here, Joyce is assuming that such beliefs like "1+1=2" are "eternal, necessary truths" since "the environment can't suddenly change to make "1+1=3", and as such the truth of such beliefs is built directly into the evolutionary explanation in a way that it is not in the case of morality. According to Joyce, we could imagine the environment would change to make killing one's

offspring good, rather than bad; however, we cannot imagine an environmental change that would make "1+1=3" true thereby making "1+1=2" false.[206]

Joyce's argument concerns counterfactuals; in the case of morality, if the environment had been otherwise, humans would have developed a totally different morality, but in the case of simple arithmetic, even if the environment had been otherwise, simple arithmetic would be the same. But is this correct? Justin Clarke-Doane challenges this assumption in his article "Morality and Mathematics: The Evolutionary Challenge". Clarke-Doane argues that is not only possible but also intelligible to imagine mathematical truths being very different just in the same way that we can imagine moral truths being very different. Very briefly, he argues that Joyce's example of simple arithmetic appears to require corresponding first-order mathematical truths about the surroundings in order to guarantee the truth of those beliefs and to provide an intuitive reason why believing "1+1=2" was more advantageous than believing "1 +1=0".[207] However, many of even our simplest mathematical beliefs (he uses the example of elementary geometrical hypotheses) could still be believed, and even promote fitness, without requiring corresponding first-order mathematical truths (i.e., mind-and-language independent truths). Instead of first order mathematical truths, Clarke-Doane claims that "the relevant truths will typically be (first-order) logical truths regarding objects in our environments" (2012: 332). So long as our ancestor's beliefs "lined up" with the physical world, and thereby promoted fitness, it is irrelevant whether they also "lined up" with any mind-and-language independent mathematical truths. Thus, rather

[206] See Joyce (2006: 182). Note, in a way, what Joyce says about morality is just repeating what Darwin originally said about the development of the moral sense; that if it developed in different animals (or a different environment), then morality would be different.
[207] See Clarke-Doane (2012: 330-31). He notes that this clearly would not apply to more complicated and abstract mathematical concepts like the axiom of mathematical induction. Joyce's argument (and those similar to it like Street) are focused on simple or elementary mathematical beliefs just as they do with the moral beliefs.

than requiring a corresponding first-order mathematical truth to explain why a mathematical belief was advantageous, it merely requires that a corresponding "hypothesis about the structure of our environment" turns out to be correct (Clarke-Doane 2012: 332).[208]

Clarke-Doane's argument presents an alternative explanation of mathematical beliefs that clearly confronts Joyce's challenge to find a truth-independent evolutionary explanation of mathematical beliefs. Counter to Joyce's claim, false mathematical beliefs could be useful, so long as those beliefs corresponded with the environment in a way that promoted fitness. If that is the case, then a more global epistemological skepticism may be warranted.

In addition to mathematical beliefs, Joyce argues that the same truth requirements—that the evolutionary explanation requires background assumptions about truth in order to explain how the belief was selected for—applies to scientific beliefs. Just as with the case of mathematics, Joyce argues that any evolved mechanism for "scientific inquiry" can only be explained if they have some connection with the truth. [209] Joyce concludes that:

> …the evolutionary genealogy of morals contrast with other cases (such as arithmetical and scientific beliefs)—in that the former does not presuppose the truth of the beliefs—is an important observation. (2006: 184)

Street also provides a very similar response. According to Street, the Darwinian Dilemma does not affect realism about non-evaluative facts "such as facts about fires, predators, cliffs and so on" (2006: 130).[210] Rather than focusing on mathematical or scientific examples, she considers basic

---

[208] He says, "the picture cannot, then, be that for any mathematical hypothesis that we were evolutionary selected to believe H, there is a (first-order) logical truth corresponding to H that captures the intuitive reason that belief in H was evolutionarily advantageous" (Clarke-Doane 2012: 330). Also note that the purpose of his argument is that the moral antirealist has to also be a mathematical antirealist, not that moral antirealism is wrong.

[209] Here, he draws on the argument of Philip Kitcher (2006), who argues that scientific judgments are different from moral judgments. Others that have made similar arguments include: Crisp (2006), Gibbard (2003), Pinker (2002), Sinnott-Armstrong (2006), and Sosa (2002).

[210] Also see Gibbard (1990) and (2003).

perceptual beliefs. She argues that the "tracking account" of truth is simply the best explanation for our capacity to make non-evaluative judgments. As such, *all* non-evaluative beliefs are most advantageous when they are correct—that is when those beliefs track the truth and the beliefs correspond with the truth.[211]

The same types of issues, however, occur here as they did for the case of simple mathematical beliefs like arithmetic. It is possible and intelligible to imagine these truths being very different in just the same way that we can imagine moral truths being very different. So why ought we presuppose the truth in any of these cases? Presupposing the truth as a requirement in these cases—for mathematical beliefs, scientific beliefs, and basic perceptual beliefs—begs the question. There is no obvious reason why the process of evolution (i.e., the mechanism of natural selection) would be sensitive to truth in one case but not in the other (if it can be said to be sensitive to the truth at all). Mathematical, scientific, and perceptual beliefs may be difficult to doubt since they often appear as though they could not be otherwise. [212] Consequently, it is easy to presuppose that those types of beliefs require truth as part of their evolutionary explanations. But just because it would be more convenient for us for these beliefs to be true—or at least accurate—does not make it so; to make such a leap is merely wishful thinking. Debunkers, like Street and Joyce, fail to provide an adequate explanation for why *the evolutionary process differs* between cases of evaluative and non-evaluative beliefs with respect to truth since they merely insist that it could not be otherwise. Additionally, there is a larger unsettled question of what truth is, which has not been adequately addressed.

[211] For related discussion, see Gibbard (1990: Chapter 6) and (2003: 253-58).
[212] As Descartes showed in his *Meditations*, there can be reasons to doubt even the most obvious common sense beliefs!

Even if debunking arguments are successful (which is a separate issue I will take up shortly), there is no obvious reason why one set of beliefs is automatically immune from the skepticism that follows from accepting evolved innate mechanisms that are responsible for various beliefs and judgments. If we take debunking seriously, then global epistemological skepticism is a serious problem that has not been dealt with adequately. Similar to Sidgwick's point that such arguments for moral skepticism destroy their own premises, Peter Railton argues that debunking arguments,

> …assume that humans are psychologically and socially equipped to carry out scientific inquiry, to produce and test hypotheses in ways that yield impartial epistemic justification, despite the fact that our perceptual, cognitive, linguistic, and deliberative capacities have all been shaped by a process of natural selection in which opportunism not impartiality, warrant, or truth rules. Why, then, isn't human epistemic pretense illusory? The hard headed argument hammers itself into the same ground into which it had previously pounded morality. (2000: 57)[213]

In other words, these arguments "debunk" themselves and global epistemological skepticisms appears inevitable. At minimum, the burden of proof lies with the debunker in very much the same way they claim it to lie with the moral realist.

[213] Note that Joyce rejects Railton's claim (Joyce 2006: 183).

Chapter Four: Realist Responses to Evolutionary Debunking Arguments

In the previous chapter, I laid out the debunking arguments of Ruse, Joyce and Street to reveal the basic structure of such arguments, the underlying assumptions (epistemological, metaphysical, and empirical) that are at work, and the types of metaethical conclusions these arguments attempt to establish. I also included several reasons to question these debunking arguments, including the global epistemological skepticism (which includes mathematics and science) that follows, albeit unintentionally, from these arguments. This chapter considers additional arguments launched against these debunkers, particularly by those who defend moral realism against debunking arguments while maintaining that an innate evolved moral sense influences morality. I will argue that while these realist defenses against evolutionary debunking ultimately fail, the debate reveals broader problems associated with using evolutionary explanations to settle the metaethical issues at stake in this debate.

Objectors to the evolutionary debunking arguments of Ruse, Joyce, and Street attempt to provide some version of moral realism that can withstand the debunkers' challenges while continuing to support the claim that there is an evolved innate moral sense.[214] These moral realists, whom I will call Darwinian Moral Realists, do not deny the connection between evolution and morality as a way to avoid the problems raised by the debunkers.[215] Darwinian Moral Realists establish their defenses by examining different assumptions—epistemological, metaphysical, and empirical—that make up the debunking argument and provide additional explanations of these assumptions that are compatible with moral realism. In other words,

[214] These arguments against the debunkers are not questioning the moral nativism that underlies these debunking arguments. For arguments of that sort, see Prinz (2009a) and Nichols (2005).
[215] Despite these similarities, Darwinian Moral Realists do not all share the same version of realism. For example, some, like Wielenberg, take a naturalistic approach while others, like Enoch, take a non-naturalistic approach.

Darwinian Moral Realists attack the premises of the debunking arguments, questioning the truth of the debunking arguments but not the logical structure of the debunking argument itself.[216] This will turn out to be a very significant oversight since the logical structure of the argument should also be questioned, and not simply the details of the premises per se.

The logical structure of debunking arguments require closer scrutiny because these arguments aim to use explanations as justifications—that is, they make the difficult move from explanation to justification that was discussed in Chapter One. Furthermore, these arguments rely on inference to the best explanation. Consequently, debunking arguments and their realist responses, encounter two significant problems. First, debunking arguments fail to provide the "work" required to make the move from explanation to justification.[217] As Matthew Braddock has recently pointed out "…even if we grant the explanatory premise, it is an opaque matter how to get from there to the debunking conclusion" (2012: 3).[218] This opaqueness—about the connection between evolutionary explanations (the premises) and the rejection of moral realism (the conclusion)—is a significant problem Darwinian Debunkers must contend with and yet one that is rarely discussed in the literature. What this opaqueness—what I will call "inferential opaqueness"—reveals are cases in which it is unclear whether or not that "work" exists and, thus, whether or not their conclusions are justified. I will argue that the nature of evolutionary explanations is responsible for this problem. Second, reliance on inference to the best explanation creates a unique problem for this debate due to the normativity entailed in this

[216] Since the debunking arguments do not claim to be deductive, I have chosen to use "truth" and "logical structure" instead of soundness and validity.
[217] Perhaps another way of thinking about this is that they are missing premises, although I'm not sure that is the only thing going on in this case.
[218] Braddock defends debunking arguments, claiming that he can provide additional premises to show how evolutionary explanations can debunk moral realism, therefore removing this problem of explanatory opaqueness. I, however, will argue against this claim.

argument type. Determining which evolutionary explanation best explains a particular trait will often require selecting one explanation from a set of competing explanations, each of which explain the trait equally as well. In such cases, a normative (or more generally evaluative) judgment must be made to pick out the *best* explanation from the set requires some normative criteria. But since normative (or more generally evaluative) beliefs are in question, determining which explanation is best, and thus making an evaluative judgment, presents a distinct problem for each side of the debate.

To begin this discussion, I will organize some of the prominent Darwinian Moral Realist arguments based on the additional realist explanations that drive their objections into two categories: (1) third factor explanations, which aim to establish additional explanations compatible with evolutionary moral realism, and (2) realist normative theories, where particular normative moral theories are introduced that claim to be consistent with evolutionary explanations.[219] Categorizing this way helps reveal the variety of realist responses launched against the debunking arguments. In the following sections, I will show why such realist defenses, despite these variations, are inadequate. However, I will only examine arguments in the first category, since the issues that arise within the first category are applicable to the second. In most cases, this inadequacy stems from failing to address the underlying structure of the debunking arguments—that is, failing to address the problems inherent in connecting evolutionary explanations to moral justifications. I conclude with some considerations of where this leaves the realist/antirealist debate over the debunking of morality and suggest that the focus

[219] These categories are not mutually exclusive; in fact, many of the arguments against debunking use some combination of them to provide support for their realist positions (while maintaining the initial evolutionary assumptions that humans have an evolved innate moral sense). Therefore, I am categorizing by argument and not by author since some authors employ more than one type of argument.

be turned to the problem of inferential opaqueness, since the nature of evolutionary explanations are the root this problem.

A. Darwinian Moral Realist Responses to Debunking

Each attempt to refute the debunkers through realism contains a unique set of realist arguments; however, these arguments can be categorized based on the particular premises of the debunking argument they take aim at. As discussed in the previous chapter, debunking arguments contain several different theoretical assumptions about epistemology, including claims about truth and knowledge that are built into their premises, and as such, evolutionary debunking arguments are typically understood as an epistemological challenge.[220] In addition to these epistemological claims, although not as obvious, the debunking arguments also make more general assumptions about normativity since morality, or more specifically, the moral truths and our ability to gain epistemic access to them (if they exist), are in question. Realist responses to debunking tend to address a number of these epistemological and normative claims, aiming to show that evolutionary explanations are compatible with realist-friendly versions of these assumptions. As such, categorizing anti-debunking arguments in terms of these claims—realist epistemological claims and realist normative claims—helps to clarify the underlying arguments at work. Therefore, the two main categories are (1) third factor explanations, which concern realist epistemological claims and aim to establish additional explanations compatible with evolutionary moral realism, and (2) realist normative theories, which concern realist normative claims where particular normative moral theories are introduced that claim to be consistent with evolutionary explanations.

[220] Note that this is not taking into consideration what metaethical conclusions they reach; see Chapter Three, Section A.i.a. for a discussion.

In the following sections, I will examine some examples of Darwinian Moral Realist anti-debunking arguments in the first category—those that contain third factor explanations. I will show that the issues that arise for third factor explanations not only apply to the second category of realist normative theories, but also to the debunking debate more generally. Close examination of the arguments will reveal some fundamental flaws with both Darwinian Moral Realist anti-debunking accounts and the debunking accounts.[221]

i. The Explanatory Challenge and Third Factor Explanations

As discussed in previous chapters, explanations play a significant role in debunking arguments. Focusing on the issue of explanation in debunking arguments, we discover an "explanatory gap" that the realist needs to deal with (at least, this is what the antirealist debunker contends).[222] Recall that the debunkers argue that the Darwinian Moral Realist needs to provide an explanation consistent with evolutionary explanations, which they then claim is not possible. In other words, debunkers argue that the best evolutionary explanations available do not support moral realism, and therefore, moral realism is an untenable view (i.e., we should embrace moral antirealism or we should be moral skeptics about realism). Framing the issue this way—as an explanatory problem—leads most anti-debunkers to argue that they can provide suitable explanations to "bridge the gap", illustrating that moral realism is consistent with evolutionary

[221] It is worth mentioning that not all of the arguments against debunking share the same interpretation of what, exactly, is being "debunked" by each evolutionary debunking argument. This, of course, affects how the Darwinian Moral Realists develop their responses (as well as the viability of their responses).

[222] I am using the term "explanatory gap" to refer the realist problem of matching up moral truths (that realists posit) with our moral beliefs and not the problem of consciousness in physicalist theories of mind.

explanations. These bridges are called "third factor explanations", a term coined by David Enoch (2010) in his attempt to take on the debunking challenge.[223]

Darwinian Moral Realists who provide third factor explanations take on this explanatory challenge by arguing that there is a correlation for reasons other than a direct connection between evaluative beliefs and evaluative truths—that is, they argue for a "third factor" or additional explanation that attempts to bridge the gap, via explanation, between mind-dependent beliefs and mind-independent truths. These third-factor explanations are meant to resolve the dilemma by presenting an alternative solution as a way out of an "irreconcilable" dilemma. Two Darwinian Realists that offer such "third factor explanations" to refute the debunkers are David Enoch (2010) and Erik J. Wielenberg (2010).[224] In this section, I will examine their third factor explanations and consider how successful such explanations really are in dealing with evolutionary debunking.[225]

In David Enoch's article "The Epistemological Challenge to Metanormative Realism: How Best to Understand it, and How Best to Cope with It", he presents a reading of Street's Darwinian Dilemma in the context of the general epistemological challenge that confronts moral realism.[226] What Enoch points out is that this challenge, and more specifically Street's Darwinian

[223] See also Wielenberg (2010). Note that in this discussion, moral beliefs and judgment, evaluative beliefs and judgments, and normative beliefs and judgments all refer to the same thing, so I will be using these terms interchangeably as they do.
[224] Skarsaune (2011) also offers a third factor explanation, but I will not consider it here in depth since it is subject to the same types of criticisms that I will raise later. However, it is interesting to note that he argues that Street's view of moral realism is a straw man since she is attacking a view a realism—a mind independent realism—that no one (in the relevant literature) really endorses.
[225] The main target here is Street, but some include Ruse and Joyce as well.
[226] Note that Enoch presents a much broader epistemological challenge aimed at the "robust" non-naturalist realist and as such there are certain parts of his argument that are not relevant to the current discussion about evolutionary debunking arguments. Enoch also agrees with Street

Dilemma, centers on explanations, claiming that "the game is an explanatory one and the winner is going to be determined on plausibility grounds" (2010: 427). As suggested earlier, considering evolutionary debunking arguments in terms of explanatory issues is an appropriate way of framing the debate. However, Enoch also attempts to show how realists can deal with Street's Darwinian Dilemma by offering an evolutionary explanation consistent with realism.[227] This response—which is merely an extension of the Darwinian Hypothesis—along with several other aspects of his argument render his response unsuccessful. His discussion, however, sheds light on the broader issues associated with evolutionary debunking arguments (like Street's Darwinian Dilemma), especially when presented as epistemological challenges.

According to Enoch, "realists owe us an account of how it is that we can have epistemic access to the normative truths about which they are realists" (2010: 413). Enoch argues that the best way to understand this challenge, what he calls the "epistemological challenge", is to see it as "the need to explain the supposed correlation between normative truths and our normative beliefs" (2010: 422). So, the epistemological challenge is merely an explanatory challenge—the realist needs to find a suitable (or reasonable) explanation for the correlation to prove that their view is plausible.[228] That is, if there is a correlation between two factors, for example a belief (A) and the truth (B), then that correlation requires an explanation. Enoch points out that many cases, perhaps even most cases, are fairly unproblematic because the correlation between two factors, (A) and (B), can be given a causal or a constitutive explanation. For example, I believe that the

---

that there is nothing essentially Darwinian about the epistemological challenge, that any causal argument will do (2010: 426).

[227] In other words, the realist must provide a "best explanation" here, exactly what Street says the realist cannot do. But, if the realist can, then they would be vindicating moral realism and filling in the explanatory gap is an attempt to justify moral realism via evolutionary explanations. See Enoch (2010: 431).

[228] Note that Enoch also suggests that these are the "best" explanations for the correlation as well.

sun rises in the morning and it happens to be true that the sun does rise in the morning. The correlation between my belief and the truth is nothing miraculous because it can be explained with a causal story: the rising of the sun every morning causes my belief about the rising sun. However, there are cases where a simple causal (or constitutive) explanation is unavailable, and these cases are problematic.[229]

Enoch uses Hartry Field's example of mathematical Platonism to illustrate just such a problematic case. Very briefly, if Platonism is true, then the correlation between my belief in a mathematical proposition and the truth of that proposition cannot be explained with a causal explanation because, according to Platonism, mathematical objects are "abstract and causally inert" (Enoch 2010: 421).[230] The only way out of such a problematic case would be to introduce a "third factor explanation", that is, some other factor, factor (C), in addition to factor (A) and factor (B), that can explain the correlation between factor (A) and factor (B). Factor (C) replaces the need for a causal explanation while taking away the mystery of how the two factors are correlated.

