A Critique of Evolutionary Debunking Arguments in Moral Philosophy

Dominic Newnham

Department of Philosophy, School of Humanities

The University of Adelaide

Submitted for the requirements for the degree of Master of Philosophy

August 2019

TABLE OF CONTENTS

ABSTRACT

Evolutionary debunking arguments (EDAs) have received much attention in recent moral philosophy. These types of arguments draw upon speculative evolutionary premises in order to challenge various philosophical viewpoints and theories. In some cases, empirical evidence has also been used to supplement the more speculative evolutionary premises in debunking arguments.

This thesis examines three prominent EDAs, from Sharon Street, Richard Joyce and Joshua Greene. Street's debunking target is the metaethical position of moral realism, particularly non-naturalistic realism. Joyce's target is the epistemic justification of moral judgements in general, leading to his conclusion of moral scepticism. Greene targets deontological approaches to moral philosophy, while maintaining that consequentialist theories are unaffected by his debunking claims. The main similarity between these three EDAs is the notion that evolutionary theory can be used to 'explain away' certain views in moral philosophy, by providing a scientific explanation of moral views that does not need to assume their truth.

The discussed EDAs face some common problems as well as problems specific to each argument. One of the main ways of resisting these arguments is to focus on the human capacity for rational reflection; it will thus be argued that our complex mental capacities enable us to overcome possible evolutionary influences on our moral thinking. This applies not only to the basic level of moral intuitions, but also to the complex moral theories that philosophers develop.

It is ultimately concluded that none of the discussed EDAs are successful. However, the possibility of more viable EDAs being developed in the future is not ruled out, provided that they can avoid the criticisms presented against these arguments.

DECLARATION

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Dominic Newnham

Signature: Date: 2/8/19

iii

ACKNOWLEDGEMENTS

iv

Chapter 1 – Introduction to Evolutionary Debunking Arguments

1.1 Introduction

In moral philosophy, one topic that has attracted considerable attention in recent years is that of evolutionary debunking arguments, hereafter referred to as EDAs. These types of arguments draw upon speculative evolutionary premises in addition to standard philosophical premises, leading to conclusions that challenge various philosophical viewpoints and theories, such as moral realism or deontological moral philosophy (Greene 2008, 2014; Joyce 2006; Street 2006). The evolutionary premises in EDAs are speculative as they cannot be directly confirmed: for instance, it is not possible to test whether tendencies towards certain moral beliefs affected the reproductive success of our early human ancestors. Various authors have acknowledged the speculative nature of evolutionary claims about morality, with some using this point to criticise EDAs (Buller 2009; Copp 2008; Joyce 2014; Shafer-Landau 2017; Street 2006). However, debunkers may also appeal to varying amounts of empirical evidence in order to supplement their speculative evolutionary premises. Thus, in addition to evolutionary biology, some debunking arguments have also referred to more experimental fields such as neuroscience (Greene 2008, 2014).

Evolutionary and empirical premises contrast with 'traditional' philosophical premises in that the latter have not typically been based on scientific claims. Apart from the fact that scientific investigations of morality are a relatively recent phenomenon, this may also be due to a long-standing assumption that philosophy as a discipline is fundamentally distinct from scientific inquiry. Although the relation between these disciplines tends to be an overlooked issue in EDA discussions, it is more explicitly noted in some works that focus on issues of relevance to certain EDAs (Audi 2014; Copp 1990; FitzPatrick 2016, 2017; Hartman 1963; Majors 2003; Sayre-McCord 1988). Debunkers' attempts to resolve persistent philosophical debates by appealing to science may be partly motivated by the view that science has been much more successful than philosophy. It is relatively easy to identify scientific success with the development of increasingly sophisticated technologies throughout human history, so even the more theoretical fields such as evolutionary biology may appear to be vindicated by their use of the same general

scientific methods. In contrast, the standing of philosophy is not helped by the fact that many philosophical debates have continued for centuries or longer. As such, EDAs are intended to reach significant philosophical conclusions by drawing upon evolutionary premises that have the apparent advantage of scientific credibility. However, it will become clear throughout the thesis that evolutionary debunking arguments do not necessarily perform any better than traditional philosophical arguments that lack any scientific claims.

After providing a general background to the concept of evolutionary debunking in this chapter, the following three chapters critically examine three prominent EDAs: the arguments are from Sharon Street, Richard Joyce and Joshua Greene respectively. Street's debunking target is the metaethical position of moral realism, particularly non-naturalistic versions of realism. Joyce aims to undermine the epistemic justification of moral beliefs in general, leading to a conclusion of moral scepticism. Greene aims to debunk deontological moral theories, while maintaining that consequentialist theories are unaffected. One of the main similarities between these EDAs is the underlying notion that evolutionary theory can 'explain away' certain views in moral philosophy, by providing a scientifically plausible explanation of moral views that does not need to assume their truth. For example, an explanation of why humans typically perceive certain actions as objectively morally wrong may simply refer to the evolutionary advantages of our ancestors' social cooperation being facilitated by shared moral beliefs. On the issue of evolutionary explanations, some authors have discussed the capacity of science to explain human morality in general, not just in relation to specific debunking arguments (Bruni, Mameli & Rini 2014; FitzPatrick 2016; Hales 2007; Mason 2010; Mogensen 2016; Singer 2005). The more specific applications of evolutionary explanations will become apparent in the chapters focused on Street, Joyce and Greene.

Ultimately, all three of the debunking arguments covered in the following chapters will be rejected due to a wide variety of problems with them; some of the problems are common to all three, while other issues mainly apply to individual arguments. A positive account of the capacity for moral knowledge will also be sketched in response to these EDAs, particularly Street's argument: the objective is to develop an account of the human capacity for rational reflection. The rational reflection account will be used to argue that our advanced reasoning capacities enable us to overcome possible evolutionary influences on our moral thinking; this applies not only to simple moral intuitions, but also to the complex moral theories that philosophers develop. This is just one of the various ways in which the discussed debunking arguments will be resisted. After rejecting the three

prominent EDAs in Chapters 2 to 4, the final chapter briefly considers some more general issues that are relevant to these types of arguments. This includes questions about the relevance of specifically evolutionary premises in EDAs, as well as the issue of whether scientific and philosophical accounts of morality necessarily compete with each other. The purpose of Chapter 5 is primarily to raise awareness of these underlying issues that affect the standing of EDAs in general. Identifying the general issues that need to be resolved is important in order for EDAs to have any chance of progressing from the unsuccessful ones presented by Street, Joyce and Greene. Ultimately, this thesis does not attempt to rule out the possibility of more successful EDAs being developed in the future, but first they need to overcome the problems facing the current arguments.

With the overall direction of the thesis now summarised, the remainder of this chapter examines the general concept of evolutionary debunking, including a brief overview of how some earlier works on evolution and ethics have led towards the more recent and widely discussed EDAs.

Compared to other philosophical arguments, the main distinctive feature of EDAs is that they provide an evolutionary explanation of some aspect of moral philosophy—such as a particular moral theory—that does not seem conducive to that viewpoint being true. Thus, one of the common themes in debunking arguments is the notion that an evolutionary explanation of human morality does not need to posit any moral truths. This has been noted by many authors in the EDA literature (Behrends 2013; Carruthers & James 2008; Das 2016; Deem 2016; Enoch 2010; Hopster 2018; Kahane 2011; Mogensen 2016; Vavova 2015; Wielenberg 2010). Debunkers then use their evolutionary explanations to support a negative conclusion regarding their targeted moral view. Debunking arguments can be applied to metaethical views such as moral realism, or to various types of normative theories, such as deontological or consequentialist theories (Kelly 2017). Since all of these philosophical viewpoints are essentially sets of particular beliefs about morality, a debunking explanation of a targeted view must be able to account for the origins of the relevant beliefs. Evolutionary explanations appear to be well-suited to this purpose, as their focus is much further back in time as compared to explanations based on more immediate influences on moral views, such as personal biases and sociological factors. In this way, evolutionary explanations could be considered more fundamental than other explanations, including philosophers' own stated reasons for favouring certain moral views. This notion can be observed particularly in Joshua Greene's debunking argument that is covered in Chapter 4, as he argues that deontologists merely rationalise their intuitive moral

judgements that can be attributed to evolutionary factors (Greene 2008). More broadly, the idea that so many aspects of human thought and behaviour can be traced back to evolutionary factors seems to be part of the appeal of EDAs, at least to those who present and support them.