Of course, just as every correlation cannot be given a causal explanation, third factor explanations are sometimes unavailable as well.[231] What Enoch and others who utilize this type of argument do not fully explore is what happens when a third factor explanation is unavailable and the correlation cannot be explained. Most, like Enoch (as well as Field and other antirealists like Street, Harman, and Joyce), assume that acceptance of realism under these conditions leads

[229] Constitutive explanations are causal, but "the explanation shows, instead, that the fact-to-be-explained is constituted by the underlying causal explanations" (Salmon 1984: 270).
[230] Also see Field (1989). If mathematical objects are causally inert, then they cannot play a role in a causal explanation.
[231] And this is what Field concludes about mathematical realism: there is no third factor explanation available to the Platonist, and therefore, mathematical Platonism should be rejected (1989: 25-30). Also see Enoch (2010: 421).

to an untenable skepticism and thus the inability to provide a proper explanation *justifies* acceptance of some form of antirealism (or at least a deep moral skepticism, which is akin to antirealism). Enoch admits that the failure to provide a satisfactory explanation of this sort is not a definitive refutation of realism; however, realism loses "plausibility points" because the *best* metaethical theory should be capable of explaining whatever needs explaining. In other words, Enoch and the debunking arguments he is discussing rely on inference to the best explanation to reach their conclusions. Despite the wide acceptance of this form of reasoning, it requires further scrutiny in these debunking cases due to the inferential opaqueness—that the connection between evolutionary explanations (the premises) and the rejection of moral realism (the conclusion) is hopelessly unclear—and normativity involved in determining the "best" explanation, issues I will come back to shortly.

When dealing with normative beliefs, Enoch argues that the robust realist (the metanormative position that argues for irreducibly non-naturalist normative truths and what Street calls the non-naturalist realist) faces a predicament similar to the mathematical Platonist. Like the Platonist, the robust realist must confront the epistemological challenge without the ability to provide a direct explanation of the correlation between moral beliefs and moral truths. Naturalist versions of metanormative realism are, in principle, better off, according to Enoch, because naturalists argue that normative truths are causally efficacious, reducible to natural facts or identical with natural facts. As a result, they can provide a causal explanation as the third factor explanation. Street, and many other antirealists like Joyce and Ruse however, do not think the naturalists will be able to provide such explanations. Like these debunkers, Enoch does not think such a naturalist view of normativity is forthcoming either. Therefore, the challenge is best

thought of as an issue for the robust realist.[232] If there is a plausible third factor explanation for the robust realist then there is no longer a dilemma.

With these considerations in mind, Enoch continues by arguing that we should not understand the epistemological challenge the robust realist faces as a question about epistemic access, justification, reliability, justification or general skeptical worries per se.[233] These ways of understanding the challenge are too broad for Enoch because they encompass more than just the normative beliefs that are in question for the robust moral realist.[234] In other words, such epistemological issues are not a special problem for normative beliefs, but rather for any type of belief. Here, Enoch acknowledges (in a footnote) that he is taking a position similar to Sayre-McCord, that "there is no distinctive epistemology of moral belief" (2010: 417).[235] Instead, for the robust realist, the epistemological challenge concerns explaining the correlation between a truth *p* and our moral beliefs that "match" this truth.[236]

Enoch recognizes that a number of theoretical assumptions, both epistemological and metaphysical (ontological) are required for setting up this type of challenge to realism. For example, when dealing with the issue of reliability, Enoch argues that the epistemological challenge assumes that moral (normative) beliefs could be reliable and that normative truths exist; otherwise, there is no epistemological issue to debate. Thus, Enoch claims that challenging these assumptions is a different type of argument against realism, another point that requires further discussion. For although questioning the epistemological and metaphysical (ontological) claims that drive the epistemological challenge may be, as Enoch says, a "different" argument,

---

[232] See Enoch footnote 26 (2010: 414-15).
[233] For a full discussion of each, see Enoch (2010: 414-20).
[234] Furthermore, they often rely on many other controversial epistemological positions. See Enoch (2010: 415 and 420).
[235] Also see Sayre-McCord (1995: 138).
[236] See Enoch (2010: 421).

this does not make them immune from criticism. And since these assumptions play an integral part of the challenge, it is possible that if these assumptions are wrong, then the epistemological challenge is not a real challenge after all.

In its narrowest, and strongest, form then the epistemological challenge is simply about the correlation between independent normative truths and normative beliefs, which is exactly what Enoch takes Street's Darwinian Dilemma to represent.[237] As previously discussed, Street's Darwinian Dilemma claims that since evolutionary forces have greatly influenced human evaluative attitudes, albeit indirectly, we must explain the relation between evolutionary forces that shape our evaluative judgments and evaluative facts, between "the selective forces that have influenced our evaluative judgments on the one hand, and the independent evaluative truths that realism posits, on the other" (2006: 109). Enoch points out that the problem for the realists is that they "are committed to just such a correlation between the independent evaluative truths and the evaluative judgments that you can expect evolutionary successful creatures to make, the normative judgments that were (roughly speaking) selected for" (2010: 425-26). In trying to explain this correlation, realists are faced with a dilemma: if the realist cannot provide a reasonable explanation for the correlation, then the correlation is an unbelievable miracle; however, the only explanation of the correlation would be a "tracking" account, which Street finds scientifically untenable. Recall, the "tracking" account maintains that evaluative judgments were selected for (increased fitness) because they were true. However, Street argues that other evolutionary explanations, such as the "adaptive link" account, which do not include reference to the truth, provide a better explanation of the evolution of evaluative judgments, thereby

[237] Enoch admits that Street's Darwinian Dilemma is the closest argument to the epistemological challenge he describes, although it's not exactly the same. In a footnote Enoch refers to Gibbard (2003) as a precursor and Joyce (2006) as a related argument (2010: 425-26).

rendering the "tracking" account unscientific. Since neither of these options is an acceptable position for the realist to take, Street argues that the realist is faced with an irreconcilable dilemma.[238]

In Enoch's terms, what causes the dilemma is a missing explanation of the correlation between evaluative judgments and evaluative facts. And, under Street's conditions of the argument, a causal explanation is untenable (the tracking account would be the causal explanation of the correlation, and this, she argues, is unscientific and hence not available). So, what is left for the realist to do? This is where Enoch reaches for that third factor explanation to bridge the gap—a third factor explanation, which he argues is a "godless pre-established harmony" that provides an explanation of the correlation between normative truths and normative beliefs and avoids both horns of the dilemma. His argument runs as follows:

> Selective forces have shaped our normative judgments and beliefs, with the "aim" of survival or reproductive success (so to speak). But given that these are by in large good aims—aims that normative truths recommend—our normative beliefs have developed to be at least somewhat in line with normative truths. (2010: 430)[239]

He continues, arguing that:

> The connection between evolutionary forces and values—the fact that survival is good—is what explains the correlation between the response-independent normative truths and our selected for normative beliefs. The fact that (roughly speaking) survival is good pre-establishes the harmony between the normative truths and our normative beliefs. (2010: 431)

So, the correlation between normative beliefs (evaluative judgments) and normative truth (evaluative truths) is explained by a third factor, the pre-established harmony provided by the claim that evolutionary aims (survival and reproductive success) are good. As such, Enoch thinks the robust realist can avoid the two horns of the dilemma by evoking this third factor explanation.

---

[238] See Street (2006: 125-28) and Enoch (2010: 426).

[239] Note that Enoch claims (in a footnote) that he is using the terms "aim" etc. loosely and that he recognizes the evolution has not "aims" or "intentions". I will take this up shortly.

Unlike the mathematical Platonist who does not have any analogous explanation—and thus ought to reject Platonism—the robust realist can explain the correlation and need not reject normative realism. In other words, if there is a plausible third factor explanation then there is no dilemma.

Erik J. Wielenberg agrees with Enoch that Street's Darwinian Dilemma is best understood in terms of an epistemological challenge that asks the moral realist to provide an explanation of the correlation between moral beliefs and moral truths.[240] As such, he too argues that all that is required is a third factor explanation. But, while Enoch offers a third factor explanation about the "the goodness of survival or reproductive success," Wielenberg offers a third factor explanation based on "certain cognitive faculties" that he uses to establish a particular model of moral knowledge (2010: 450). Unlike Enoch, Wielenberg's third factor explanation is also a causal explanation, connecting evolved cognitive faculties with beliefs about moral rights.

Wielenberg starts by arguing for the existence of what he calls "moral barriers," or a basic set of moral rights, that were fitness-enhancing. These moral barriers stem from "viewing ourselves as possessing boundaries that may not be transgressed no matter what" (Wielenberg 2010: 445). And, belief in such "moral barriers", such as a right not to be killed, were favored by evolutionary forces. As a result, humans evolved certain cognitive faculties that entail the existence of these moral barriers, which, in most western cultures, are known as moral rights.[241]

---

[240] Again, these Darwinian Moral Realists assume that there is only an epistemological challenge here, but as I argued in the previous chapter, I think the conclusion is not merely a moral skepticism, but rather a full blown antirealism. This will turn out to be another significant weakness of these responses.

[241] According to Wielenberg, moral barriers include the concepts of rights and duties, depending on cultural background. For details of his evolutionary explanation of moral barriers, see Wielenberg (2010: 444-47).

Like Enoch, Wielenberg recognizes that a number of theoretical assumptions, both epistemological and metaphysical (ontological), are required for setting up this type of response. The existence, or at least the possibility of the existence of such moral rights, is assumed based on the structure of the debunking arguments themselves. Wielenberg argues that, "such arguments are not aimed at showing that there are not moral truths. Rather, such arguments are aimed at showing even if there are moral truths, human beings lack knowledge of such truths" (2010: 447).[242] He argues, therefore, that moral barriers do exist, and our lack of epistemological access is not sufficient for showing that they do not exist. And, like Enoch, he does not think that taking this liberty begs the question.

Next, Wielenberg presents a view of moral knowledge that combines elements of reliabilism, foundationalism, and intuitionism. In the case of moral beliefs, he argues that some moral beliefs are justified basic beliefs (or foundational beliefs) formed by certain evolved cognitive faculties. These cognitive faculties also possess a reliable belief-forming process and, thus, produce generally reliable beliefs, which includes moral beliefs.[243] The truth of these basic moral beliefs is self-evident and unanalyzable; in other words, humans have intuitions about certain moral claims that justify their truth and they are brute facts for which no further explanation is available.[244] Adopting an ethical intuitionism similar to Michael Huemer's, Wielenberg claims that, "we have an intuition (an intellectual appearance) that the moral claim is true and come to believe the moral claim on the basis of this intuition" (2010: 448).[245]

---

[242] Note that this is something I challenged in the previous chapter.
[243] Here, Wielenberg uses Peter Carruthers argument that evolutionary processes tend to produce true beliefs. See Wielenberg (2010: 450) and Carruthers (1992: 112-13).
[244] See Wielenberg (2010: 460).
[245] Here, Wielenberg cites Huemer (2005: 99-104). Wielenberg's view also clearly resembles the British Moralists/sentimentalist tradition, removing god as the guarantee of truth and replacing it with an evolutionary story.

With this view laid out, he then argues that any creature that evolved to possess these cognitive faculties will also possess moral barriers (certain basic rights) because the cognitive faculties entail the existence of moral barriers. As a result, these evolved creatures are likely to believe in those moral barriers as well. Thus, there is a causal connection between these evolved cognitive faculties and moral beliefs, which explains why moral beliefs are correlated with moral rights (moral barriers).[246]

Applied to Street's Darwinian Dilemma, Wielenberg argues that this third factor explanation, like Enoch, avoids both horns of the dilemma. Wielenberg understands Street's Darwinian Dilemma to entail the following thesis (what he calls the "Evolutionary Debunking Thesis"):

> If S's moral belief that P can be explained without appealing to the truth of P, then S's moral belief that P being true would be a lucky coincidence. (2010: 458)

According to Wielenberg, his third factor explanation can explain why "S's moral belief that P being true" would not be just a lucky coincidence. Instead, his model of moral knowledge would explain this correlation between S's moral belief and the truth of the moral belief in the following way:

> [T]he presence of the very cognitive faculties that cause (or at least causally contribute to) my belief that I have certain rights also entails that I have those very rights. Furthermore, the process that produces this belief is a reliable one, and hence the belief constitutes knowledge. (2010: 458)

Thus, S's moral belief that P is true is the result of a reliable belief forming process that stems from cognitive faculties that produce this initial thought about moral rights; in other words, we should trust our cognitive faculties since our minds do not generally play tricks on us. Presented

[246] According to Wielenberg, "when evolutionary processes produce beings that think they have rights, those beings generally do have rights, and when evolutionary processes produce other kinds of beliefs, those beliefs tend to be true" (2010: 451).

this way, Wielenberg claims that his view is akin to Steven Pinker's view that "our moral sense has evolved to mesh with the intrinsic logic of ethics" and that there is nothing mysterious or coincidental about it (2006: 151).[247]

Wielenberg's response to the Darwinian Dilemma is reminiscent of Descartes's foundationalist response to the epistemological skepticism presented in the *Meditations.* The moral skepticism generated from debunking arguments is similar to a Cartesian skepticism, where evolution plays the role of the evil demon, tricking us into believing certain moral claims regardless of their truth or existence. Just as Descartes resolves this skepticism by introducing a benevolent God that provides us with trustworthy cognitive faculties, Wielenberg argues that evolution produces cognitive faculties that are reliable, so we cannot doubt the basic, foundational moral beliefs they produce. But whereas it may be easy (or easier) to see how a benevolent God would grant us reliable cognitive faculties (if such a God exists), it is much less clear why evolution produces reliable cognitive faculties. Here, Wielenberg offers an argument presented by Peter Carruthers, in which Carruthers says "organisms (of the sort that act on beliefs) will only survive, in general and in the long run, if they base their actions on beliefs that are true, or at least close to truth" (1992: 112). Additionally, Carruthers assumes that other mechanisms closely related to forming beliefs, like perception and memory, are generally reliable. Wielenberg thinks this is enough to support the claim that all evolved cognitive faculties are reliable.

In response to Joyce's debunking argument, Wielenberg defends his view that evolved cognitive faculties are reliable.[248] In Joyce's discussion of evolutionary approaches that vindicate

[247] Also see Wielenberg (2010: 459).
[248] I will put aside his discussion of Ruse because of his interpretation of Ruse is contentious. See Braddock (2012: 17).

(rather than debunk) morality, Joyce considers the possibility that there are cases in which false beliefs turn out to be fitness-enhancing, and thus some of our cognitive faculties could be unreliable. Wielenberg understands Joyce's argument against vindicating morality to entail the following thesis:

> If S's moral belief that P can be explained without appealing to the truth of P, then S's moral belief that P is a product of an unreliable process. (2010: 462)

This thesis, according to Wielenberg, is only supported by an analogy—one that Wielenberg thinks is too weak to establish the claim that moral beliefs are produced by an unreliable process. Without recounting the details of Joyce's analogy (which was discussed in Chapter Three, Section A.i.b.), Wielenberg merely points out that unlike Joyce's case of magical plant beliefs that are false and that were produced by an unreliable process, we have no reason to assume that moral beliefs are false, and thus it does not follow that they were produced by an unreliable process.[249]

B. Problems with Third Factor Explanations

The most obvious problem with third factor explanations is the contentious nature of evolutionary explanations embedded in them. As most philosophers acknowledge, these explanations are speculative or hypothetical. Consequently, most arguments that rely on evolutionary explanations are, therefore, couched in conditional form (e.g. *if* this evolutionary explanation is correct, then this is what we can say about morality).[250] So although the specific

[249] Joyce responds to Wielenberg by arguing that we have no good reason to believe that moral beliefs are *not* false—we cannot assume, without other good reasons, that the "pill-pusher" in the case was benign rather than malign or mischievous (2016b: 12-13). I will return to this argument in more detail in Chapter Five.

[250] For example, Enoch says this about his own third factor explanation: "[I] have no idea whether this explanation actually works (or whether the phenomenon it is supposed to explain is actually a real phenomenon). All that is crucial for me is that it could work, and that its structure is exactly similar to that of the explanation I am really after" (2010: 431).

evolutionary explanations offered as part of these third factor explanations are based off of work in evolutionary psychology and sociobiology, they still consist of speculative scientific explanations that should not be taken for granted.[251] The conditional or hypothetical nature of these explanations raise concerns about the reliability of the claims derived from them. If the evolutionary explanations used to support a particular view turn out to be false, then the claims derived from them would also fall apart. Thus, there is a worry that all such conditional or hypothetical arguments are built on a house of cards and this seems a poor strategy for resolving any philosophical debate.[252]

An additional concern with such explanations is the issue of bias. Since there are often many competing evolutionary explanations for any given trait—and especially for a trait as abstract and contentious as "morality" or a "moral sense"—some views may be guilty of cherry-picking; that is, they may be guilty of endorsing evolutionary explanations that fit neatly with a particular philosophical view rather than which explanation is actually best. In addition (and closely related) to cherry-picking, some may also be guilty of committing the Whig fallacy. According to historian Herbert Butterfield who coined the term, the Whig fallacy occurs when, "studying the past for the sake of the present" results in a distortion of the past because one attempts to see the past in terms of the present (1931/1965: 16). This type of historical interpretation leads to an artificial selection of which past events are considered relevant or significant and, thus, an erroneous "abridgement of history" occurs (1931/1965: 24). It also

[251] For example, Wielenberg relies on the work of Hamilton (1963), Sober and D.S. Wilson (1998), and Buss (2004).

[252] This problem does not necessarily apply to all conditional or hypothetical arguments. Such argument types may be useful in other debates (e.g., about possible worlds or modal logic). I am only arguing that in relation to evolutionary explanations and the way these arguments intend to use such conditional/hypothetical arguments is problematic.

falsely presents history as an inevitable course of progress to the present state.[253] In the case of evolution of morality, it is possible that interpreting evolutionary history in terms of our current view of morality or a moral sense leads to a similar mistake—and thus a misunderstanding—about the evolutionary past of humans and the development of morality.

As a result, one must proceed with extreme caution when entering such a debate. Even if we could resolve or avoid these problems, which some might say is possible (at least in principle), a more troubling issue emerges. The nature of evolutionary explanations turns out to be not only inherently speculative or hypothetical, but also historical and irregular in such a way that renders them unsuitable for resolving this metaethical debate.

Although the issues above are for both sides of the debate (for both the debunkers and those who argue against the debunkers), I will start by focusing on the particular issues facing the realist responses identified in the previous section. Both Enoch and Wielenberg run into additional problems concerning the nature of evolutionary explanations when they extend these explanations to make specific normative claims.[254] The first concern that emerges for both Enoch and Wielenberg is the viability of each of their proposed third factor explanations.