Although debunkers supposedly have the plausibility of evolutionary theory on their side, they face resistance from the majority of philosophers who are doubtful of the debunking capacity of these evolutionary accounts. In any discipline, resistance to unfamiliar new approaches is to be expected. Philosophy has traditionally been a discipline in which abstract theories and arguments are developed independently of the considerations that feature in EDAs: namely the evolutionary origins of our moral views. Despite being outnumbered by the proponents of traditional philosophical arguments that do not engage with evolutionary theory, some debunkers have argued that evolutionary accounts of morality can shift the burden of proof over to defenders of more conventional views. For instance, Richard Joyce argues that the epistemic justification of moral beliefs is undermined by his evolutionary genealogy of morality, thus placing the burden of proof on those who wish to reinstate the justification of any moral beliefs (Joyce 2006, 2016). However, since this thesis argues that certain philosophical views are not threatened by the EDAs from Street, Joyce and Greene, it will be maintained that debunkers are the ones who owe us better arguments.

If the overall argument of the thesis is convincing, it will support a fairly modest outcome: the philosophical positions targeted by debunkers—such as moral realism—will return to whatever standing they have in the absence of evolutionary challenges. As such, until any better EDAs are developed in the future, debunkers' targeted views should continue to be assessed from a non-evolutionary perspective. It is beyond the scope of the thesis to consider how well these targeted philosophical viewpoints fare against non-evolutionary challenges. Since it is being suggested that the rejection of current EDAs should lead us back to non-evolutionary arguments, it is worth examining how EDAs rose to prominence in the first place.

1.2 The Rise of Evolutionary Debunking Arguments

Evolutionary debunking arguments have become increasingly relevant in recent years, with the amount of literature on the topic increasing significantly since 2006. This was the year in which Sharon Street's "A Darwinian Dilemma for Realist Theories of Value" (Street

2006) and Richard Joyce's *The Evolution of Morality* (Joyce 2006) were published, both instigating a great deal of debate. Since then, many philosophers have become more concerned with the origins and causes of moral beliefs and theories, particularly the ways in which these origins may affect their plausibility. Although the surge of interest in EDAs is relatively recent, modern debunking arguments have surely been influenced to some degree by earlier attempts at deriving philosophical conclusions from evolutionary premises. One notable early example of the debunking approach is a 1986 article by Michael Ruse and Edward O. Wilson, who expressed dissatisfaction with the traditional approach to moral philosophy:

> For much of this century, moral philosophy has been constrained by the supposed absolute gap between *is* and *ought*, and the consequent belief that the facts of life cannot of themselves yield an ethical blueprint for future action. For this reason, ethics has sustained an eerie existence largely apart from science. Its most respected interpreters still believe that reasoning about right and wrong can be successful without a knowledge of the brain, the human organ where all the decisions about right and wrong are made. Ethical premises are typically treated in the manner of mathematical propositions: directives supposedly independent of human evolution, with a claim to ideal, eternal truth (Ruse & Wilson 1986, p. 173).

As such, Ruse and Wilson intended to challenge the view of ethics being a mysterious phenomenon that cannot be explained scientifically. Although it is now more common to grant that human morality can be scientifically explained, this does not mean that there is now widespread acceptance of Ruse and Wilson's views from the time. One particularly contentious aspect of their early work is their suggestion that moral philosophy should become a branch of science, which is a more radical approach than that of recent debunking arguments: "While many substantial gains have been made in our understanding of the nature of moral thought and action, insufficient use has been made of knowledge of the brain and its evolution. … The time has come to turn moral philosophy into an applied science" (Ruse & Wilson 1986, p. 173). Although moral philosophy is not considered a scientific discipline by recent debunkers, knowledge of the human brain and our evolved mental capacities is used in modern EDAs. Among the EDAs examined in this thesis, Sharon Street and Richard Joyce are primarily concerned with evolutionary speculations about our mental capacities, whereas Joshua Greene combines evolutionary claims with empirical findings from neuroscientific studies of brain activity.

Returning to the early debunking approach exemplified by Ruse and Wilson, the applications of their view can be demonstrated by considering a specific example from

their work. One particular moral issue that they mention is sibling incest: due to its potential negative effects on reproductive fitness, the ubiquitous moral opposition to it is amenable to an evolutionary explanation: "Formal incest taboos are the cultural reinforcement of the automatic inhibition, an example of the way culture is shaped by biology" (Ruse & Wilson 1986, p. 184). This point raised the idea that examining morality from a biological rather than purely philosophical perspective could have implications for the moral status of particular issues. However, apart from this brief consideration of one specific moral issue, Ruse and Wilson's view that moral philosophy can be directly informed by evolutionary biology is not adequately supported by the rest of their discussion. Much of their discussion is purely descriptive, lacking in examples of normative conclusions that could be derived from biological facts. Merely knowing about the biological basis of incest avoidance does not tell us whether it is right or wrong, unless we include some normative assumptions. For example, the assumption could be that we all ought to adopt 'natural' human behaviour, but this would require a consensus on how to define natural versus unnatural behaviour. Peter Singer has highlighted the important role of such assumptions; this can be seen in his book *The Expanding Circle* which was originally published in 1981, before Ruse and Wilson's article: "Where an ethical belief is explicitly based on an assumption about what is natural for human beings, there is no difficulty in seeing how biology can be a tool of criticism" (Singer 2011, p. 68). Singer's point does not necessarily help Ruse and Wilson's case, as it is not clear how biological facts could help us identify what is 'natural' behaviour for humans. This is just one example of the difficulties facing any attempt to derive moral conclusions from biological facts.

Compared to Ruse and Wilson's bold claims about moral philosophy and science, Singer's view from the same period was closer to that of modern EDAs: "Neither evolutionary theory, nor biology, nor science as a whole, can provide the ultimate premises of ethics. Biological explanations of ethics can only perform the negative role of making us think again about moral intuitions which we take to be self-evident" (2011, p. 84). The general concept of a 'negative role' is evident in the recent EDAs from Street, Joyce and Greene: these debunkers combine biological explanations of morality with independent philosophical premises in order to reach debunking conclusions. Street describes widespread moral belief tendencies as evolutionary adaptations, combining this speculation with the premise that such adaptations would be highly unlikely to align with moral facts posited by moral realists (Street 2006). Joyce develops the idea of humans having evolved an adaptive 'moral sense' that explains the concepts in our moral

judgements; he argues that we should be sceptical of moral judgements since they can be explained purely in terms of this adaptation, without needing to assume their truth (Joyce 2006). Greene describes deontological moral philosophy as a rationalisation of adaptive emotion-based responses to certain factors in moral dilemmas; this is combined with assumptions about the moral irrelevance of those factors, thus forming part of his argument against deontology (Greene 2008, 2014). Each of these EDAs utilises evolutionary claims in the negative role of 'explaining away' certain philosophical targets, such as moral realism or deontological philosophy. Although this approach avoids the problems with Ruse and Wilson's attempts at directly drawing ethical conclusions from evolutionary biology, it will become clear in the following chapters that evolutionary premises have a limited impact in the discussed EDAs. As such, it will turn out that even the more modest role for evolutionary explanations provides little advantage to evolutionary arguments over more traditional philosophical arguments that do not appeal to science.

As indicated by the idea of evolutionary explanations only playing a negative role, Singer's early views on biological explanations in moral philosophy can be seen as a pathway from Ruse and Wilson's ambitious views of the time to the somewhat more modest approach of recent EDAs. In discussing the evolutionary origins of the human conception of morality, Singer noted that despite the complexity and diversity of our moral views, they can nevertheless be traced back to relatively simple patterns of social behaviour seen in other animals:

> [W]hile the diversity of ethics is indisputable, there are common elements underlying this diversity. Moreover, some of these common elements are so closely parallel to the forms of altruism observable in other social animals that they render implausible attempts to deny that human ethics has its origin in evolved patterns of behavior among social animals (Singer 2011, p. 29).

Now that it is widely accepted that human morality can be explained in evolutionary terms, modern philosophical discussions concerning evolution and ethics are mainly focused on the implications this may have for the plausibility of various moral theories and the reliability of moral intuitions. Singer's observation regarding the common elements of morality also provides an early example of a point that features in some recent EDAs, such as Street's argument. Street specifically refers to the ubiquity of certain moral belief tendencies, such as the view that it is right to care for one's own offspring more than we care for strangers (Street 2006, p. 115). By pointing out how widespread certain moral beliefs are, she appeals to the idea that this can be explained in evolutionary terms. However, while critics of EDAs often grant for the sake of argument that debunkers'

evolutionary claims may be plausible, it will be demonstrated in later chapters that debunkers' reliance on speculation is problematic.