In Enoch's case, he offers a third factor explanation that consists of an evolutionary explanation that includes a "godless pre-established harmony" between normative truths and normative beliefs based on the "fact that survival is good". His evolutionary explanation not only conflates the "good" with the "natural" (a clear violation of the naturalistic fallacy, although one that those immersed in the debate dismiss), but it also rests on a folk notion of evolutionary

[253] See Butterfield (1931/1965: 2, 55-58).
[254] So many evolutionary explanations are contentious. As Gould (1978) once described it, most evolutionary explanations are akin to Rudyard Kipling's "Just so Stories". There are much deeper questions about the entire field of evolutionary psychology (and sociobiology) that could be raised, but I will not discuss those issues here.

theory—that is, on the way that ordinary people talk about evolution. Folk evolutionary theory assumes that evolution and natural selection have a *telos*—an aim or a goal in mind. This is the common misunderstanding of evolution, which is reminiscent of the teleological views of those like Spencer that are now deemed unscientific.[255] Enoch's third factor explanation requires discussion of evolution in terms of having "aims" and "goals" inherent in it in order to work in the manner he intends. In a footnote, Enoch acknowledges that referring to evolutionary aims is merely a manner of speaking because, taken as a scientific theory, evolution and natural selection have no aims:

> Of course, evolution has neither a mind, nor an aim in mind. Talk of the evolutionary ''aim'' in the text is meant as shorthand for the usual respectable, non-teleological, evolutionary way of putting things. (2010: 430)

Although Enoch claims to understand that evolution is a non-teleological process, without these folk notions his third factor explanation falls flat. Enoch summarizes his third factor explanation in the following way:

> Selective forces have causally shaped our normative beliefs; that survival is good (non-causally but closely) related to many normative truths; and so that survival (or whatever the evolutionary "aim" is) is good *explains* the correlation between our normative beliefs and the normative truths. (2010: 431)

Thus, in order to explain how beliefs match with the truth, there must be some particular aim or goal for which evolution (natural selection) consistently strives for to account for the way selective forces shaped our beliefs in a particular direction—that is, towards a direction of "fit" between our beliefs and the truth. If we acknowledge the lack of intentionality and the

[255] Perhaps Enoch would like to argue for a "folk evolution as theory" similar to the way that David Lewis (1972) argues for a "folk psychology" as theory to explain mental states. I don't think, however, that this would be a very promising route to take since it will be difficult to use platitudes about evolution that ordinary people are inclined to support to develop a cohesive theory about evolution in the same manner that platitudes about the mind ordinary people are inclined to support might be used to establish a theory about mental states (psychology). In others words, platitudes about how the mind works seem very different from those that people might have about evolution.

randomness of evolution (natural selection), then we are back to the unexplained "miracle" that Street (and other debunkers) are worried about for the realist. Without "aims" or "goals" Enoch's argument for this explanation makes little sense—both the "pre-established harmony" and "the fact that survival is good" require evolution to be a teleological process if they are really to *explain* how normative beliefs match up with the normative truths. When you strip his argument of the folk notions there is nothing left to account for the connection.

Enoch admits there is somewhat of a miracle at work here—but he thinks is too small to worry about:

> We are, after all, epistemologically lucky to have evolved in an environment in which having by-and-large true beliefs is presumably conducive to survival and reproductive success…the pressure to supply such an explanation is not particularly strong, nor does it pose a particular problem for the realist. (2010: 434)

But if Enoch's third factor explanation is meant as a response to Street's Darwinian Dilemma (or other debunking arguments), it is difficult to see how the pressure to supply an explanation is not strong—that is exactly what the debunkers are asking for. And, without the evolutionary folk notions as part of the explanation, the miracle left is much bigger than he thinks. Consequently, Enoch's third factor explanation does not successfully bridge the explanatory gap presented by the debunkers.

Wielenberg's third factor explanation does not fare any better. His claim that humans evolved certain cognitive faculties that entail the existence of moral barriers (or moral rights) and produce reliable moral beliefs is equally contentious. Of course, it is not clear that humans possess the moral barriers he describes or that such barriers, if they exist, are the result of an evolutionary process. The mere existence of a group of people possessing a conception of moral rights (or any trait for that matter) does not prove that such conceptions (or traits) have an evolutionary origin. But these issues stem from the speculative nature of evolutionary

explanations and questions regarding the scientific validity of such claims mentioned earlier so I will set them aside for now. A more significant problem this explanation faces comes from claims regarding the reliability of the cognitive faculties that produce moral beliefs.

Although Wielenberg's criticism that Joyce's argument against the reliability of evolved cognitive faculties relies on a weak analogy is legitimate, such criticism does not guarantee that *all* evolved cognitive faculties, are in fact reliable.[256] Evolutionary reliabilism—which William Ramsey defends as "the view that natural selection tends to favor reliable belief-producing mechanisms" because truth is not only a causally relevant feature of beliefs, but is causally relevant in a way that enhances fitness—is a debatable position (2002: 16,18).[257] Wielenberg endorses evolutionary reliabilism without providing any significant defense of it. Thus, the burden of proof lies with Wielenberg to prove that such faculties are, in fact, reliable and not merely assume they are simply because we have no obvious reason to believe otherwise. And if the reliability of the cognitive faculties are not fully addressed, then his argument runs the risk of circularity.

To defend this claim then, Wielenberg presents Carruthers's arguments for evolutionary reliabilism. He does this, however, without any serious consideration of the arguments that could be lodged against it.[258] And while a connection between access to the truth and fitness would support the claim that human cognitive faculties are reliable, this solution merely shifts the explanatory burden from the connection between normative truth and normative beliefs to the connection between normative truth and evolutionary fitness. In other words, it requires an additional third factor explanation to explain the connection between truth and fitness. If the

[256] Joyce (2016b) has recently defended his analogy, which I will discuss below.
[257] Note that Ramsey's argument is a response to Plantinga (2002).
[258] For example, see Plantinga (2002).

explanation for the connection between truth and fitness is not satisfactory, the explanation between truth and belief falls apart as well.

Even if a satisfactory account of the connection between truth and fitness is available, the argument still runs the risk of circularity. As previously mentioned, Wielenberg's response to the Darwinian Dilemma is reminiscent of Descartes's foundationalist response to the epistemological skepticism presented in the *Meditations.* Just as many argue that Descartes's response to epistemological skepticism is viciously circular, Wielenberg's response to moral skepticism runs the same risk of circularity. Remember, how the Cartesian circle is typically understood.[259] Descartes' argument runs something like this:

> (1) If God is benevolent (non-deceiving), then when I perceive *x* clearly and distinctly, then *x* is true.
>
> (2) I clearly and distinctly perceive a benevolent God exists.
>
> (3) Therefore, it is true that a benevolent God exists.
>
> (4) If a benevolent God exists, then when I perceive *x* clearly and distinctly, then *x* is true.
>
> (5) Therefore, if I perceive *x* clearly and distinctly, then *x* is true.

Now, Wielenberg's argument runs something like this:

> (1) If we've evolved to have reliable cognitive faculties (non-deceiving), then when I perceive moral beliefs (i.e., that we have moral barriers or rights*)* that were fitness enhancing, then moral beliefs are true.
>
> (2) The cognitive faculties that produce moral beliefs are reliable because moral beliefs were fitness enhancing.
>
> (3) Therefore, the cognitive faculties that produce moral beliefs are reliable.
>
> (4) If the cognitive faculties that produce moral beliefs are reliable, then when I perceive moral beliefs that were fitness enhancing, then moral beliefs are *true.*

[259] See Descartes*, Meditations*, Meditation 3. As mentioned in Chapter One, I'm putting aside any debate over whether or not this is the correct interpretation of Descartes.

(5) Therefore, if I perceive moral beliefs that were fitness enhancing, then moral beliefs are true.[260]

Instead of the Cartesian criteria of knowledge (clear and distinct perceptions) appearing in both the premises and the conclusion to prove that reliable perceptions produce knowledge (true belief), evolutionary fitness appears in both the premises and conclusion of Wielenberg's argument to prove that our moral beliefs were formed by a reliable process that produces knowledge (true belief). As a result, his third factor explanation does not bridge the explanatory gap as successfully as he claims.

Wielenberg recognizes that his argument appears question-begging; however, he thinks this appearance is the result of assumptions he makes about the truth of certain moral claims that he is permitted to take. As a result, he denies circularity due to the context of the argument. Wielenberg says "it is important to keep in mind the thesis I seek to refute: the thesis that even if there are moral facts, humans lack moral knowledge" (2010: 442). This response to the circularity of his argument is insufficient. First, it does not deal with the fact that he establishes the reliability of the evolved cognitive faculties by assuming evolutionary fitness and not merely the truth of certain moral claims. Second, even if he is correct that the thesis he seeks to refute permits an assumption of some moral facts, it does not guarantee that those moral facts are fitness enhancing. To assume that the truth of moral facts are fitness enhancing begs the question.

Setting aside the individual problems that each third factor explanation faces, there are several broader issues these types of arguments face, which I alluded to earlier. These broader problems are applicable to the second category of realist responses, realist normative theories,

[260] Some might argue that I've only created a straw-man out of Wielenberg's argument; however, I have merely presented a simplified version of his argument. The fact that the truth of moral beliefs is guaranteed by the reliability of evolved cognitive faculties that is proven by the truth of the moral beliefs themselves remains. Thus, this is simply a case of the circular reasoning that Plantinga (2002) was worried about.

where particular normative moral theories are introduced that claim to be consistent with evolutionary explanations. The first problem rests with the theoretical assumptions—epistemological, ontological, moral, and scientific—that are embedded in these third factor explanations and the realist normative theories. The second problem stems from the logical structure of the argument—that is, that debunking arguments and third factor responses both rely on inference to the best explanation. The third and final problem is inferential opaqueness—the opaqueness of the connection between evolutionary explanations (the premises) and the rejection (or acceptance) of moral realism (the conclusion) that is at stake in the debunking debate. I will now look at each one of these issues in detail.

i. Theoretical Assumptions

To begin, third factor explanations require a significant number of theoretical assumptions to make them plausible. The same is true of realist normative theories, although I will not account for the details of the argument because the same problems apply.[261] As previously mentioned, both Enoch and Wielenberg acknowledge that they are making these assumptions and disregard any potential criticism about such assumptions. However, these assumptions are central to the debate, and thus criticism of them cannot be disregarded so easily. In particular, the assumption that moral truths (or, in Wielenberg's case moral barriers) actually exist not only plays an essential role in their third factor explanations, but also lies at the heart of the realist/antirealist debate that evolutionary debunking arguments address.

Enoch and Wielenberg both argue that the epistemological challenge presented by evolutionary debunking arguments assumes that moral (normative) beliefs could be reliable and that normative truths exist; otherwise, there is no epistemological issue to debate. However, this

[261] For example, see Copp (2008).

response is unsatisfactory for two reasons. First, both the reliability of moral beliefs and the existence of moral truths are at stake in the debunking debate. The epistemological challenge presented by debunking arguments intends to question the existence of moral truths, either to prove they do not exist (as in the case of Ruse and Street) or at least to decrease the likelihood that they exist (as in the case of Joyce).[262] And whether or not moral beliefs are reliable plays an essential role in the debunking arguments as well—if moral beliefs are not reliable then the truth of such beliefs are in question. As such, it begs the question to make ontological and epistemological assumptions about moral truths. Joyce has recently responded to this issue, addressing Wielenberg's assumptions about moral facts. Although he does not accuse him of begging the question (as I have), he raises concerns about Wielenberg's assumptions regarding the nature of moral facts. Joyce complains:

> Wielenberg thinks he can help himself to a certain view of moral facts without pausing to make the view plausible, since he reads me as claiming that even if moral facts existed, *and whatever their nature*, the EDA [evolutionary debunking argument] would reveal our beliefs about them to be unjustified. (2016a: 12)

Joyce continues, arguing that the nature of moral facts is exactly what's at stake in the debate, and that he intends his debunking argument to place the burden of proof on those like Wielenberg (and thus I assume Enoch as well) who want to argue for those moral facts. He wants them to present arguments for the nature of moral facts to engage in the debunking debate (although he quickly assumes that he would defeat these views without issue). Without such arguments, Joyce points out that Wielenberg's views on the nature of moral facts are unjustified. This leads to the second problem—that as a defense of moral realism their argument is half-baked. The best Wielenberg and Enoch can argue is *if* these moral truths exist then we could

[262] As discussed in Chapter Three, all of the debunking arguments employ Harman-style arguments to undermine beliefs about moral truths (and/or properties) based on the explanatory inertness of moral truths (properties).

have access to them, but there is no reason to believe that such moral truths do, in fact, exist. Independent arguments for why such moral truths exist are still necessary.

ii. Inference to the Best Explanation

Next, these arguments rely on inferences to the best explanation to reach their conclusions. Despite the wide acceptance of this form of reasoning, it requires further scrutiny in these cases of debunking due to both the normativity involved in determining the "best" explanation and the inferential opaqueness mentioned earlier. I will start with the first problem for inference to the best explanation—normativity.

Evolutionary debunking arguments, and the Darwinian Realist's third factor explanations that respond to them, all rely on inference to the best explanation. If the nature of the debate is explanatory, then one way to understand how this debate functions is as a competition between different applications of evolutionary theory (of course, those come with a lot of theoretical baggage, but I will address that problem again shortly). As Enoch pointed out, providing a satisfactory explanation of the connection between moral beliefs and moral truths is cashed out in terms of "plausibility points" because the *best* metaethical theory should be capable of explaining whatever needs explaining.[263] When looking closely at the debate, we find competing sets of evolutionary explanations—some that claim to explain the connection between moral beliefs and moral truth and others that explain the existence of moral beliefs while making no appeal to moral truths. While debunkers like Street and Joyce reject the possibility that any scientifically sound evolutionary explanations would appeal to the truth, this is exactly what the Darwinian Moral Realist attempts to provide. To determine which evolutionary explanation is correct, we must rely on inference to the best explanation: whichever evolutionary explanation

[263] See Enoch (2010: 427).

best explains the phenomena in question is the correct theory. And when there exist two (or more) explanations that explain the phenomena in question equally as well, then appeals to other qualities such as simplicity, parsimony, or other aesthetic qualities come into play.[264]

"Best", however, is a normative (evaluative) term. Thus determining which explanation is best requires a normative (evaluative) judgment. This causes two distinct problems in the debunking debate: one for the debunkers and one for the Darwinian Moral Realists who rely on third factor explanations.

First, the Darwinian Moral Realist who employs a third factor explanation ends up presenting a circular argument; they rely on normativity to explain normativity. In an attempt to explain moral beliefs, or more broadly normative (evaluative) beliefs, and how such normative beliefs are connected to the truth, third factor responses invoke evolutionary explanations that must be selected on normative grounds. But it is precisely those normative grounds that are in question. Granted, normative beliefs about human behavior (morality) and scientific claims (or theories) may seem so radically different that one may argue that they cannot be compared. But why would the content (moral or scientific) of a normative judgment make any difference? And, if the argument is that we have evolved the capacity (a moral sense or some set of cognitive faculties) that is capable of making normative (evaluative) judgments, what reason is there to believe that we evolved to differentiate between types of content that may be at stake when making normative judgments? Within the debunking debate, many, on both sides, have explicitly argued that it is the capacity to make normative judgments (or more broadly speaking, evaluative judgments), which were selected for and constitute an innate moral sense rather than any particular, content specific moral claims.

[264] Here I have in mind something like Quine's (1955), (1966), and (1981) response to this type of problem.

This leads to the problem that debunkers face. The form of the debunking argument, inference to the best explanation, undermines their own argument. If they are correct, and we have evolved to make evaluative judgments without any reference to the truth (so we have evaluative beliefs but there is no connection with the truth), then the evaluative judgment required to decide between competing evolutionary explanations would also not be reliable. Just as we could not know if our moral beliefs were correct since there is no reason to believe that we have moral beliefs connected to the truth, we also could not know if our evaluative beliefs that were used to decide between competing evolutionary explanations were reliable. This is especially problematic in cases where two explanations explain the phenomena equally as well and other aesthetic qualities such as simplicity or parsimony are required.

Most debunkers, particularly Street and Joyce, argue that scientific beliefs are different from moral beliefs. So while moral beliefs may have evolved in such a way to become unreliable, scientific beliefs have evolved in such a way to become reliable. Since I addressed these arguments and the problems associated with them in Chapter Three, I will not recount them here. But let's grant them the principle of charity and assume they are correct. Even if scientific beliefs were different from moral beliefs, the particular problem raised by inferences to the best explanation still remains whenever we need to decide between two competing evolutionary explanations that explain the phenomena equally well. At this point, we would have to make an evaluative judgment that could not be made on purely "scientific" grounds, but rather on those aesthetic qualities mentioned earlier. Beliefs about such aesthetic qualities will obviously be more akin to moral beliefs, and thus the problem still remains.

Street recognizes this problem in her recent work. She argues that global *evaluative* skepticism would follow from her (and any similar) debunking argument (2015: 691). But,

according to Street, such global evaluative skepticism is only a problem for *mind-independent* conceptions of value (i.e., the realist conceptions of values that are under attack in the debunking arguments). Similar to her earlier debunking arguments, she concludes that such global evaluative skepticism is unacceptable, and thus we have good reason to reject mind-independent conceptions of value. *Mind-dependent* conceptions of value (i.e., antirealism), however, do not run into the same problems because the causal origins (in this case the evolutionary origins) of values are irrelevant when values are mind-dependent. In other words, values are created or constructed so their evolutionary origins cannot undermine them. For Street, the ability of the mind-dependent conception of value to avoid global evaluative skepticism vindicates this conception of value—it provides a good reason to believe the mind-dependent conception while at the same time leaves us with no good reason to believe the mind-independent conception (Street 2015: 292-93).[265] As with her previous argument, Street's argument supporting antirealism is insufficient. Street fails to fully explore whether or not mind-dependent values are plausible or capable of providing an adequate conception of value. The ability of a mind-dependent conception of value to avoid global evaluative skepticism alone does not establish or guarantee the truth of a mind-independent conception of value. As a result, her argument for a mind-dependent conception of value (antirealism) is incomplete.

iii. Inferential Opaqueness

The third and final problem is inferential opaqueness—the opaqueness of the connection between evolutionary explanations (the premises) and the rejection (or acceptance) of moral realism (the conclusion) that is at stake in the debunking debate. In other words, it is not obvious

[265] Street (2015) provides a detailed discussion about undermining and vindicating genealogies. She also accepts the evolutionary explanations used to establish her position as a "fixed point" from which she can argue (2015: 692). I will address both of these points in Chapter Five.

how we can derive metaethical conclusions from evolutionary explanations. Closer examination will reveal that the nature of evolutionary explanations render them incapable of settling any metaethical issues.[266] In addition, we find that there are many theoretical assumptions getting attached to the evolutionary explanations that determine the conclusion, and not necessarily with good reason. It is these theoretical assumptions, and not the evolutionary explanations, that are responsible for reaching conclusions about the proposed metaethical positions.

To understand the problem of inferential opaqueness, it is necessary to look more closely at the nature of evolution explanations. Evolutionary explanations attempt to explain the existence of some particular trait or set of traits via evolutionary theory. Evolutionary theory, however, is both irregular and historical and these features create a certain set of limitations on the explanations it generates. These limitations not only account for inferential opaqueness, but also render evolutionary explanations incapable of resolving metaethical issues.[267]

The irregularity of evolutionary theory stems from the nature of the mechanisms that drive the process of evolution. Traits that evolved because they made an organism better suited for its environment, and therefore increased fitness (the ability to survive and reproduce), are considered adaptations. Adaptations are produced by natural selection, which includes positive, negative, directional, stabilizing, disruptive, and kin selection. Each of these types of selection is related to fitness, although they are related in different ways. Evolution also includes processes that are not related to fitness, such as genetic drift (sometimes referred to as neutral selection) and spandrels (byproducts). As a result, fitness does not always determine the existence of

[266] I take metaethical issues to include both epistemological and ontological issues, which are both at stake in this debate (as mentioned earlier).
[267] It may be possible to argue that these features of evolutionary theory also create problems for nativist views more generally; however, this type of argument is beyond the scope of this dissertation.

certain traits or even the survival of some organisms. Finally, all types of natural selection are also connected to random changes that occur in both the organism (e.g., genetic mutations) and the environment (e.g., changes in climate). The evolutionary process itself is non-directional, that is, it is non-teleological. So although natural selection may be a natural law, it is not a precise universal law. Unlike other scientific theories, and the scientific explanations generated from them, evolutionary theory contains no universal laws or rules that create the uniformity and predictability found in many other areas of science. [268] It is in this sense then, that evolutionary theory can be considered "irregular."

In addition to the irregularity of evolutionary theory, it is also historical because it provides a natural history that accounts for the existence of living organisms. As such, evolutionary explanations use a set of past events (e.g., a mutation, a change in environment, etc.) to explain the existence of some currently existing trait (e.g. the brain size of homo sapiens) or a formerly existing trait (e.g., the brain size of homo neanderthalensis). These explanations can only account for current or past traits but cannot predict future traits.[269] Any predictions using evolutionary explanations would merely be hypothetical, or what Michael Scriven refers to as a "hypothetical probability prediction" in his discussion of the roles of explanation and predication in evolutionary theory (1959: 478). Such predictions concern what is likely to occur given a certain set of hypothesis (e.g., what the environment will be like in the future, which genetic changes might occur, etc.) but not what will actually occur. This inability to predict is compounded by the fact that evolution is an irregular process, as previously described. The

[268] For example, physics could be considered "regular" in that it contains universal laws or rules that create uniformity and predictability.

[269] Some argue that predictions are possible but these types of "predictions" are only about what you will or can expect to find in a population now based on past events (some use the term "retrodiction" to refer to this type of prediction). However, you still cannot make accurate predictions about *future* cases.

evolutionary process is inherently random and thus accurate predictions are impossible. As Scriven points out: "…it is a feature of the irregular subjects that, unlike classical atomic physics, the irregularity-producing factors lie outside their range of observation and are not predictable by reference to any factors within this range" (1959: 478). As a result, hypothetical probability predictions are also non-falsifiable and thus, in terms of making scientific predictions, they are essentially useless.[270]

The inability to use evolutionary theory to predict future cases puts limitations on evolutionary explanations and raises a unique concern when it comes to morality. Debunking arguments use the evolutionary explanations of an innate moral sense to account for the capacity to make moral judgments and, thus, the nature of moral judgments themselves. The exact nature of moral judgments (or even more generally evaluative judgments) is clearly up for debate. However, one uncontentious aspect of moral judgments (and perhaps morality more generally) is that we make moral judgments not just about past events, or current events, but future events as well. For example, we might make the moral judgment that killing an innocent child *was* wrong, or that killing an innocent child *is* wrong, or that killing an innocent child *would* be wrong. Moral (evaluative) judgments are sometimes about future cases, so any explanation of those judgments should also be capable of accounting for future cases. Evolutionary explanations cannot, even in principle, account for moral judgments about future cases because evolutionary theory cannot make accurate predictions. So, even if we could explain why we have evolved to make our current or past moral judgments based on past events, we cannot predict what our future moral judgments ought to be based on these evolutionary explanations. The evolutionary

[270] As Scriven points out, "hypothetical predictions do not have any value for actual prediction…hence there will be cases where we can explain why certain animals and plants survived even when we could not have predicted that they would" (1959: 478). He also argues that in some cases, such as evolution, you cannot even predict *in principle*.

explanation can account for what is or what was but not for what will be because it is incapable of making accurate predictions.

Someone might argue that although the evolutionary explanation cannot account for future moral judgments, it could still account for the capacity for making moral judgments since it is the capacity that makes the judgment about future cases, not the evolutionary explanation itself. In other words, the inability to predict future events is a problem for direct evolutionary explanations of moral judgments but it is not a problem for indirect explanations of moral judgments i.e., the capacity to make moral judgments.

The problem with this objection ties into a more general issue about biological explanations and debunking arguments recently addressed by Andreas L. Mogenson. Mogensen points out that debunking arguments are often guilty of confusing proximate causes with ultimate causes. He describes the difference between these causes as follows:

> Proximate causes are causes of a trait that operates within an organism's own lifetime: these might include the immediate triggering causes or the developmental factors responsible for its acquisition and expression. Ultimate causes belong to its evolutionary history: an explanation in terms of natural selection or phylogeny is an explanation in terms of ultimate causes. (2015: 197)

The debunkers' claim that we have evolved to have a capacity to make moral judgments via truth-indifferent selection pressures is an evolutionary explanation about the ultimate causes of moral judgments. However, the claim that *all* moral beliefs (moral judgments) are, therefore, truth-indifferent does not follow from this. Mogensen argues that moral beliefs (moral judgments) also have proximate causes. So, even if we have evolved to have a capacity to make moral judgments that do not require moral truth (facts), it is still possible that proximate causes

exist that require moral truth (facts).[271] Since evolutionary explanations contain a combination of both proximate and ultimate causes, a full explanation must account for both types of causes that would play a role in any evolutionary explanation of any trait, including morality. Mogensen argues that the debunking arguments of Ruse, Joyce, and Street all fail to account for proximate causes, and as a result, they are guilty of jumping to the conclusion that all moral beliefs (judgments) evolved to be truth-independent.[272]

This distinction between proximate causes and ultimate causes is relevant to the objection regarding indirect moral judgments about future cases via the evolved capacity for moral judgments. Just as any adequate evolutionary explanation requires consideration of both ultimate and proximate causes, it requires both direct and indirect explanations as well. So, even if we have evolved a capacity to make moral judgments that includes the ability to predict future cases (despite evolution's inability to predict future cases) because it was fitness enhancing, this evolutionary explanation will only include the ultimate causes of moral judgments. The claim that *all* moral beliefs (moral judgments) are, therefore, capable of handling future cases does not follow because a full explanation must include proximate causes as well. And unlike the case of truth and moral facts, it is not possible that proximate causes will be capable of including future cases because any direct evolutionary explanations cannot, even in principle, account for moral judgments about future cases. As previously mentioned, a satisfactory explanation of moral judgments should be capable of accounting for future cases because moral (evaluative) judgments are sometimes about future cases and yet evolutionary explanations are limited in this respect.

---

[271] Mogensen, of course, is not implying that those moral facts will be required to explain proximate causes, only that it is possible and that the debunking arguments fail to address these types of causes.

[272] For the full argument, see Mogensen (2015: 199-203).

The irregular and historical nature of evolution raises another significant concern. Evolution is not only a complex process, but also one that relies on knowledge of past events that make it extremely difficult, if not impossible, to provide a complete account of any given trait. One might worry that there is an epistemic problem associated with *all* evolutionary explanations—since we can never know the entire evolutionary history of any given trait, evolutionary explanations are inherently flawed, and thus, incapable of providing knowledge. Consequently, one might be tempted to disregard all evolutionary explanations. But many scientists (and philosophers) have argued that evolutionary explanations are still capable of providing scientifically acceptable explanations that are justified and informative even if they are always incomplete. In other words, incomplete explanations can be sufficient in certain cases when they provide a satisfactory account of the phenomena in question.[273]

So, we cannot throw out all evolutionary explanations on account of the impossibility of full epistemic access. But we should take a more moderate attitude towards evolutionary explanations. The partial, irregular, and historical nature of these explanations should give us pause when considering how we want to use these explanations to do other things, such as making claims about how to change human behavior or, more generally, making claims about human nature or supporting other theoretical views.[274] Since these explanations are sufficient, rather than complete, they are only successful for a particular inquiry. Sufficient explanations are only useful (reliable or perhaps true) in some cases and not all cases. Take, for example, an evolutionary explanation of the human eye. A sufficient explanation of the human eye is not necessarily a sufficient explanation of human eyesight, and would definitely not be sufficient to

[273] See Scriven (1959: 480).
[274] The problem of full epistemic access argument would defeat all of these arguments because it would be rejecting all evolutionary explanations in principle; however, that is not the kind of argument I want to make, as this would mean not entering the debate at all.

explain a much broader concept such as color. There are, therefore, certain limitations inherent in these types of explanations. And as a result, caution is required to avoid mistakenly overextending the reach of these explanations. This type of mistake occurs when evolutionary explanations about particular traits are used to explain more general, universal or abstract features of any organism. While the explanation may be sufficient to explain a particular trait there is no reason to assume that it is sufficient for explaining a more general feature of the organism (or the species as a whole). The debunking debate is guilty of just such a mistake.

As previously discussed, the debunking debate treats morality as a trait that is susceptible to an evolutionary explanation.[275] The debunking arguments of Ruse, Street, and Joyce all propose some version of an innate moral sense, which is an evolved mechanism that includes the capacity to make moral judgments. The exact nature of moral judgments, however, is a metaethical issue since there are epistemological, semantic, and ontological issues associated with moral judgments. What's at stake in the debunking debate is not the existence of an evolved innate moral sense, but the ability of the moral judgments (or moral beliefs) produced by an innate moral sense to connect with the moral truth (if there is any moral truth). Even if the evolutionary explanations they provide for an innate moral sense are correct, it cannot be assumed that an explanation sufficient to explain an innate moral sense also includes a sufficient explanation that accounts for the full nature of moral judgments, or more broadly the ontological structure of morality and our epistemic access to it. This would be similar to the case mentioned above, where one attempts to explain the nature of color via an evolutionary explanation of the human eye. Such an explanation would not only be incomplete, but also wildly insufficient. Although some of the debunking arguments, like Joyce, are careful to insist that they are merely

[275] Of course this is a contentious claim, but following the principle of charity, assume that morality is an evolved trait.

raising an epistemological problem and not attempting to reach any particular metaethical conclusions directly from these evolutionary explanations, the epistemological problems raised still include metaethical assumptions based on the evolutionary explanation of the moral sense. In other words, the epistemological problem is meant to be a result of the evolutionary explanation. Thus, even the mildest forms of debunking attempt to extend the evolutionary explanation of the capacity to make moral judgments to also explain the metaethical nature of those moral judgments, and in some cases morality more generally.[276]

The problem of inferential opaqueness—the opaqueness of the connection between evolutionary explanations (the premises) and the rejection (or acceptance) of moral realism (the conclusion)—stems then not only from the particular nature of evolutionary explanations, but also from the nature of explanations more generally and the relationship between explanations and justification discussed in Chapter One. Recall that the success of any explanation is intimately linked to the initial inquiry. As a result, there are various levels of explanation for the phenomena; the success of an explanation depends on the level of explanation being sought and the level of explanation being sought depends on the initial inquiry and the goal of that inquiry. Even when an evolutionary explanation for one aspect of an evolved trait turns out to be satisfactory, that same explanation may not be satisfactory for another aspect of the same trait. We have seen this already with the example of the human eye, an issue that is only compounded by the irregular and historical nature of evolutionary explanations. In the same way, the evolutionary explanations used in debunking arguments (on both sides of the debate) often get

[276] Metaethics covers a wide range of issues, including but not limited to epistemological, ontological, and semantic questions. It is not my goal to define the exact nature of metaethics, but merely to point out that evolutionary explanations are not useful for explaining metaethics, whatever it may be.

mistakenly overextended as described above, which causes the connection between evolutionary explanations and metaethical conclusions to become confused. In other words, when we look more closely at the connection between evolutionary explanations and metaethical claims, we notice that the evolutionary explanation fails to explain the metaethical claims successfully. Even a sufficient evolutionary explanation of the moral sense (an evolved innate moral capacity) cannot explain the metaethical nature of morality more generally.

Furthermore, we see that the goal of the inquiry was not merely explanatory but justificatory as well. The relationship between explanation and justification, also discussed in Chapter One, therefore, needs to be revisited.

Remember, explanations and justifications are different because they have distinct purposes. Unlike explanations, justifications provide the "should" or "ought" story behind a phenomena rather than the "why" and "how".[277] Moreover, the goal of a justification differs from the goal of an explanation. Justifications aim to show why something is right or correct, whereas explanations aim to provide clarification and understanding. As such, a justification is also a defense and requires (or includes) an argument while an explanation does not. Most importantly, there is no automatic transition between an explanatory statement and a justificatory statement; you cannot automatically conflate the two types of statements since extra work needs to be done to turn an explanation (or explanatory reasons) into a justification (or justificatory reasons). Additional information is needed to show why these explanatory statements lend support to the justificatory statement.

In the case of debunking arguments, evolutionary explanations are used to justify metaethical conclusions (on both sides). Not only are these evolutionary explanations

[277] This is not to imply that justification only applies to future events; there are justifications regarding past events as well, i.e., what "should" or "ought" to have been done.

unsatisfactory because they cannot explain any metaethical claims sufficiently, but they also fail to provide the work—the additional information—required to turn explanations into justifications. This failure explains the opaqueness inherent in all evolutionary debunking arguments. The partial, irregular, and historical nature of evolutionary explanations makes them hopelessly incomplete, and thus the opaqueness cannot be remedied by any evolutionary explanations on their own. The only additional information available in these arguments are metahethical assumptions, and it should be clear at this point that this is not a successful strategy.

There are two significant problems that arise when using metaethical assumptions to convert evolutionary explanations into metaethical justifications in debunking arguments. First, if the additional information used to connect the evolutionary explanations to a metaethical conclusion is a set of metaethical assumptions, the argument simply begs the question. This was the problem raised earlier in this chapter in Section B.i on theoretical assumptions: assuming the metaethical conclusions in the premises of the argument is simply a case of circular reasoning—a problem the debunking debate seems hopelessly trapped in. Second, and more significantly, if metaethical assumptions are used to bridge the gap between the evolutionary explanations and the metaethical conclusion, then the evolutionary explanations are not the relevant explanations doing the work to turn the explanation into a justification. Instead, the metaethical assumptions, which are not explained or supported by any evolutionary explanation, are responsible for the metaethical conclusion. As a result, the evolutionary explanations do not actually play any significant role in the debunking arguments. In fact, they are not essential to the debate at all and, as such, could be eliminated from the debate without changing the structure of the argument. So, additional metaethical assumptions result in either a circular argument or in an argument that

does not depend on any evolutionary explanations—an odd place to be if you are presenting an evolutionary debunking argument or responding to one!

Street recognizes that evolution is not essential to her debunking argument. She argues that her Darwinian Dilemma represents not just an evolutionary argument against moral realism, but also a more general causal argument against the realist. As such, Street claims that you could substitute any other scientifically viable causal explanation for the Darwinian hypothesis, and a similar problem will arise for the realist. Street, therefore, may not need to worry about the problems I've raised here because they are focused on the particular nature of evolutionary explanations and the limitation inherent in them. But, she would need to modify her argument significantly, eliminating any evolutionary explanations from her argument and provide a causal explanation that might lead to the same antirealist conclusions she reached with the Darwinian Dilemma. As of now, Street has not presented such an argument, and I remain skeptical that any will be forthcoming. Furthermore, since my concern is with evolutionary explanations and not scientific explanations of morality more generally, the fact that evolutionary explanations are not necessary to make her debunking argument simply reinforces my point—that evolutionary explanations do not play any role in determining metaethical conclusions.

Although Joyce does not relinquish evolutionary explanations as readily as Street, he admits that there are significant numbers of metaethical assumptions being made in the debunking debate. In "Ethics and Evolution", Joyce clarifies his debunking argument from 2006, adding that:

> [W]hile I have argued that empirical evolutionary discoveries are sufficient to create a substantive burden for the moralist, I also recognize the need to appeal to *a priori* metaethical methods in bolstering the challenge. (2013a: 19)

Thus Joyce acknowledges that certain metaethical claims, independent of any evolutionary explanations, are needed to fully establish his debunking argument. What he fails to recognize is

that the burden or challenge for the realist would not exist without them. The burden or challenge for the realist Joyce refers to is an *epistemological* challenge regarding the *truth* of moral beliefs and judgments produced by a moral sense and *epistemic access* to that *truth*. So, without epistemological conceptions such as "truth", "belief", "judgment", etc., and some view about what counts as a justified belief, it is not possible to establish the challenge. If evolutionary explanations were sufficient to create a burden for the moral realist, then there would be no need to appeal to *a priori* metaethical methods. The evolutionary explanations are not doing the work in debunking arguments as Joyce claims.

Joyce has also argued that evolutionary explanations (or scientific data more generally) cannot resolve debates concerning the cornerstone of the debunking argument—moral nativism. In his article "Many Moral Nativisms", Joyce points out that the concept "moral judgments" is indeterminate and, thus, sufficiently pliable to allow for different legitimate specifications. As a result, he says:

> One possible consequence of this is that on some legitimate conceptions of *moral judgment* moral nativism is true, but on other equally legitimate conceptions moral nativism is false. And if there is no satisfactory way of deciding among these conceptions then the debate over moral nativism would be undecidable—not just in the sense that we lack decisive data, but in the sense that there is really no fact of the matter. (2013b: 550)

This indeterminacy means that more than one "morality" can fit with the evolutionary facts. So, the best-case scenario is that evolutionary explanations of morality are incomplete. The worst-case scenario is that evolutionary explanations are empty and uninformative. Either way, evolutionary explanations will not resolve the debate over moral nativism. If the issue of moral nativism cannot be settled, then it is curious why anyone, even the realist, would need to take any evolutionary debunking argument seriously. And, the realist should not attempt to establish their metaethical conclusion based on evolutionary explanations either or risk falling into the same trap. I will return to these points again and explore their implications in Chapter Five.

C. Conclusion

Darwinian realists take the challenge proposed by debunkers seriously and attempt to find various ways to explain why moral truths (facts) are necessary to explaining moral beliefs despite those beliefs' evolutionary genealogy. What I have shown in this chapter is not only that such attempts to answer the debunking challenge contain flawed arguments, but also that closer examination of the debunking debate reveals information about the nature of evolutionary explanations that both sides overlook—that evolutionary explanations are not capable of settling the metaethical issues that are at stake in this debate due to their incomplete, irregular, and historical nature.

Chapter Five: The Limitations of Evolutionary Ethics

In this final chapter, I will present the final arguments against evolutionary debunking—the genetic fallacy, a challenge to skepticism, and the indeterminacy of moral concepts. These arguments reveal the inability of the debunking debate to resolve the issues the debunkers and their opponents seek to settle, which includes whether or not to accept or reject moral realism or moral antirealism, or alternatively, whether or not to accept moral skepticism. Furthermore, this discussion will highlight the limitations of evolutionary ethics more generally, even with its new metaethical twist, while providing a direction for a more productive debate in metaethics.

A. The Genetic Fallacy

I have yet to fully address what might seem like the most obvious counterargument to debunking arguments—the genetic fallacy. Genetic reasoning is a type of reasoning in which a claim is either undermined or supported based on the historical or causal origin of that claim. This type of reasoning is fallacious when the historical or causal origin of the claim in question is irrelevant to the evaluation (i.e., the truth or falsity) of the claim.[278] In other words, a claim should not be evaluated on where it came from but rather on its current status. Both sides of the evolutionary debunking debate (and more generally evolutionary ethics) use genetic reasoning—in the debunking debate the evolutionary origins of moral beliefs either undermine or vindicate certain metaethical positions. So why can't these arguments simply be dismissed based on the genetic fallacy?