Singer in his early work also discussed the problems with philosophers relying on ungrounded claims, although he was more aligned with debunkers in that he criticised traditional philosophy for its reliance on intuitions. Thus, he observed that many philosophers have been content with ethics merely systematising pre-existing moral intuitions:

> Almost all the thinking we do about ethics involves connecting one ethical judgment to another, more fundamental one. Even moral philosophers who develop theories about what we ought to do rarely press deeper. Some of them explicitly say that philosophy can do no more than systematize our moral intuitions. We can criticize one moral intuition on the basis of others, they say, but we cannot criticize all or most of our moral intuitions at once (Singer 2011, p. 70).

As such, Singer suggested that science could facilitate progress in moral philosophy by allowing us to discover the underlying causes and origins of our moral intuitions. In his words, "Science provides leverage against some ethical principles when it helps us understand why we hold our ethical principles. What we take as an untouchable moral intuition may be no more than a relic of our evolutionary history" (Singer 2011, p. 70). Thus, Singer assumed that certain causal explanations of intuitions can have a debunking effect. Although recent debunkers maintain that evolutionary explanations can have such an effect, defenders of moral beliefs and theories are generally unconcerned by these types of purported explanations of their views. Since the thesis rejects the three main EDAs that are discussed, the non-debunking view of evolutionary explanations will be favoured. Thus, although the explanations offered by debunkers will be criticised in various ways, the general concept of evolutionary accounts of morality will not be rejected; rather, the aim is to oppose the idea that such explanations can debunk the targeted philosophical views.

This chapter has briefly examined early examples of debunking ideas based on evolutionary considerations, in order to introduce the general approach of EDAs. The more recent debunking arguments that are examined in the following chapters have built upon some of these earlier ideas, although it will be demonstrated that they are still far from being plausible. Chapter 2 will now begin the main focus of assessing modern EDAs by examining Street's argument.

<u>Chapter 2 – Sharon Street's Evolutionary Debunking Argument</u>

2.1 Overview of Street's Argument

Sharon Street's influential EDA against moral realism is one of the most prominent debunking arguments in the literature, having received much attention from other philosophers (Braddock 2016; Brosnan 2011; Clarke-Doane 2012; Copp 2008; Das 2016; Deem 2016; Enoch 2010; FitzPatrick 2014, 2015; Hanson 2017; Hopster 2018; Kahane 2011; Klenk 2017; Mogensen 2016; Shafer-Landau 2013; Skarsaune 2011; Street 2006; Tropman 2014; Vavova 2015; White 2010; Wielenberg 2010). This chapter examines several major problems with the argument, which may also apply to any other EDAs that depend on similar ideas. The current section begins by outlining the main points in Street's argument, with later sections focusing on the issues that are raised.

Street's EDA aims to undermine evaluative realism by drawing upon evolutionary biology. Evaluative realism is broader than moral realism: it is the view that evaluative beliefs can be true or false in virtue of their relation to mind-independent evaluative facts. As Street describes this position, "There are at least some evaluative facts or truths that hold independently of all our evaluative attitudes" (2006, p. 110). Moral facts are only one type of evaluative fact: non-moral evaluative facts may concern reasons in favour of certain actions that have no apparent moral relevance. However, the discussion from this point onwards will proceed as though Street is directly targeting moral realism, with any of her references to evaluative facts or beliefs being referred to as moral facts or beliefs. Many philosophers have taken this approach when discussing Street's argument, including David Copp (2008).

Street's conception of moral beliefs includes the approval or disapproval of certain actions and states of affairs, and judgements about what we have a reason to do, or 'ought' to do (Street 2006, p. 110). Street also notes that there are both non-natural and naturalistic versions of moral realism. It will be argued in this chapter that naturalistic moral realism is safer from Street's argument. For now, it is important to note that her argument directly targets non-natural moral realism; Street recognises that it is a more complex issue whether naturalistic realism is also undermined by her argument (Street 2006, p. 112). She describes her targeted position of non-natural realism as the following view: "[E]valuative facts or truths are not reducible to any kind of natural fact, and are not the kinds of things

that play a role in causal explanations; instead, they are irreducibly normative facts or truths" (Street 2006, pp. 111-112). In contrast, moral naturalists assert that morality can be accounted for in a completely natural conception of the world, which rules out the existence of any non-natural or supernatural entities.

Since Street's argument draws upon speculative evolutionary biology, she begins with the following statement:

> I try to rest my arguments on the least controversial, most well-founded evolutionary speculations possible. … [A] cognitive trait such as the widespread human tendency to value the survival of one's offspring may, according to evolutionary psychology, be just as susceptible to evolutionary explanation as physical traits such as our bipedalism or our having opposable thumbs (Street 2006, pp. 112-113).