Proponents on both sides of the debunking debate reject the genetic fallacy on the following grounds: according to evolutionary ethicists, the evolutionary origins of a moral belief are relevant to current status of certain beliefs and, therefore, considering such origins does not

[278] Sometimes the genetic fallacy is classified as an ad hominem (Copi 1953). For a debate on the classification of the genetic fallacy, see Klement (2002).

amount to a fallacious form of genetic reasoning. Since the claims in question are beliefs (or judgments) about morality, the argument is framed as an epistemological issue in which the belief formation process is considered relevant to the status of the moral beliefs. According to the debate, both the truth of the belief itself and/or whether or not the believer is justified in holding that belief are influenced by the evolutionary history that is responsible (either directly or indirectly) for the belief in question. Consequently, the evolutionary origins of a moral belief are considered relevant to the evaluation of such beliefs. Setting aside the issues of evolutionary explanations of the origin of a belief, this general strategy is fairly common in epistemological debates concerning the justification of beliefs. It is generally accepted that how a belief is formed is relevant to whether or not the believer is justified in holding that belief (independently of the truth or falsity of the belief). So, for example, if someone forms a belief from a random, and thus unreliable, process such as flipping a coin, then the belief they hold is not justified regardless of the truth or falsity of that belief.[279] Taken a little further, some even argue that there are cases in which the causal history of a belief directly affects the truth or falsity of the belief itself and not simply the status of the believer. In such cases, the truth or falsity of a belief is inferred based on the origins of the belief.[280]

Looking at two types of genetic reasoning, those that aim to vindicate the truth of a belief and undermine (debunk) the truth of a belief, Klement (2002) differentiates between arguments that propose a direct causal history of a belief and those that propose a belief forming process that results in a belief. In the vindicating arguments, the direct causal history can guarantee the

[279] See Kahane (2011: 105-6).
[280] Klement (2002) points out that this is a much more controversial position. Also note that since these non-fallacious forms of the genetic arguments consider the belief forming process, they may not easily translate to genetic arguments at the level of normative prescriptive theory (e.g., theory debunking).

truth of a belief, whereas a highly reliable belief forming process only makes it more likely that a belief is true. Klement provides the following argument form to illustrate an example of this type of non-fallacious genetic argument:

> Person(s) S believes *p*.
> S's belief that *p* is the result of belief-forming process *f*.
> Belief-forming process *f* is *highly reliable*, i.e., it produces true beliefs much more often than it produces false beliefs.
> Therefore, S's belief that *p* is true. (2002: 387)

In the case of debunking arguments, guaranteeing the truth or falsity of a belief is much more difficult. A belief that is not the result of a causal chain is not necessarily false, and a belief that is not the result of a reliable belief forming process is not necessarily likely to be false. In these debunking cases, the issue is usually about the justification for holding the belief rather than the truth or falsity the belief. To guarantee that a belief is false, the belief forming process would need to be not simply unreliable but anti-reliable—that is, much more likely to produce false beliefs than true beliefs. Thus Klement provides the following example of an inductively strong form of this type of negative argument:

> Person(s) S believes *p*.
> S's belief that *p* is the result of belief-forming process *f*.
> Belief-forming process *f* is *highly anti-reliable*, i.e., it produces false beliefs much more often than it produces true beliefs.
> Therefore, S's belief that *p* is false. (2002: 387)

This argument form, however, is rare. According to Klement, it is extremely difficult to present a plausible account of what would make our beliefs not simply unreliable but anti-reliable.[281]

Turning back to the debunking arguments of both Ruse and Street, a brief analysis reveals that they both offer arguments that fail to meet the criteria for the limited non-fallacious

[281] See Klement (2002: 386-89).

forms of genetic reasoning.[282] Ruse's debunking argument does not directly target moral beliefs, but rather the subjective feelings and sentiments that produce moral beliefs (they affect what we think about morality). He argues that since the subjective feelings and sentiments that produce moral beliefs possess evolutionary origins, all moral beliefs (i.e., moral judgments) are false. More specifically, these moral beliefs are false in regards to their objectivity—we think our moral beliefs are objectively true, but this is false due to the evolutionary origins of those beliefs. So, Ruse attempts to argue that moral beliefs are formed from a process that reliably produces false beliefs about morality that appear objectively true. Following Klement's argument forms, Ruse's argument looks something like this:

> Person S believes moral proposition *p.*
> S's belief that *p* is the result of a belief-forming process *f* (an evolved innate moral sense produces subjective feelings and sentiments which in turn produce moral beliefs).
> Belief-forming process *f reliably produces false beliefs*.
> Therefore, S *falsely* believes that *p* is *objectively true*.

[282] These moral realists, who want to rely on evolutionary explanations to establish the truth or the justification of moral beliefs, also use genetic reasoning. Again, using Klement's positive form of the genetic argument, the evolutionary vindicating argument would take the following form:

> Person S believed moral proposition p.
> S's belief that *p* is the result of a belief-forming process *f,* which is an evolved innate moral sense (belief forming mechanism).
> Belief-forming process *f* is a *highly reliable process*, i.e. it produces true beliefs much more often than it produces false beliefs.
> Therefore, S's belief that *p* is justified.

Thus, any vindicating argument that does not follow this form will be fallacious. The biggest hurdle for the vindicators (i.e., the moral realists) is establishing the claim that the belief-forming process is highly reliable. As previously discussed in Chapter Four, Section A, the most common strategy to refute debunking arguments is not to deny any of the debunkers' premises but to add an additional premise (a third factor explanation) that explains why, despite the unreliable belief-forming process, moral beliefs are true. For a discussion of the issues these types of approaches face, see Chapter Four, Section B.

In this case, the belief-forming process is not simply anti-reliable, and thus likely to produce false beliefs, it also guarantees, according to Ruse, the falsity of these beliefs and that such beliefs will appear to be objectively true to the believer.

Before addressing the argument form, it is relevant to point out that Ruse fails to establish the claim that the belief-forming process reliably produces false beliefs—his only argument is that the moral sense (which is indirectly responsible for moral beliefs) evolved independently of the truth (i.e., the moral sense was selected for because it was fitness-enhancing not truth-preserving). Even if this account were correct, it is not sufficient to establish that the moral sense reliably produces *false* beliefs; there is no reason to assume that being fitness-enhancing implies that beliefs will be false. Similarly, if natural selection is not sensitive to the truth of moral beliefs, Ruse cannot argue that the belief forming process is anti-reliable (likely to produce more false beliefs than true beliefs). As such, a more charitable version of Ruse's argument could be presented as:

> Person S believes moral proposition *p*.
> S's belief that *p* is the result of a belief-forming process *f* (an evolved innate moral sense produces subjective feelings and sentiments which in turn produces moral beliefs).
> Belief-forming process *f* is *unreliable*, i.e., it produces beliefs in a manner that does not guarantee the truth.
> Therefore, S *falsely* believes that *p* is *objectively true*.

Two problems arise that reveal the fallacious form of this argument. First, an unreliable belief process does not guarantee the falsity of a belief, only that one may not be justified in holding a belief. Second, Ruse establishes this claim about the content of moral beliefs—that moral beliefs contain the belief of objectivity—not as the result of the evolutionary process that leads to these false moral beliefs but rather from an *a priori* claim that in order for morality to function, we must believe that our moral beliefs (i.e., moral judgments) are objective. As a result, Ruse's

argument commits the genetic fallacy—his claims extend well beyond the limits of the non-fallacious forms of genetic reasoning that are available.

Street's debunking argument does not fare any better. She argues that the evolutionary history of moral beliefs reveals a belief-forming process that is "off-track", and therefore, it is an unreliable process. Street's argument can be formulated as follows:

> Person S believes moral proposition *p*.
> S's belief that *p* is the result of a belief-forming process *f,* which is an evolved innate moral sense (belief-forming mechanism).
> Belief-forming process *f* is *unreliable*, i.e., it produces beliefs in a manner that does not guarantee the truth.
> Therefore, S's belief that *p* is *likely false*.

Notice that Street argues that the belief-forming process is *unreliable* and not *anti-reliable*. Thus, similar to Ruse's argument, Street's argument will run into trouble when attempting to establish the falsity of moral beliefs. Under the conditions of Street's own argument, it would be impossible to argue that the evolved belief-forming process (an innate moral sense) is anti-reliable since she claims that the truth or falsity of a belief is irrelevant to the natural selection process that produced it (hence why our moral beliefs are "off-track"). But if the truth or falsity of moral beliefs were irrelevant to the selection process, then it would be just as much of a lucky coincidence that the moral sense produced consistently false beliefs as true ones (or, that it is more likely to produce false beliefs than true beliefs). Street's conclusion, however, is that our moral beliefs are likely to be false. Based on her own argument, the innate moral sense that produces moral beliefs is merely unreliable, so the best she could argue is that we are unjustified in holding these beliefs, not that they are likely to be false. As such, Street's argument commits the genetic fallacy because she cannot establish her conclusion based on her proposed evolutionary genealogy of moral beliefs.

This, however, is not the end of Street's argument. Street presents an additional conclusion. She argues that since moral beliefs are likely to be false, moral realism is likely to be false, and thus moral antirealism is likely to be true. As just discussed, Street fails to establish the conclusion that moral beliefs are likely to be false, so any further claims based on this conclusion are no good. But even if moral beliefs were likely to be false, this wouldn't necessarily prove that moral realism was false or that moral antirealism was true. Further premises, which are not provided by the evolutionary explanations themselves, are needed to support this conclusion.

Finally, there is Joyce's debunking argument to consider. He argues that moral belief originates from an innate moral sense, which is an unreliable belief-forming mechanism, so moral beliefs are unjustified. At first glance, this argument appears to be the most successful since it is the only non-fallacious form of genetic reasoning in the debunking debate. Joyce's argument is as follows:

> Person S believes moral proposition *p*.
> S's belief that *p* is the result of a belief-forming process *f,* which is an evolved innate moral sense (belief-forming mechanism).
> Belief-forming process *f* is a *highly unreliable* process.
> Therefore, S's belief that *p* is *unjustified*.

As previously discussed, unlike Ruse and Street, Joyce does not attempt to establish that all moral beliefs produced by an evolved innate moral sense are false, but rather that all moral beliefs are *unjustified*. Setting aside any issue that might arise related to various accounts of justification that are up for debate, Joyce's argument appears to escape the problems faced by Ruse and Street.

Joyce, however, adds that since all moral beliefs are unjustified, moral skepticism follows. The further claim of moral skepticism is where Joyce runs into problems. The earlier version of Joyce's (2006) debunking argument argues for a moral skepticism akin to a version of moral

error theory. In this case, Joyce then extends his argument too far since the moral error theory he endorses (a form of moral fictionalism) claims not merely that all moral judgments are unjustified, but that all moral judgments are false. As a consequence, this early version of his argument would clearly count as a fallacious form of a genetic argument—he cannot conclude that all moral judgments are false from the premise that the belief-forming process is unreliable.

Later versions of Joyce's argument, however, take a step back from error theory and merely claim that moral skepticism is an agnostic position, one in which we have no reason to believe that moral beliefs are true or false—we cannot know if they are true or false because they are formed by an unreliable belief-forming process so all moral beliefs are unjustified. This, as Joyce says, is a more modest claim, and it does not appear to commit the genetic fallacy. To support what Joyce refers to as a "modest form" of his debunking argument, Joyce (2016b) argues that his debunking argument is meant simply as a challenge to the moral realist and not as a way to establish any particular metaethical position. The burden of proof, according to Joyce, lies with the moral realist to explain why we should think we are justified in holding our moral beliefs despite their evolutionary history.[283] He further argues that the likelihood of the moral realist meeting this challenge is very low.

In order to escape the genetic fallacy, Joyce must weaken his conclusion so that the moral skepticism he endorses only provides a reason to suspend judgment; it cannot establish the truth

[283] Note that the only moral realists that need be concerned are those that accept Joyce's (or any other modest debunker's) evolutionary explanation of an evolved innate moral sense. The target of the challenge is quite small and thus not very significant when placed in the larger context in the debate over moral realism. Furthermore, Joyce's epistemological challenge to moral realism is also not unique. It is merely a reinvention of the same standard objections to moral realism—that even if moral facts (or properties) exist and/or even if moral beliefs are true, the realist needs to explain how we justify our moral beliefs. These arguments are also a variation of the classical epistemological skeptical hypothesis (e.g., Descartes' evil genius) applied to morality, despite his denial of this claim, which I will discuss below.

value of moral beliefs. Being unjustified in holding a belief does not provide any grounds for or against moral realism (or antirealism)—it does not settle issue nor does it push the debate closer to a resolution. The much bolder conclusions of Ruse and Street about the truth or falsity of moral beliefs would push the debate forward; however, their conclusions cannot be established without committing the genetic fallacy. Joyce takes the conclusion that moral beliefs are unjustified, however, to be a problem for the moral realist but not for the antirealist. Joyce presents the challenge as something only the realist must worry about. But a genuinely agnostic position claims only that the truth values of moral beliefs are unknown (or, perhaps even unknowable).[284] There is, by definition, no reason to favor one side over the other. So if all moral beliefs are unjustified then the challenge is equal for both sides of the debate. Independent arguments are required from both sides of the debate to establish their position.

Joyce assumes that an unreliable belief-forming process is enough to push the debate in favor of the antirealist. As previously discussed, we must differentiate between belief-forming processes that are anti-reliable, unreliable, reliable, or reliably false. Unreliable processes produce beliefs in a manner that does not guarantee truth—but they do not guarantee falsity either. We have no reason to believe that the belief is true or false, so we must suspend judgment. Only a reliably false belief-forming process would guarantee that moral beliefs were false. And, only an anti-reliable belief-forming process would guarantee that moral beliefs are more likely to

[284] Although Joyce appears to only argue that the truth values of moral beliefs are unjustified, it is also possible that an agnostic view could hold that the truth conditions for moral beliefs are unknown. So, a moral agonistic could argue that not only do we not know if our moral beliefs are true or false, but we also don't know under what conditions moral beliefs are true or false.

be false. Thus, without additional reasons, an unreliable belief-forming process should render us agnostic about the beliefs, without reasons to favor one side over the other.[285]

One final issue emerges from Joyce's modest debunking that is worth considering since it is relevant to any form of the debunking. It is not clear why moral beliefs must remain unjustified, i.e., why moral beliefs are not revisable. Even if Joyce is correct and moral beliefs are initially derived from an unreliable process (again, which I have previously argued he fails to establish), this would not preclude the believer from using other belief-forming processes that *are* reliable to revise those beliefs. In fact, this seems a more accurate picture of how we come to hold justified beliefs. We likely form many true beliefs in unreliable ways, or at least in ways that are not directly reliable (e.g., taking for granted what an authority says is true). We revise these beliefs through a process of rational reflection until we become justified.

Take the case of beliefs about motion. Numerous studies have shown that commonsense beliefs about the natural motion of objects, which are primarily formed by direct perceptual observations, are systematically wrong.[286] For example, when a running person drops a ball, it continues to move forward at the same speed as the runner. However, research shows that people often mistake where the ball will land. McCloskey (1983) asked college students where a ball dropped by a walking person would land and found the following:

> Only 45 percent of the students knew the ball would travel forward as it fell. Forty-nine percent thought the ball would fall straight down and land directly under the point where it was released; 6 percent thought the ball would move backwards as it fell. (1983: 123)

[285] The debate over the truth conditions of moral beliefs seems to take for granted a realist theory of values. Joyce (2016b) argues that this is not necessary for the debunking debate. Some forms of debunking, such as Street's, do invoke a realist theory of value, where moral beliefs are both truth-apt and mind-independent. However, he claims his epistemological debunking assumes that moral beliefs are truth-apt but mind-dependent and, thus, does not require a realist theory of value.

[286] For a discussion of such studies, see Proffitt (1999).

In this case, more than half of the subjects held mistaken beliefs about the motion of the ball. Similar studies that asked subjects about the motion of falling objects, projectile motion of objects, and curvilinear motion of objects found similar results. Very often, subjects were wrong about their beliefs.[287]

McCloskey referred to such mistaken beliefs as an "intuitive physics" (also called naïve physics or folk physics).[288] Common sense beliefs about motion are formed from our direct perceptual observations. However, such beliefs tend to match the disproven medieval "impetus theory" of motion rather than Newtonian physics, which is the accepted theory of motion. McCloskey hypothesized that such mistaken beliefs were likely the result of misperceptions. In other words, perceptual beliefs cause beliefs about motion, but those perceptions are sometimes in error (e.g., illusions). These beliefs, however, are revisable. According to McCloskey, it is possible to revise mistaken beliefs about motion formed via perceptual observations by learning about Newtonian mechanics.[289] Such beliefs could also be revised through additional perceptual observations (e.g., observations of the same motion in different situations or comparing the mistaken belief with other perceptual beliefs). Thus, even the same mechanism that produced the mistaken belief can be used to revise that same belief.

What this example shows is that when beliefs are formed by an unreliable belief-forming process they can still be revised.[290] The belief-forming process that produces beliefs about motion is unreliable because we cannot tell the difference between correct or incorrect beliefs based on the method itself. Until the mistaken beliefs are pointed out, we cannot know if we are

[287] See McCloskey (1983).
[288] See Stich (1996) and Proffitt (1999).
[289] McCloskey acknowledges that such revision is not easy. But it is possible.
[290] I am suspicious of any view that claims beliefs cannot be revised. And, if Quine (1951/1980) is correct, then not only are all beliefs revisable, but no belief is immune from revision.

holding the correct beliefs directly from our perceptual observations of motion. However, when the mistaken beliefs are discovered, some outside independent reasons, such as learning about Newtonian physics, can fix our mistaken beliefs. Once these beliefs about are revised, we can have reliable beliefs about motion. So if other types of beliefs formed that were from unreliable belief-forming processes can be revised, why would moral beliefs be exempt from this type of revision?

Street (2006) argues that such revision to moral beliefs would still be wildly "off-track" because the initial beliefs were so "off-track" to start. Such beliefs are "off-track" because they are formed from a process that is truth independent. This feature, truth independence, makes the belief-formation process unreliable. In light of these considerations, Street considers the well-known process of reflective equilibrium as a means to get moral beliefs back "on-track" with the truth. According to Street, the rational process of reflective equilibrium involves evaluative judgments, so the process itself will also be "off-track". Street argues:

> And reflection of this kind isn't going to get one any closer to evaluative truth, any more than sorting through contaminated materials with contaminated tools is going to get one closer to purity. (2006: 124)

However, since the non-fallacious forms of debunking can only claim that the belief-forming process is unreliable, not anti-reliable or reliably false, there is no way to establish how "off-track" moral beliefs turn out to be. To guarantee that beliefs are so far "off-track" that revisions would be hopeless requires an anti-reliable process or a reliably false process—neither of which can be established via evolutionary explanations that claim truth independence. An anti-reliable process of belief formation would make it more likely that all moral beliefs were false while a reliably false process of belief formation would guarantee the falsity of moral beliefs. In the first case, it is more likely that moral beliefs are "off-track" than "on-track". Thus, moral beliefs are

likely to be far away from the truth. In the second case, moral beliefs are guaranteed to be "off-track" and, thus, far away from the truth. But since the belief-forming process is unrelated to the truth—the evolutionary explanation of moral beliefs claims truth independence—the process is merely unreliable. An unreliable belief-formation process produces beliefs in a manner that does not guarantee the truth, but it does not guarantee falsity either. Nor does it automatically make it more likely that such beliefs are either true or false. In this case, there is no way to determine if unreliable beliefs are true or false based solely on the process from which they were formed. As a result, an unreliable belief-forming process would lead to unjustified beliefs but it would remain unknown how far "off-track" or "on-track" these beliefs were.

If moral beliefs stem from an unreliable belief-forming process then Street cannot claim that such beliefs are so far "off-track" that they are hopeless—i.e., they are too far "off-track" to be revised.[291] The correct stance towards such beliefs should be one of uncertainty; we do not know yet how "on" or "off" track moral beliefs are. Deciding how "on" or "off" track beliefs are is just what the process of reflective equilibrium seeks to resolve. Reflective equilibrium revises beliefs by checking for consistency—that is, how well the beliefs in question cohere with one's entire set of beliefs. From what Street (and others in the debate) have argued, it is safe to assume that she takes many of our non-moral beliefs (e.g., basic perceptual beliefs, mathematical beliefs, scientific beliefs, etc.) to be "on-track". Moral beliefs are not only checked against other moral beliefs, but our entire set of beliefs. As a result, there is a way to get such beliefs back "on-track"

[291] Of course, one may have independent reason for rejecting the process of reflective equilibrium; however, if it is a legitimate process by which we can revise beliefs, then it cannot be rejected on the grounds that Street claims. I am also suspicious of the idea that beliefs that are way "off-track" beliefs cannot be revised, but I will not address that issue here and just take it for granted as part of the argument.

(or at least to discover how "on" or "off" track they turn out to be) by comparing them with "on-track" beliefs.

This, of course, would not satisfy a critic like Street because she also argues that the tools for reflective equilibrium should be suspect. Since the process of reflective equilibrium requires making evaluative judgments, the method itself turns out to be unreliable. However, if Street is correct, then she must reject the process for non-moral beliefs as well, which she may be willing to do. But rejecting all evaluative judgments includes rejecting even those evaluative judgments made in pursuit of science. As discussed in Chapter Four, this includes all arguments by inference to the best explanation, including her own claims that the best explanation for the evolution of moral beliefs is non-truth tracking. As such, Street cannot exclude moral beliefs (or the tools) from being revised based on other, independent grounds.[292] Since moral beliefs can be revised, there is no reason (within this debate) why unjustified moral beliefs could not become justified—so having unjustified moral beliefs need not lead to a prolonged state of moral skepticism.

B. The 'Skeptical Challenge' Challenge

Modest debunking arguments (like Joyce's later version) manage to avoid the genetic fallacy, but they do so at a great expense. At best, modest debunking arguments present a challenge to both sides of the debate: if the modest debunker can successfully establish her conclusion, then this challenge is for both moral realist and the moral antirealist alike. Both the moral realist and the moral antirealist require additional arguments to answer this challenge and to move on from skepticism. And while Joyce's modest version of his debunking argument

[292] The argument against reflective equilibrium also cannot entail that such revisions of moral beliefs are impossible on the grounds that moral beliefs are all false (or that antirealism is true). This begs the question.

cannot be easily dismissed on the grounds that it uses fallacious reasoning, there are still other aspects of his argument (and these types of argument) that raise concerns.

The primary worry is that debunking arguments undermine more than moral beliefs. If Joyce is correct that beliefs formed from an evolved innate moral sense are always unreliable, then it is also very likely that all of our beliefs are unreliable. This is a claim that Joyce and other debunkers deny. But, if all beliefs possess some type of evolutionary history (either directly or indirectly), and the evolutionary process (natural selection) functions independent of the truth, then all our beliefs are unjustified. I've already discussed this objection in Chapter Three, so I will not dwell on it here; however, it is something worth considering again. Any evolutionary debunking argument in morality runs the risk of global epistemological skepticism that undermines the justification of other beliefs (e.g., perceptual beliefs, thereby indirectly undermining beliefs about evolution) required to make the argument.

Katia Vavova presents a variation of this objection in her recent article on evolutionary debunking that deserves attention. She argues that the skepticism that debunkers aim for entails much more than they would like. She points out that all debunking arguments include some premise akin to the following:

> NO GOOD: If you have no good reason to think that your belief is true, then you cannot rationally maintain it. (2013: 6)

If this premise is true, then it applies not only to cases of moral (or evaluative) beliefs, but to all types of beliefs. "NO GOOD" provides criteria by which we can decide if we should maintain or reject our beliefs. However, this criterion leads to a much more inclusive skepticism than the debunker intended. This premise implies that we are required to have 'good' reasons to rationally maintain our beliefs. These 'good' reasons for maintaining a belief must also be independent

reasons; in other words, they cannot be the same as the beliefs that are under investigation. She then takes the case of perceptual beliefs to illustrate the problem:

> Unless we are skeptics, we should grant that sensory perception is a perfectly good belief forming method. *Ceteris paribus*, if you perceive that *p*, you are rational in concluding that *p*. Do we have good reason to think that perception would lead us to true beliefs about our surroundings? Not if 'good' reason is understood as an appropriately independent reason: for if we set aside all that is in question, we must set aside all beliefs gained by perception. This includes all scientific beliefs, like the beliefs that evolutionary theory is true. (2013: 7)

Thus, the standard of establishing 'good' reasons for our beliefs ends up calling into doubt all of our beliefs. And, as her example illustrates, if we can doubt beliefs about evolutionary theory, then the evolutionary debunker has inadvertently debunked her own argument.[293] We are led back to the original global epistemological skepticism that Sidgwick alerted us to in "The Theory of Evolution in its Application to Practice".

Of course, this is something debunkers like Joyce are not willing to embrace. Debunkers insist that this type of global epistemological skepticism is unfounded because there is an important distinction between moral and non-moral beliefs. Unlike moral beliefs, non-moral beliefs (e.g., perceptual beliefs, mathematical beliefs, scientific beliefs, etc.) are selected for via a process that is relevant to truth. As a result, non-moral beliefs stem from a reliable belief-forming process so they are justified.[294] Joyce (2006) further argues that the skepticism raised by the modest debunking argument is not a standard form of the skeptical challenge, like Descartes' Evil Demon or brain-in-vat scenarios, which result in extreme epistemological skepticism. These forms of the skeptical challenge are difficult to take seriously. The modest debunker's skeptical

[293] Vavova also argues that this problem is applicable to any debunking (or skeptical) argument, not just evolutionary debunking. See Vavova (2013: 6-8). Additionally, Vavova offers a possible fix for the debunker; they can establish 'good reasons' for thinking that moral beliefs are not true. She argues, however, that this approach is equally unsuccessful. I will return to her discussion in the next section.

[294] See discussion in Chapter Three for details of their argument.

challenge, according to Joyce, is not so easily dismissed because it rests on empirical evidence rather than a fabricated, outlandish possibility. In other words, the reason for moral skepticism is real whereas the other type is merely a far-fetched possibility. As Joyce says,

> It is not just that in this case we can *make up* a consistent hypothesis according to which a bunch of our ordinary beliefs are false; rather, it is that we might have empirical evidence supporting the hypothesis that explains how these beliefs came about but does not require that they be true…If the everyday standards for being morally justified take account of empirical data concerning human evolution, then if these data ultimately show moral beliefs to be unjustified it will be by ordinary epistemic standards. (2006: 186-87)

Claims about being empirically founded, however, will not get Joyce (or any form of debunking) off the hook so easily.

Entering into the debunking argument, there are certain background beliefs about the nature of morality that are necessary for the argument to work. We must already know something about moral beliefs in order for the argument to make sense. To start, we need to know what morality *is*—what is right or wrong, good or bad in regards to human behavior—so that we can differentiate between moral beliefs and non-moral beliefs. Evolutionary debunking arguments, for example, consider certain behaviors (e.g., altruistic acts) and emotions (e.g., sympathy) as moral rather than non-moral.[295] These also lead to a distinction between moral beliefs and non-moral beliefs, both of which are part of the evolutionary story that establishes the processes that form beliefs. This distinction is especially important to debunkers since they do not want to include all beliefs in their skeptical conclusion. Assumptions about the nature of moral beliefs (or judgments), including features of moral beliefs such as truth-aptness, are also part of the background beliefs about morality. As such, debunkers must presuppose certain claims about morality in order to establish the argument that leads to a skeptical conclusion.

[295] This holds for any form of modest debunking and not just evolutionary debunking arguments.

The problem with the skeptical argument should now be obvious—if beliefs about the nature of moral beliefs are themselves moral beliefs, and all moral beliefs are unjustified, then beliefs about the nature of moral beliefs must also be unjustified. The modest debunker faces her own irreconcilable dilemma: on the one hand, if she maintains that all moral beliefs are unjustified, then she cannot establish the skeptical challenge, which entails certain moral beliefs be justified; on the other hand, if she wants to maintain the skeptical challenge, she must accept that some moral beliefs are justified. In either case, the skeptical challenge is self-refuting—the truth of the challenge implies its falsity.[296] It cannot be true that all moral beliefs are unjustified.[297] As a result, even modest forms of debunking are unsuccessful.

The modest debunker might try to refute this challenge by arguing that beliefs *about* the nature of morality are not moral beliefs. Consequently, non-moral beliefs *about* morality can be justified even if all moral beliefs are unjustified, thereby avoiding the dilemma. But how can our beliefs about morality be justified if all of our moral beliefs are unjustified? The burden of proof lies with the debunker to show how beliefs about morality can be justified despite all moral beliefs being unjustified. If beliefs about morality are not moral beliefs, then what kind of beliefs are they? How are these beliefs formed? Is that belief-forming process reliable? Can we have a justified belief about unjustified beliefs? If the content of a belief contains an unjustified belief, how does this affect the belief? In other words, the debunker must explain how beliefs about the nature of moral beliefs are not only distinct from moral beliefs but also why they are justified. So

[296] Here, I have in mind something akin to Putnam's response to the 'brain-in-vat' skeptical hypothesis. See Putnam (1981: 1-21).
[297] I don't mean to imply that moral realism is true, only that if moral beliefs are truth-apt, then some moral beliefs are justified, whatever it means to be justified.

even modest forms debunking turn out to be problematic—not on the grounds of committing the genetic fallacy but by possibly presenting a self-refuting argument.[298]

The same problem applies the stronger debunking arguments as well. This stronger form of debunking in question is akin to Ruse's argument. If all moral beliefs are reliably false (or anti-reliable), then all beliefs about the nature of morality are reliably false (or anti-reliable). But all beliefs about the nature of morality cannot be reliably false (or anti-reliable). Therefore, all moral beliefs are not reliably false (or anti-reliable). The other ambitious debunking argument by Street merely claims that moral beliefs are formed from an unreliable process. In the case of Ruse, I have already argued that he cannot establish that moral beliefs are reliably false if the evolutionary process is truth independent. For the same reason, neither side can reasonably argue that moral beliefs are formed from an anti-reliable process either. Remember, they can at best establish that the belief-forming process is unreliable (that is why they are already guilty of the genetic fallacy). Thus, the stronger evolutionary debunking argument does not lead to the same skeptical problem as the modest debunking argument because it cannot establish its conclusion.

Despite the failure of the stronger debunking arguments to establish their conclusion (that moral beliefs are likely to be false) or the weaker modest conclusion (moral skepticism), Vavova (2013) presents a similar argument against the strongest forms of debunking that is worth considering. Her target is evolutionary debunking arguments that claim the belief-forming

[298] Some might object that this response to the skeptical challenge entails more than merely evolutionary debunking arguments. It seems to close off the epistemological route for establishing antirealist positions in metaethics by requiring that some moral beliefs be justified. Although, this is not my intention, I accept this possible ramification. Closing off the epistemological route does not, however, prove moral realism is correct, and it still leaves open other routes for the antirealist to establish her position (e.g., some forms of non-cognitivism).

process is anti-reliable. She shows how even the strongest forms of debunking still run into problems. [299]

Vavova agrees with my claim that the challenge is not simply for the moral realist. While I maintain that moral skepticism would lead to an equal split of the burden, she shifts the burden of proof to the debunker. The debunker must show that we have 'good' reasons to believe that our moral beliefs (or evaluative beliefs) are mistaken and not simply claim that we have 'no good' reason to believe them. The debunker needs to show that there are independent non-question begging reasons for thinking that moral beliefs are mistaken. These 'good' reasons cannot include any of the moral beliefs that are in question since the premise of the argument is that such moral beliefs are untrustworthy so they cannot be relied on.

The debunker's 'good' reason for thinking that moral beliefs are mistaken is that adaptive beliefs (beliefs selected for based on fitness) are independent of true moral (evaluative) beliefs. As a result, moral beliefs are "off-track"—they are selected on the basis of fitness not truth. Vavova argues that this 'good' reason, however, is not good after all. If we are to evaluate these 'good' reasons without any assumptions about the moral beliefs that are in question, then we cannot make any assumptions about what morality is like. But the debunker cannot establish the claims that we have 'good' reasons to believe that our moral beliefs are erroneous without the claim that adaptive beliefs and moral beliefs are independent. And, it is not possible to separate adaptive beliefs from moral beliefs without reference to the moral beliefs that are in question. If the debunker loosens the standards for what counts as a 'good' reason, then there is 'no good'

[299] It is worth considering because a non-evolutionary strong debunking argument that establishes anti-reliable (or reliably false) belief-forming processes is possible, even if unlikely. It seems unlikely that there is any reliably false or anti-reliable belief-forming processes that could not be revised once the believer became aware of the problem.

reason to think we are in error.[300] If the debunker keeps the standards for what counts as 'good' reasons, then she cannot talk about morality; there is no way to establish the premises necessary to reach the conclusion.[301]

Vavova has shown that, like the modest debunking arguments I discussed, the stronger versions of debunking are equally self-refuting. Even if the evolutionary explanations were removed (as they should be because they are problematic for other reasons already discussed), it is impossible to construct a skeptical challenge to morality without making some reference to morality. For example, one might provide a historical account of the origins of moral beliefs that gives 'good' reason to think that moral beliefs were formed in an unreliable manner. The historical debunker concludes that all moral beliefs are unjustified. But the same problem arises—in order to establish the claim that moral beliefs (rather than non-moral beliefs) are formed in an unreliable way, we need to know something about morality to identify which beliefs are moral. Thus, all moral debunking arguments require some assumptions about morality. The evolutionary debunking, therefore, cannot establish the skeptical challenge that it claims to.

C. Evolutionary Explanations, Moral Assumptions & the Indeterminacy of Moral Concepts

At this point, it should be clear that the use of evolutionary explanations in both the arguments intended to debunk (or undermine) metaethical positions and their respective counterarguments are highly problematic. Setting aside particular worries about the arguments of Ruse, Joyce, and Street and the realist counterarguments of Enoch and Wielenberg discussed in Chapters Three and Four, one common issue emerges: the debunking debate relies on a set of explanations that are not capable of supporting the various metaethical claims that are at stake.

---

[300] Vavova then argues that means we have good reason to think we are not mistaken, but I disagree that this follows automatically; it leaves it an open question that requires further justification.

[301] See Vavova (2013: 8-15).

As argued in Chapter Four, evolutionary explanations cannot provide adequate explanations of the moral phenomena in question (e.g., an evolved innate moral sense or the moral beliefs and judgments produced by a moral sense) due to the nature of the explanations itself. As a result, these explanations are limited and cannot provide justification for any metaethical view. The initial debunking arguments, and the debate that follows, cannot get off the ground if evolutionary explanations cannot adequately address the moral phenomena in question (i.e., if the *explanans* does not provide adequate clarification for the *explanandum*). Reliance on evolutionary explanations to resolve issues in metaethics is, therefore, a simple non-starter.

Interestingly, both Street (2006) and Joyce (2016) have made a point to stress that their debunking arguments need not hinge on evolutionary explanations. As previously mentioned, Street (2006) argues that any causal argument will suffice to debunk morality. More recently Joyce has said,

> The first thing to note about this sketched argument is that the evolutionary perspective is, strictly, dispensable. Were we to explain our moral beliefs by reference to, say, developmental and socialization processes, then, so long as these processes similarly nowhere imply or presuppose that our or anyone else's moral judgments are true, the same epistemological conclusion could be drawn. (2016: 2)

Such claims appear inconsistent with the evolutionary debunking arguments they have both presented (and in the case of Joyce, defended on several occasions).[302] If evolutionary explanations are not a necessary part of debunking arguments, then such explanations do not play a vital role in the debate over moral realism. Instead, debunking is merely a general strategy that any empirically minded antirealist might want to employ. This leads to a further issue related to empirical (or naturalistic) ethics and another inconsistency in Joyce's position that exemplifies the central problem with the debate.

[302] See Joyce (2016a), (2016b), (2013a), (2013b), (2013c), and (2006).

Joyce (2013b) has argued that the use of scientific data or empirical facts to resolve conceptual debates in philosophy has certain limitations. Conceptual debates in morality—for example, how to define a moral judgment—cannot be settled with any empirical data. In fact, Joyce goes so far as to claim that a certain level of indeterminacy will always persist within these debates because conceptual issues cannot be resolved via empirical data. In regards to the nature of moral judgments he says,

> It's not a matter of our not knowing which is the correct conception (because we lack data); it's that there is no unique fact of the matter. (2013b: 16)

To illustrate this problem, Joyce presents two equally viable conceptions of moral judgments, a *liberal* conception and a *strict* conception. The *liberal,* less demanding, conception of moral judgments identifies moral judgments as emotional traits that possess more general psychological roles. The *strict,* more demanding conception describes moral judgments as possessing strong cognitive elements.[303] These competing conceptions of moral judgments have different implications. A *liberal* conception of moral judgments tends to support anti-nativist views, whereas a *strict* conception tends to support nativist views. According to Joyce, both conceptions are legitimate. Joyce extends these considerations to differing conceptions of *moral rightness* to make his point:

> In other words, there may be some legitimate precisification of the concept *moral rightness* (for example) according to which rightness is a real property of certain actions; but there may be other equally legitimate precisifications according to which no such property exists anywhere. (2013b: 17)

Differing conceptions of *moral rightness* lend support to opposing views on moral realism. Neither conception, however, can be ruled out by any amount of empirical data.

[303] Joyce (2013b) says these are general categories; there will be variations of each conception, but broadly speaking they will fit into one or the other conception.

Consequently, the debate over moral nativism cannot be resolved on the basis of empirical facts alone. Whether or not moral nativism is correct will depend on the conception of moral judgment employed. Even if the scientific debate regarding these matters were resolved, two different and equally viable conceptions of moral judgments may persist. As a result, Joyce says,

> The upshot is that both moral nativism and moral non-nativism may be perfectly defensible positions, and may remain so even when all data are in. (2013b: 18)

Empirical data (which would include evolutionary explanations) cannot settle these disputes because more than one legitimate conception of the concept in question may be available and we have no good way, or at least a good way that relies on empirical data, of deciding between competing conceptions. According to Joyce, there are indeterminate concepts, like 'moral judgments' and 'moral rightness', that play a significant role in the moral nativist debate, and as a result there is no amount of data that will settle the debate. Thus, much like the Quinean picture of science, the data underdetermines the theories in question, and we must choose which theory to follow on independent grounds.[304] Joyce recognizes that this might be an unpopular position, especially among naturalists, but he thinks it is nonetheless true.[305]

I happen to agree with Joyce's claims here—that there will be cases, especially in moral philosophy, where no amount of empirical data can resolve what is a purely conceptual, or theoretical, debate. Analyzing the standard distinction between descriptive and normative ethics helps clarify this point. Any work done in the descriptive realm, which is the empirical study of our moral beliefs and practices, requires some notion of morality before analyzing the data. All empirical data are not morally significant, thus some criteria must be used to distinguish between

[304] See Quine (1975).
[305] See Joyce (2013b).

data that counts as moral and data that does not.[306] The task of descriptive ethics is not to describe all beliefs people hold (or practices they perform), but rather to focus on those beliefs (and practices) that are considered moral. For example, the belief that 'water is H_2O' is typically not a moral belief, whereas the belief that 'waterboarding is wrong' is typically a moral belief. [307] And, it is not every evaluative belief that descriptive ethics considers either but a particular set of evaluative beliefs that are considered moral. For example, the belief that 'water is good' is typically a non-moral belief, whereas the belief that 'waterboarding is bad' is typically a moral belief. There is much debate over which data are moral and which are not moral, and that is not an issue I aim to resolve here.[308] My point is simply that moral concepts are required in order to pick out (from the available data) which data is moral and which is not moral. Those moral concepts must exist *before* interpreting the data. Even if one argues that they are informed by the data, there still must be some concept of morality, at some other level (or perhaps a higher-order concept) that one brings to the table before sorting through the data. Determining what morality is, or what moral concepts are, is, therefore, a debate (or at least a significant portion of the debate) that takes place on purely conceptual grounds.[309] As a result, designating the subject

[306] I may be taking the claim all data is not moral for granted, but I have a hard time taking any objection to this claim seriously; would anyone really argue that something like "there are three jelly beans on the table" is moral?

[307] I'm not trying to argue for any particular criterion, only that there is a criterion used to make this kind of distinction and that we do make these kinds of distinctions in descriptive ethics.