Although she is aware of her reliance on speculation, Street maintains that there is enough certainty for her argument to be taken as a serious threat to realism: "[W]hile I am skeptical of the *details* of the evolutionary picture I offer, I think its *outlines* are certain enough to make it well worth exploring the philosophical implications" (Street 2006, p. 113). This element of speculation is criticised later on, but for now it is important to keep this in mind when considering her evolutionary claims. Street's evolutionary account begins with the following speculations about the evolutionary origins of our moral beliefs:

> The forces of natural selection have had a tremendous influence on the content of human evaluative judgements. This is by no means to deny that all kinds of other forces have also shaped the content of our evaluative judgements. No doubt there have been numerous other influences: some of them were perhaps evolutionary factors other than natural selection — for example, genetic drift; and many other forces were not evolutionary at all, but rather social, cultural, historical, or of some other kind. And then there is the crucial and *sui generis* influence of rational reflection that must also be taken into account … I am discounting none of these other influences. My claim is simply that one enormous factor in shaping the content of human values has been the forces of natural selection, such that our system of evaluative judgements is thoroughly saturated with evolutionary influence (Street 2006, pp. 113-114).

Apart from the bold speculation that our judgements are 'thoroughly saturated' with evolutionary influences, it is also notable that Street acknowledges the influence of other factors such as rational reflection. As will become clear later, this point is particularly important to focus on when opposing her argument.

Street's evolutionary account quickly moves on to the details of how evolutionary influences have ultimately influenced our current moral judgements. The first point is that the survival and reproductive success of early humans would have been affected by their tendencies towards different types of moral beliefs, with some tendencies being much more likely to get passed down than others. For example, it would be harmful to think that endangering one's own life is good, whereas it would be beneficial to think that one's own survival is good and therefore worth promoting (Street 2006, pp. 114-115). Assuming that tendencies towards different moral beliefs affected the behaviour of early humans, we would expect different rates of reproductive success between individuals, depending on their particular belief tendencies. This allows for an evolutionary explanation of the prevalence of certain belief tendencies in the present day, as the process of natural selection would favour humans whose tendencies were more conducive to survival and reproduction. Thus, at least some of our beliefs and values seem to be explainable by this evolutionary account.

Apart from straightforward beliefs about value, beliefs concerning moral 'rightness' and 'wrongness' could also have evolutionary explanations; for instance, social cooperation requires some level of agreement on right and wrong behaviour. It is fairly self-explanatory that social cooperation would benefit individuals living in a hostile environment, although there is debate around the biological issue of individual versus group selection. Street does not examine these biological details since they are not particularly relevant to her objections against realism. Ultimately, any plausible account of natural selection can be used as long as it is assumed that certain belief tendencies were more conducive to survival and reproduction.

Presumably, early humans would not have held moral beliefs of the same level of complexity as those we have now. As such, Street speculates that natural selection operated on "proto" versions of the more complex moral beliefs that we currently have (Street 2006, p. 114). This claim is important since it provides a link between our evolutionary history and contemporary morality, including the complex reasoning involved in moral philosophy. The idea is that many aspects of contemporary morality can be traced back to basic proto-moral beliefs that served the simpler needs of early humans. Street also asserts that despite the diversity of moral judgements in the present day, they exhibit patterns that are constant throughout history and across different cultures. She lists several examples of such beliefs:

(1) The fact that something would promote one's survival is a reason in favor of it.

(2) The fact that something would promote the interests of a family member is a reason to do it.

(3) We have greater obligations to help our own children than we do to help complete strangers.

(4) The fact that someone has treated one well is a reason to treat that person well in return.

(5) The fact that someone is altruistic is a reason to admire, praise, and reward him or her.

(6) The fact that someone has done one deliberate harm is a reason to shun that person or seek his or her punishment (Street 2006, p. 115).

These examples are supposed to support the idea that our diverse moral judgements can be reduced to a straightforward evolutionary explanation, since natural selection likely would have favoured humans with these belief tendencies. To illustrate how such tendencies could be favoured, Street provides an explanation of example (1): people who tend towards this belief would do more to promote their survival than those who lack this belief, thus allowing them to pass down this tendency to their offspring. This would gradually increase the proportion of individuals with this belief tendency over each subsequent generation (Street 2006, p. 116).