[308] The relationship between the descriptive and the normative is also contentious, but that is irrelevant; regardless of how you think they inform or interact with one another, some idea about what counts as morality is required before the debate commences.

[309] Even naturalist versions of morality must have some vision of what counts as moral. Generally, the naturalist realist argues that moral properties are identical to some natural properties, while the naturalist antirealist argues that since moral properties reduce to natural properties, moral properties are not real (or are unnecessary). Regardless of one's position, the meaning of *morality* must be established at the outset.

matter presupposes some moral concepts, and no amount of empirical data will resolve these conceptual issues.

What I find curious about Joyce's position is how he can support the claim that no amount of empirical data can resolve such moral debates while arguing for moral skepticism via an evolutionary debunking argument. His evolutionary debunking argument rests on premises that consist of a set of scientific claims that are, to use Street's phrase, "saturated" with moral concepts. In fact, the entire debunking debate is filled with scientific data that is "saturated" with moral concepts. Focusing on Joyce's argument, the evolutionary explanation of an innate *moral* sense that produces *moral* beliefs evolved independently of the *moral* truth. This is the reason we cannot trust our *moral* beliefs (i.e., we are not justified in holding moral beliefs). Starting with the scientific data—Joyce's evolutionary explanation of a moral sense—he must employ particular moral concepts, such as moral judgments, to discuss the evolution of a moral sense. Again, this is something that Joyce does not deny, but acknowledging this will not protect him against the issues associated with it. The same applies to the other positions in the debunking debate.

My first concern was already addressed in the previous section. The need to assume moral concepts renders all evolutionary debunking arguments fallacious because they must assume the truth of the classification of the concepts they aim to undermine. Simply put, they are self-refuting. But this is only a problem for the debunkers. My second concern applies to both debunkers and vindicators alike. Joyce, or any philosopher involved in the debate, could be cherry-picking moral concepts. That is, they could easily be selecting particular moral concepts that support their intended conclusion. Decisions made at the conceptual level will ultimately affect the conclusion, so one could choose moral concepts that are more favorable to their view.

As Joyce himself points out, different conceptions of *moral judgments* support different positions regarding moral nativism, and different conceptions of *moral rightness* support different positions regarding moral realism. It is easy to see how someone arguing for or against moral nativism (or for or against moral realism) would simply favor the moral conception that best fits her position, and not necessarily for any good reason. The same goes for deciding which evolutionary explanations (or, more broadly, empirical data), one selects. I raised a similar concern earlier in regards to Darwin's view of the moral sense in Chapter Two and again in Chapter Four with respect to the use of evolutionary explanations. Of course, just as in the case of Darwin and/or the use of evolutionary explanations, this wouldn't make his or any view *wrong,* but it should raise some suspicion and cause us to proceed with caution.

My final concern, however, considers a more significant problem facing evolutionary ethics. If the scientific explanation includes moral concepts, then there is, as Joyce has pointed out, no amount of empirical data that will resolve the debate between competing accounts of an innate moral sense (including the arguments against an innate moral sense). These competing accounts will need to be decided on independent grounds and, when it comes to moral concepts, at a purely theoretical level. What this reveals is that the evolutionary explanations do little to direct the debate.[310] They do so little, that Joyce and Street are willing to replace them with other types of empirical explanations. Street concludes her discussion of her proposed Darwinian Dilemma by stating the following:

> At the end of the day, then, the dilemma at hand is not distinctly Darwinian, but much larger. Ultimately, the fact that there are any good scientific explanations of our evaluative judgments is a problem for the realist about value. (2006: 155)

And as just mentioned above, Joyce (2016b) has recently has argued that evolutionary explanations are dispensable. If such explanations are dispensable or replaceable, then it is

[310] Or, possibly, whatever scientific explanation one substitutes.

evident that the evolutionary explanations do not resolve the metaethical debate. Furthermore, if any reason for doubting moral beliefs will do, then other empirical explanations are equally as dispensable. Both Street and Joyce accept this line of reasoning. As Street (2006) says, any good scientific explanation of evaluative judgment that raises doubt for the moral realist is sufficient to get the argument started. Similarly, Joyce suggests that other empirical explanations could replace evolutionary explanations. He even extends beyond the natural sciences and includes sociological explanations as another candidate for explanation—in Joyce's words "so long as these processes similarly nowhere imply or presuppose that our or anyone else's moral judgments are true, the same epistemological conclusion could be drawn" (2016b: 2).

The dispensability of evolutionary explanations clearly shows that evolution is not an essential element of the debate. But the dispensability of all empirical explanations tells us something more. Debunkers are seeking any empirical story that raises doubt. However, just as Joyce pointed out with the case of moral nativism, whether or not the empirical data can undermine (or vindicate) moral beliefs depends on what we think about moral beliefs. As a result, the debate over moral realism cannot be resolved on the basis of empirical facts alone. Whether or not moral realism or antirealism is correct will depend on the conception of moral belief employed. Even if the scientific debate regarding these matters were resolved, two different and equally viable conceptions of moral beliefs may persist. In other words, the debate is not taking place at the empirical level but at the conceptual level. Such explanations are, therefore, not sufficient to claim that we are or are not justified in holding moral beliefs or to support any form of moral skepticism or any other metaethical position.

With these considerations in mind, it becomes obvious that the foundation of Joyce's moral skepticism (or any similar argument on either side of the debunking debate) rests on

premises that cannot be secured via data. It is futile to argue over evolutionary explanations of moral beliefs (or more broadly of morality) when such explanations play no significant role in determining the outcome of the debate. *The real debate then is not about the evolutionary origins of morality but rather the moral concepts that are used to understand and interpret the empirical data.*

D. The Limitations of Evolutionary Ethics and Future Directions

The evolutionary debunking debate attempts to use evolutionary explanations to reach metaethical conclusions, either undermining moral realism, vindicating moral realism, vindicating moral antirealism, or vindicating moral skepticism (i.e., skepticism about moral realism). This approach, however, not only fails to resolve this metaethical debate successfully, but also shows little promise of ever resolving these issues. Using evolutionary explanations to reach metaethical conclusions is a flawed method, and therefore should be avoided.

In this essay, I've raised many problems that both sides of the debunking debate must contend with. A brief overview of some of the central problems for evolutionary metaethical approaches, however, is sufficient to show why such an approach will be fruitless. To begin, most debunking arguments (and their vindicating responses) are logically flawed because they are guilty of committing the genetic fallacy, despite their denial of this. The few debunking views that escape this fallacy are saddled with another logical problem: they will likely be self-refuting.

Two more insurmountable problems arise for both debunking and vindicating arguments. First, the evolutionary explanations employed in the debate (on both sides) are more limited than they think. As previously discussed, certain features, such as being both irregular and historical (and often incomplete), limit the ability of evolutionary explanations to explain certain

phenomena adequately, which includes metaethical claims. As a result, evolutionary explanations are not capable of justifying the metaethical conclusions that are stake in the debate. Second, the theoretical assumptions required to enter the debate reveal not only that the evolutionary explanations are not essential to the debate, but also that they are dispensable. The theoretical assumptions, in particular moral concepts, are what affect the metaethical conclusions, not the evolutionary explanations. From these three problems, neither side of the debate is capable of establishing their positions via evolutionary explanations.

While the focus of this essay has been on evolutionary metaethics, many of the same problems may extend to evolutionary ethics more broadly construed. For example, normative (prescriptive) views in evolutionary ethics must also confront the genetic fallacy. And while some views might escape this logical problem, the same two insurmountable issues that metaethical views face will remain. As discussed in Chapter Four, evolutionary explanations are equally unsuccessful in explaining normative (prescriptive) claims adequately as they are for explaining metaethical claims adequately. As a result, evolutionary explanations may not be helpful in resolving debates over competing moral theories or even more specifically, how we ought to behave.[311] Certain theoretical assumptions, in particular moral concepts, are also a necessary feature of normative (prescriptive) views in ethics. Again, as I previously argued, if evolutionary explanations cannot fully account for moral concepts, then these explanations are not likely to play a significant role in supporting any particular ethical theory. Evolutionary ethics is therefore, limited by which moral concepts are employed, and thus whatever is decided at this theoretical level regarding those concepts will inform how we interpret the evolutionary

[311] Here, I have in mind views like Singer (2005) and Greene (2003) and (2008), which attempt to argue that evolutionary explanations support consequentialist theories and undermine deontological theories. For a discussion of some objections to these views, see Selim Berker (2009) and Kahane (2011).

explanations being used.[312] Although this brief analysis is not sufficient to show that the entire field of evolutionary ethics—metaethics, normative ethics and prescriptive ethics— fails, it provides good reason to be suspicious of it. And, even though it is beyond the scope of this essay, I suspect that evolutionary ethics can only at best be a very limited enterprise.

While many are drawn to evolutionary ethics, I hope to have shown that it is not a useful approach to resolving metaethical issues. Farber's concern that there is little promise in contemporary versions of evolutionary ethics was correct but not because, as he claimed, "most contemporary arguments are simply recycled versions of earlier ones" (1994: 2). Contemporary debates in evolutionary ethics are distinct from their predecessors—many battle it out at the metaethical level rather than at the level of prescriptive theory. However, shifting the debate into the realm of metaethics does not help these approaches succeed. The use of evolutionary explanations in the meatethical debate over moral realism turns out to be a non-starter. So, there is still good reason to think that there is little promise for such evolutionary approaches.

What these debates inadvertently reveal is where metaethical debates should be taking place—at the theoretical level. Both the moral assumptions required for the debate and the indeterminacy of moral concepts force us to remain on the theoretical level. In order to establish any metaethical conclusions, we must first figure out what we are arguing about; we must first determine what morality *is*. This means developing views about the content of morality (how we distinguish between moral and non-moral), as well as the nature of morality (which includes the nature of moral beliefs, judgments, etc.). Disputes over these conceptual views of morality will not be settled via evolutionary explanations, nor will they be resolved by any empirical data. Therefore, the debate must take place at the theoretical level.

[312] For a discussion regarding concerns that the use of empirical data often comes "philosophically loaded", see Appiah (2008) and Berker (2009).

Many, especially anyone with naturalist leanings, will worry that this commits us to ignoring science and traps us into doing "armchair" philosophy. The lesson, however, should not be to ignore scientific explanations but to accept their limitations. The nature of the discourse constrains us; the best we can do is to attempt to establish good reasons for accepting or rejecting certain moral concepts via theoretical debate. That is all we can do. And this should come as a relief to philosophers, for that is what we are trained to do.

Bibliography

Alexander R.D. (1987). The Biology of Moral Systems. Aldine de Gruyter, New York.

Allhoff, F. (2003). Evolutionary Ethics from Darwin to Moore. History and Philosophy of the Life Sciences, Vol. 25, No. 1, 51-79.

Axelrod, R. (1984). The Evolution of Cooperation. New York: Basic Books.

Ayala, F. J. (2010). The Difference of Being Human: Morality. Proceedings of the National Academy of Sciences of the United States of America, 107 (Supplement 2), 9015–22.

Bailey, T. (2013). Nietzsche the Kantian? In J. Richardson and K. Gemes (eds.), The Oxford Handbook on Nietzsche, (134-59). Oxford: Oxford University Press.

Bain, A. (1872). Mental and Moral Science. London: Longmans, Green and Co.

Barrett, J. (1991). Really Taking Darwin and the Naturalistic Fallacy Seriously: An Objection to Rottschaefer and Martinsen. Biology and Philosophy, Vol. 6, Issue 4, 433-37.

Baumrin, B.H. (1968). Is There a Naturalistic Fallacy? American Philosophical Quarterly, Vol.5. No. 2, 79-89

—— (1986). Moral Blindness. Metaphilosophy, Vol. 17, No. 4, 205-13.

Blackburn, S. (1993). Essays in Quasi-Realism. New York: Oxford University Press.

—— (2000). Ruling Passions. New York: Oxford University Press.

—— (2009). Pascal's Wager. In Steven M. Cahn (ed.), Exploring Philosophy: An Introductory Anthology, (231-34). Oxford University Press.

BonJour, L. and Sosa, E. (2003). Epistemic Justification. Internalism vs. Externalism, Foundations vs. Virtues. Malden, MA: Blackwell.

Boyd, R. (1988). How to Be a Moral Realist. In G. Sayre-McCord (ed.), Essays on Moral Realism , (181-228). Ithaca: Cornell University Press.

Braddock, M. (2012). Debunking Challenges to Moral Realism. Unpublished dissertation thesis, Duke University, Durham, NC.

Brink, D.O. (1989). Moral Realism and the Foundations of Ethics. Cambridge: Cambridge University Press.

Buss, D.M. (1999). Evolutionary Psychology: The New Science of Mind. Boston: Allyn and Bacon.

—— (2004). Evolutionary Psychology: The New Science of the Mind. New York: Pearson Education.

Butterflied, H. (1931/1965). The Whig Interpretation of History. New York: W.W. Norton & Company.

Carruthers, P. (1992). Human Knowledge and Human Nature. Oxford: Oxford University Press.

Clarke-Doane, J. (2012). Morality and Mathematics: The Evolutionary Challenge. Ethics, Vol. 122, No. 2, 313-40.

Copi, I. (1953). Introduction to Logic. New York: Macmillan.

Copp, D. (2008). Darwinian Skepticism About Moral Realism. Philosophical Issues, 18, 184-204.

Cottingham, J. (1991). Descartes. Malden, MA: Wiley-Blackwell.

Couzens Hoy, D. (1983). Nietzsche, Hume and the Genealogical Method. In Y. Yovel (ed.), Nietzsche as Affirmative Thinker: Papers Presented at the Fifth Jerusalem Philosophical Encounter, 251-53.

Crisp, R. (2006). Reasons and the Good. New York: Oxford University Press.

Cummins, D. (1996). Evidence of Deontic Reasoning in 3 and 4-year old Children. Memory and Cognition, 24:6, 823-29.

Darwall, S., Gibbard, A. and Railton, P. (1992). Toward Fin de siècle Ethics: Some Trends. The Philosophical Review. Vol. 101, No. 1, 115-89.

Darwin, C. (1838-1839). *Notebook N:* Metaphysics and Expression. In K. Rookmaaker (transcr.) and P.H. Barrett (ed.), *Darwin Online,* CUL-DAR126.

—— (1838-1851). 'Books to be Read' and 'Books Read' Notebook. In K. Rookmaaker (transcr.), Darwin Online, CUL-DAR119.

—— (1838-1858). 'Books [read]' Notebook. In K. Rookmaaker (transcr.) and J. van Wyhe (ed.), Darwin Online, CUL-DAR120.

—— (1852-1860). 'Books Read' and 'Books to be Read' Notebook. In K. Rookmaaker (transcr.), Darwin Online, CUL-DAR128.

—— (1861). On the Origin of the Species by Means of Natural Selection. London: John Murray.

—— (1871). The Descent of Man and Selection in Relation to Sex. New York: D. Appleton and Company.

—— (1881). Letter to William Graham. Darwin Correspondence Database. URL = <https://www.darwinproject.ac.uk/entry-13230>.

—— (1980). Metaphysics, Materialism, and the Evolution of the Mind: Early Writings of Charles Darwin, P.H. Barrett and H. E. Gruber (transcrs.), USA: The University of Chicago Press.

Dawkins, R. (1976). The Selfish Gene. Oxford: Oxford University Press.

Descartes, R. (1641/1996). Meditations on First Philosophy. In J. Cottingham (ed.), Descartes: Meditations on First Philosophy: With Selections from the Objections and Replies, (1-62). Cambridge: Cambridge University Press.

de Waal, F.B.M, and Luttrell, L. (1988). Mechanisms of Social Reciprocity in Three Primate Species; Symmetrical Relationship Characteristics or Cognition. Etiology and Sociobiology, 9, 101-18.

Doney, W. (1955). The Cartesian Circle. Journal of the History of Ideas 16, No. 3, 324-38.

Driver, J. (2012). Consequentialism. New York: Routledge.

Dworkin, R. (1996). Objectivity and Truth: You'd Better Believe It. Philosophy and Public Affairs 25, 87-139.

Dwyer, S. (1999). Moral Competence. In K. Murasugi and R. Stainton (eds.), Philosophy and Linguistics, (169-90). Boulder, CO: Westview Press.

Elbe, S. (2001). 'We Good Europeans…': Genealogical Reflections on the Ideas of Europe. Millennium - Journal of International Studies, 30, 259-83.

Enoch, D. (2010). The Epistemological Challenge to Metanormative Realism: How Best to Understand It, and How to Cope with It. Philosophical Studies, 148 (3), 413-38.

Farber, P. L. (1994). The Temptations of Evolutionary Ethics. Berkeley: University of California Press.

Field, H. (1989). Realism, Mathematics, and Modality. New York: Blackwell Publishing.

FitzPatrick, W. (2014). Morality and Evolutionary Biology. In E. N. Zalta (ed.), The Stanford Encyclopedia of Philosophy. URL= <http://plato.stanford.edu/archives/fall2014/entries/morality-biology/>.

Fodor, J. (2000). The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology. Cambridge, MA: MIT Press.

Fodor, J. and Piattelli-Palmarini, M. (2010). What Darwin Got Wrong. New York: New York: Picador.

Gewirth, A. (1970). The Cartesian Circle Reconsidered. Journal of Philosophy, 67, No. 19, 668-85.

Gibbard, A. (1982). Human Evolution and the Sense of Justice. Midwest Studies in Philosophy, 7, 31–46.

—— (1990). Wise Choices, Apt Feelings. Cambridge, Mass.: Harvard University Press.

—— (2003). Thinking How to Live. Cambridge, MA: Harvard University Press.

Gill, M. (2007). Moral Rationalism vs. Moral Sentimentalism: Is Morality more like Math or Beauty? Philosophy Compass, Vol. 2, Issue 1, 16-30.

—— (2010). Ethics and Sentiment. In J. Skorupski (ed.), The Routledge Companion to Ethics. New York: Routledge.

Gould, S.J. (1978). Sociobiology: The Art of Storytelling. New Scientist 80 (1129), 530-33.

—— (1997). Darwinian Fundamentalism. New York Review of Books, 44 (10), 34–7.

Greene, J. (2003). From Neural 'Is' to Moral 'Ought': What are the Moral Implications of Neuroscientific Moral Psychology? Nature Reviews Neuroscience 4, 847-50.

—— (2008). The Secret Joke of Kant's Soul. In W. Sinnott-Armstrong (ed.), Moral Psychology: The Neuroscience of Morality, Vol. 3, (35-79). Cambridge, MA: MIT Press

Griffiths, P. (2002) What is Innateness? Monist, 85, 70-85.

Geuss, R. (1994). Nietzsche and Genealogy. European Journal of Philosophy, 2, Issue 3, 274-92.

Hamilton, W.D. (1963). The Evolution of Altruistic Behavior. American Naturalist 97, 354-56.

—— (1964). The Genetical Evolution of Social Behavior, I and II. Journal of Theoretical Biology 7, 1-52.

Harman, G. (1977). The Nature of Morality: An Introduction to Ethics. USA: Oxford University Press.

—— (1999). Moral Philosophy and Linguistics. In K. Brinkmann (ed.), Proceedings of the 20th World Congress of Philosophy: Volume 1: Ethics, (107-15). Bowling Green, Ohio: Philosophy Documentation Center.

Harman, O. (2012). Is the Naturalistic Fallacy Dead? (And if so, Ought it Be?) Journal of History and Biology, 45, 557-72.

Hobbes, T. (1651/2009). Leviathan. USA: Oxford University Press.

Hudson, W. D. (1969). The Is-Ought Question: A Collection of Papers on the Central Problem in Moral Philosophy. London: Macmillan.

Huemer, M. (2005). Ethical Intuitionism. New York: Palgrave Macmillan.

Hume, D. (1738/1978). Treatise on Human Nature, L.A. Selby-Bigge (ed.), New York: Oxford at the Clarendon Press.

—— (1751/1991). An Enquiry Concerning Principles of Morals. In D.D. Raphael (ed.), British Moralists: Vol. II: Hume—Bentham and Index, (59-94). Indianapolis: Hackett Publishing Company.

Huntley, W. B. (1972). David Hume and Charles Darwin. Journal of the History of Ideas, Vol. 33, No. 3, 457-70.

Hutcheson, F. (1726/2004). An Inquiry into the Original of Our Ideas of Beauty and Virtue in Two Treatises. W. Leidhold (ed.), Indianapolis: Liberty Fund. URL=<http://oll.libertyfund.org/titles/2462>.

—— (1738/1991). An Inquiry Concerning Moral Good and Moral Evil. In D.D. Raphael (ed.), British Moralists: Vol. I: Hobbes—Gay, (261-99). Indianapolis: Hackett Publishing Company.

—— (1742/1991). Illustrations upon the Moral Sense. In D.D. Raphael (ed.), British Moralists: Vol. I: Hobbes—Gay, (305-21). Indianapolis: Hackett Publishing Company.

James, S. M. (2010). An Introduction to Evolutionary Ethics. USA: Wiley-Blackwell.

Joyce, R. (2006). Is Human Morality Innate? In P. Carruthers, S. Laurence & S. Stich (eds.), The Innate Mind: Culture and Cognition, (257-79). Oxford, UK: Oxford University Press.

—— (2005). Moral Fictionalism. In M.E. Kalderon (ed.), Fictionalism in Metaphysics, (287-313). New York: Oxford University Press.

—— (2006). The Evolution of Morality. Cambridge, MA: The MIT Press.

—— (2013a). Ethics and Evolution. In H. LaFollette and I. Persson (eds.), The Blackwell Guide to Ethical Theory, (123-47). Malden, MA: Wiley-Blackwell.

—— (2013b). The Many Moral Nativisms. In K. Sterelny, R. Joyce, B. Calcott, and B. Fraser (eds.), Cooperation and its Evolution, (549-72). Cambridge, MA: MIT Press.

—— (2013c). The Evolutionary Debunking of Morality. In J. Feinberg and R. Shafer-Landau (eds.), Reason and Responsibility, (527-34). Boston, MA: Cengage.

—— (2015). Moral Anti-Realism. In E. N. Zalta (ed.), The Stanford Encyclopedia of Philosophy. URL = <http://plato.stanford.edu/archives/fall2015/entries/moral-anti-realism/>.

—— (2016a). Evolution, Truth-tracking, and Moral Skepticism. In R. Joyce (ed.), Essays in Moral Skepticism, New York: Oxford University Press. Draft available at: URL= <http://personal .victoria.ac.nz/richard_joyce/acrobat/joyce_2016_evolution.truthtracking.moral.skepticism.pdf>

—— (2016b). Reply: Confessions of a Modest Debunker. In U. Leibowitz and N. Sinclair (eds.), Explanation in Mathematics and Ethics, New York: Oxford University Press. Draft available at: URL=<http://personal.victoria.ac.nz/richard_joyce/acrobat/joyce_2016_confessions.of.a.modest.debunker.pdf>.

Kahane, G. (2011). Evolutionary Debunking Arguments. Nous, 45 (1), 103-25.

Kalderon, M. (2005). Moral Fictionalism. Oxford: Clarendon Press.

Kant, I. (1785/1993). Grounding for the Metaphysics of Morals, J. W. Ellington (trans.). Indianapolis: Hackett Publishing Company, Inc.

—— (1886). The Metaphysics of Ethics, J.W. Semple (ed. and trans.) with Introduction by Rev. Henry Calderwood. Edinburgh: T. & T. Clark.

Kitcher, P. (1993). The Evolution of Human Altruism. Journal of Philosophy, 90, 497–519.

—— (2006). Biology and Ethics. In D. Copp (ed.), The Oxford Handbook of Ethical Theory, (163–85). New York: Oxford University Press**.**

—— (2011). The Ethical Project. Cambridge, MA: Harvard University Press.

Klement, K.C. (2002). When is Genetic Reasoning Not Fallacious? Argumentation, 16, 303-400.

Korsgaard, C. M. (2010). Reflections on the Evolution of Morality. The Amherst Lecture in Philosophy, 5, 1–29.

Krimsky, S. and Gruber, J. (eds.). (2013). Genetic Explanations: Sense and Nonsense. USA: Presidents and Fellows of Harvard College.

Lamark, J.B. (1809/1984). Zoological Philosophy. An Exposition with Regard to the Natural History of Animals, H. Elliot (trans.). Chicago: University of Chicago Press.

Lamb, T. D. (2011). Evolution of the Eye. Scientific American, Vol. 305, Issue 1, 64-69.

Lecky, W. E. H. (1869). History of European Morals from Augustus to Charlemagne, Vol. 1. New York: D. Appleton and Company.

Leiter, B. (2000). Nietzsche's Metaethics: Against the Privilege Readings. European Journal of Philosophy, 8, Issue 3, 277-97.

—— (2002). Nietzsche on Morality. London: Routledge.

—— (2004) The Hermeneutics of Suspicion: Recovering Marx, Nietzsche and Freud. In B. Lieter (ed.), The Future of Philosophy, (74-105). Oxford: Clarendon Press.

Lewens, T. (2007). Darwin. New York: Routledge.

Lewis, D. (1972). Psychophysical and Theoretical Identifications. Australasian Journal of Philosophy, 50, 249–58.

Lillehammer, H. (2003). Debunking Morality: Evolutionary Naturalism and Moral Error Theory. Biology and Philosophy, 18 (4), 567-81.

—— (2004). V—Moral Error Theory. Proceedings of the Aristotelian Society, 104, 95–111.

Lumsden, C.J. and Wilson, E.O. (1981). Genes, Mind, and Culture: The Coevolutionary Process. Cambridge, MA: Harvard University Press.

Machan, T.R. (1978) Herbert Spencer: A Century Later. In H. Spencer The Principles of Ethics Vol. 1. Indianapolis: Liberty Classics. URL=< http://oll.libertyfund.org/pages/machan-spencer-a-century-later>.

Machery, E. and R. Mallon (2010). Evolution of Morality. In J. Doris (ed.), The Moral Psychology Handbook. Oxford: Oxford University Press.

Mackie, J. (1977). Ethics: Inventing Right and Wrong. London: Penguin.

Mameli, M. and Bateson, P. (2007). The Innate and the Acquired: Useful Clusters or a Residual Distinction from Folk Biology? Developmental Psychobiology, 49, 818-31.

Mameli, M. (2008). On Innateness: The Clutter Hypothesis and the Cluster Hypothesis. Journal of Philosophy, 105, 719-37.

Marx, K. and Engles, F. (1845-88/1998) The German Ideology: Including Thesis on Feuerbach and a Critique of Political Economy. New York: Prometheus Books.

Mason, K. (2010) Debunking Arguments and the Genealogy of Religion and Morality. Philosophy Compass, 5, 570-78.

Mikhail, J. (2002). Law, Science, and Morality: A Review of Richard Posner's *The Problematics of Moral and Legal Theory.* Stanford Law Review, 54, 1057-1127.

Mogensen, A.L. (2015). Evolutionary Debunking Arguments and the Proximate/Ultimate Distinction. Analysis, 75 (2), 196-203.

Moore, G.E. (1903/2004). Principia Ethica. New York: Dover Publications, Inc.

Murphy, J.G. (1982). Evolution, Morality, and the Meaning of Life. Totowa, NJ: Rowman and Littlefield.

Nagel, T. (1997). The Last Word. Oxford, UK: Oxford University Press.

—— (1986). The View from Nowhere. Oxford: Oxford University Press.

Nichols, S. (2005). Innateness and Moral Psychology. In P. Carruthers, S. Laurence, and S. Stich (eds.), The Innate Mind: Structure and Contents, (353-70). New York: Oxford University Press.

Nietzsche, F. (1887/1994). On the Genealogy of Morals, K. Ansell-Pearson (ed.) and C. Diethe (trans.). Cambridge: Cambridge University Press.

—— (1882/2001). The Gay Science, B. Williams (ed.) and J. Nauckhoff and A. Del Caro (trans.). Cambridge: Cambridge University Press.

—— (1878/1986). Human, All too Human, R.J. Hollingdale (trans.). Cambridge: Cambridge University Press.

Nozick, R. (1981). Philosophical Explanations. Cambridge, MA: Belknap Press of Harvard University Press.

Okasha, S. (2013). Biological Altruism. In E. N. Zalta (ed.), The Stanford Encyclopedia of Philosophy. URL =<http://plato.stanford.edu/archives/fall2013/entries/altruism-biological/>.

Owen, D. (1997). Maturity and Modernity: Nietzsche, Weber, Foucault, and the Ambivalence of Reason. New York: Routledge.

Pigden, C., (ed.). (2010). Hume on Is and Ought. New York: Palgrave Macmillian.

Pinker, S. (2002). The Blank Slate. New York: Penguin Group.

—— (2006). Evolution and Ethics. In J. Brockman (ed.), Intelligent Thought. New York: Vintage Books.

Pitt, J., (ed.). (1988). Theories of Explanation. New York: Oxford University Press.

Plantinga, A. (2002). Naturalism Defeated. In Beilby (ed.), Naturalism Defeated: Essays on Plantinga's Evolutionary Argument Against Naturalism. New York: Cornell University Press.

—— (2000). Warranted Christian Beliefs. New York: Oxford University Press.

Plato. Phaedo. (360 B.C.E. /2002). In G.M.A Grube (trans.), Five Dialogues, (93-154). Canada: Hackett Publishing Company, Inc.

—— Euthyphro. (360 B.C.E. /2002). In G.M.A Grube (trans.), Five Dialogues, (1-20). Canada: Hackett Publishing Company, Inc.

Prinz, J. (2008). Is Morality Innate? Moral Psychology, 1, 367-406.

—— (2009a). Against Moral Nativism. In D. Murphy and M. Bishop (eds.), Stich and his Critics, (167-89). Malden, MA: Wiley-Blackwell.

—— (2009b). The Emotional Construction of Morals. New York: Oxford University Press.

Quine, W.V.O. (1975). On Empirically Equivalent Systems of the World. Erkenntnis, 9: 313–28.

—— (1966). On Simple Theories of a Complex World. In The Ways of Paradox. New York: Random House.

—— (1981). Theories and Things. Cambridge, MA: Harvard University Press.

—— (1951/1980). Two Dogmas of Empiricism. In A Logical Point of View, (20-46). Cambridge, MA: Harvard University Press.

—— (1955/1980). Posits and Reality. In A Logical Point of View, (246-54). Cambridge, MA: Harvard University Press.

Rabinowitz, A. (2012). Ethical Realism and the Darwinian Dilemma. Unpublished master's thesis, Colorado State University, Fort Collins, Colorado.

Railton, P. (1986). Moral Realism. Philosophical Review 95, 163-207.

—— (2000). Darwinian Building Blocks. In L. Katz (ed.), Evolutionary Origins of Morality. Thorverton: Imprint Academic.

Ramsey, W. (2002). Naturalism Defended. In Beilby (ed.), Naturalism Defeated: Essays on Plantinga's Evolutionary Argument Against Naturalism. New York: Cornell University Press.

Raphael D.D., (ed.). (1991). British Moralists: Vol. I: Hobbes—Gay. Indianapolis: Hackett Publishing Company.

—— (1991). British Moralists: Vol. II: Hume—Bentham and Index. Indianapolis: Hackett Publishing Company.

Richards, R. (1986). A Defense of Evolutionary Ethics. Biology and Philosophy, 1, 265-93.

—— (1987). Darwin and the Emergence of Evolutionary Theories of Mind and Behavior. Chicago: The University of Chicago Press.

Rousseau, J. (1791/1923). The Social Contract and Discourses by Jean-Jacques Rousseau, G.D. H. Cole (trans.). London and Toronto: J.M. Dent and Sons.

Ruben, D., (ed.). (1993*).* Explanation. Oxford: Oxford University Press.

Ruse, M. (1986). Evolutionary Ethics: A Phoenix Arisen. Zygon. 95-112.

—— (1986b). Taking Darwin Seriously. Oxford: Basil Blackwell.

—— (1995). Evolutionary Naturalism. London: Routledge.

—— (2006) Is Darwinian Metaethics Possible (And If It Is, Is It Well Taken)? In G. Boniolo and G. De Anna (eds.), Evolutionary Ethics and Contemporary Biology, (13-26). New York: Cambridge University Press.

—— (2009) The History of Evolutionary Thought. In M. Ruse and J. Travis (eds.), Evolution: The First 4 Billion Years, (1-48). USA: President and Fellows of Harvard College.

—— (2009b). Philosophy After Darwin: Classic and Cotemporary Readings. USA: Princeton University Press.

—— (2009c). "Homo Sapiens": from "Altruism" to Altruism. In M. Ruse (ed.), Philosophy After Darwin: Classic and Cotemporary Readings, (489-509). USA: Princeton University Press.

Salmon, W. (1984). Scientific Explanation and the Causal Structure of the World. Princeton, NJ: Princeton University Press.

—— (1989). Four Decades of Scientific Explanation. Minneapolis: University of Minnesota Press.

Sayre-McCord, G. (no year). Moral Realism. URL=<http://www.unc.edu/~gsmunc/Papers/Moral%20Realism.pdf.>.

—— (1995). Coherentist Epistemology and Moral Theory. In W. Sinnott-Armstrong and M. Timmons (eds.), Moral Knowledge? New Readings in Moral Epistemology, (137–89). Oxford: Oxford University Press.

Scanlon, T.M. (1998). What We Owe to Each Other. Cambridge: Harvard University Press.

Scriven, M. (1959). Explanation and Prediction in Evolutionary Theory. Science Vol. 130, No 3374, 477-82.

Shafer-Landau, R. (2003). Moral Realism: A Defense. Oxford: Oxford University Press.

Shaftesbury, A. A. C. (1699/1991). An Inquiry Into Virtue and Merit. In D.D. Raphael (ed.), British Moralists: Vol. I: Hobbes—Gay, (169-88). Indianapolis: Hackett Publishing Company.

Sidgwick, H. (1874/1981). The Methods of Ethics. Indianapolis: Hackett Publishing Company.

—— (1876). Evolutionary Theory and in its Application to Practice. Mind 1, No. 1. 52-67.

—— (1886). The Historical Method. Mind 11, No. 42. 203-19.

Singer, P. (2005). Ethics and Intuitions. Journal of Ethics 9, 331-52.

Sinnott-Armstrong, W. (2006). Moral Skepticism. Oxford: Oxford University Press.

Skarsaune, K. O. (2011). Darwin and Moral Realism: Survival of the Iffiest. Philosophy Studies, 152, 229-43.

Slote, M. (2010). Moral Sentimentalism. New York: Oxford University Press.

Smith, A. (1790/1991). The Theory of Moral Sentiments. In D.D. Raphael (ed.), British Moralists: Vol. II: Hume—Bentham and Index, (201-54). Indianapolis: Hackett Publishing Company.

Sober, E. (1994). From a Biological Point of View. USA: Cambridge University Press.

Sober, E. and Wilson, D.S. (1998). Unto Others: The Evolution and Psychology of Unselfish Behavior. Cambridge, MA: Harvard University Press.

Sosa, E. (1991). Reliabilism and Intellectual Virtue. In E. Sosa (ed.), Knowledge in Perspective. Cambridge: Cambridge University Press

—— (1999). How to Defeat Opposition to Moore. Philosophical Perspectives, 13, 141-53.

—— (2002). Reliability and the A Priori. In John Hawthorne & Tamar Gendler (eds.), Conceivability and Possibility, (369-84). New York: Oxford University Press.

Spencer, H. (1851). Social Statics: or, The Conditions essential to Happiness specified, and the First of them Developed. London: John Chapman.

—— (1879). Data of Ethics. London: Williams and Norgate.

—— (1879/1892) The Principles of Ethics Vol. I. New York: D. Appleton and Company.

—— (1885). Essays Speculative and Practical. New York: J. Fitzgerlad.

—— (1898). The Principles of Ethics Vol. II. New York: D. Appleton and Company.

Sripada, C. S. and Stich, S. (2007). A Framework for Psychology Norms. In P. Carruthers, S. Laurence, and S. Stich (eds.), The Innate Mind: Vol 2: Culture and Cognition, (280-301). New York: Oxford University Press.

Stewart, D. (1863). Account of the Life and Writings of Adam Smith LL.D. In Sir W. Hamilton (ed.), The Collected Works of Dugald Stewart, Vol. 10, (1-203). Edinburgh: Thomas Constable and Co.

Stich, S. (1993). Moral Philosophy and Mental Representation. In M. Hechter, L. Nadel, and R.E. Michod (eds.), The Origin of Values, 215-28. New York: Aldine de Gruyter.

Street, S. (2006). Darwinian Dilemma For Realist Theories of Value. Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition, Vol. 127, No. 1, Selected Papers from the American Philosophical Association, Pacific Division, 2005 Meeting, 109-66.

—— (2009). Evolution and the Normativity of Epistemic Reasons. Canadian Journal of Philosophy, Vol. 9, Supplement 1, 213-48.

—— (2015). Does Anything Really Matter or Did We Just Evolve to Think So? In G. Rosen, A. Byrne, J. Cohen, and S.V. Shiffron (eds.), The Norton Introduction to Philosophy, (685-95). New York: W.W. Norton & Company.

Thompson, P. (1999). Evolutionary Ethics: Its Origins and Contemporary Face. Zygon, 34 (3), 472-84.

Trivers, R. (1971). The Evolution of Reciprocal Altruism. Quarterly Review of Biology, 46, 35-57.

Tivers, R. (1985). Social Evolution. CA: The Benjamin-Cummings Pub Company.

Vavova, K. (2014). Debunking Evolutionary Debunking. In R. Shafer-Landau (ed.), Oxford Studies in Metaethics, Vol. 9, (76-101). Oxford, UK: Oxford University Press.

Vogel, J. (2000). Reliabilism Leveled. The Journal of Philosophy, 97:11, 602-23

Walter, A. (2006). The Anti-naturalistic Fallacy: Evolutionary Moral Psychology and the Insistence of Brute Facts. Evolutionary Psychology, 4, 33-48.

Wielenberg, E. J. (2010). On the Evolutionary Debunking of Morality. Ethics, 120 (3), 441-64.

Williamson, T. (2000). Knowledge and its Limits. New York: Oxford University Press.

Wilkins, J.S. and Griffiths, P.E. (2013). Evolutionary Debunking Arguments in Three Domains: Fact, Value, and Religion. In J. M. G. Dawes (ed.), A New Science of Religion, (133-46). New York: Routledge.

Wilson, D. S., Dietrich, E. and Clark, A.B. (2003). On the Inappropriate Use of the Naturalistic Fallacy in Evolutionary Psychology. Biology and Philosophy, 18 (5), 669-81.

Wilson, E.O. (1975). Sociobiology: The New Synthesis. Cambridge, MA: Harvard University Press.

Wilson, E.O. and Ruse, M. (1986). Moral Philosophy as Applied Science. Philosophy, Vol. 61, No. 236, 173-92.

Woolcock, P. G. (1999). The Case Against Evolutionary Ethics. In M. Ruse and J. Maienschein (eds.), Biology and the Foundation of Ethics, (276-306). Cambridge University Press.

No Author Indicated. (1869). The Westminster Review. London: Turner and Co. July and October. Vol. XXXVI.