Street clarifies that moral beliefs cannot themselves be genetically inherited, at least not in a complex form with specific content. Rather, it is the more basic evaluative tendencies that she claims to be genetically heritable traits (Street 2006, p. 119). Street defines such a basic tendency in these terms: "[A]n unreflective, non-linguistic, motivational tendency to experience something as 'called for' or 'demanded' in itself, or to experience one thing as 'calling for' or 'counting in favor of' something else. We may think of these as 'proto' forms of evaluative judgement" (Street 2006, p. 119). She speculates that variations in these basic tendencies may have been based in genetic differences throughout most of our evolutionary history. This is a scientific claim that cannot be easily validated. However, the finer details of evolutionary biology are not as important as the way Street uses these speculations to support her argument: the claim is that natural selection has directly influenced our basic evaluative tendencies, which then significantly affect the content of our more complex moral beliefs. Thus, the influence of natural selection on the specific content of moral beliefs has been indirect (Street 2006, pp. 119-120).

The indirect nature of evolutionary influences on our moral beliefs would seem to allow plenty of room for other influences, most notably the human capacity for rational

reflection. This capacity has clearly been developed to a significant degree throughout human evolution. However, despite the extensive development of human mental capacities compared to the basic reasoning our early ancestors would have required, Street downplays the evolutionary autonomy of rational reflection: "[H]ad the general content of our basic evaluative tendencies been very different, then the general content of our full-fledged evaluative judgements would also have been very different, and in loosely corresponding ways" (Street 2006, p. 120). Thus, despite our capacity for 'full-fledged' reflective beliefs that are more sophisticated than our basic evolved belief tendencies, it is claimed that they are nevertheless contingent on evolutionary factors. As such, our capacity to critically reflect upon our moral beliefs is just another mental capacity that is apparently saturated with evolutionary influence (Street 2006, p. 124). This would suggest that the influence of evolutionary forces such as natural selection is inescapable, even affecting the complex moral theories that philosophers develop. Importantly, this raises the question of whether any domains of knowledge are relatively unaffected by evolutionary influences, such as science and mathematics. Comparing moral knowledge to scientific knowledge turns out to be an important way in which moral realists can resist Street's EDA, as will be argued later in this chapter.

The discussion so far has covered the evolutionary premises of Street's argument. The next part of her EDA is the crucial point where she presents a 'Darwinian dilemma' for moral realists:

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

Street considers both of these options to be untenable, thus encouraging us to drop moral realism in favour of anti-realism. On one horn of the dilemma, the realist denies any relation between evolutionary forces and independent moral truths. According to moral realism, moral facts are mind-independent in the sense that they are not determined by the views of any actual or possible believers. Thus, the facts would remain the same whether or not anyone correctly recognised them. Street asserts that if there is no relation between evolutionary influences and mind-independent moral facts, then evolution must be considered a distorting influence on our moral beliefs (2006, p. 121). In this case, the only way realists could deny a 'distorting' influence on this horn of the dilemma would be to assert that our beliefs reliably track moral facts by pure coincidence. However, it seems

incredibly unlikely that such a coincidence would occur for any substantial portion of our moral beliefs (Street 2006, pp. 121-122). Thus, denying any relation leads realists to an unacceptable epistemic position, as depending on pure chance is not a sufficiently reliable method for recognising mind-independent moral facts. To reinforce the idea that a great deal of moral 'luck' would be required, Street asserts that the range of moral beliefs that humans actually have is only a tiny selection from a huge universe of 'logically possible' beliefs. In other words, there are countless other coherent sets of moral beliefs that humans could have ended up with, had our evolutionary history been different (Street 2006, p. 122).

On the other horn of the dilemma, the realist asserts that there is a relation between evolutionary influences and independent moral facts. Street considers this option to have more initial plausibility than denying a relation, since it does not require a purely coincidental alignment of our evolved evaluative tendencies with moral truths (Street 2006, p. 125). Street suggests that a realist taking this option would posit a moral truth-tracking mental capacity. The realist might argue that just as truth-tracking perceptual beliefs promoted our ancestors' survival by allowing them to detect predators and food, moral beliefs could have benefited them by reliably tracking moral facts, which would then guide their social behaviour (Street 2006, pp. 125-126). This will be called the 'truth-tracking' account. Street notes that since this is a hypothesis about evolution, it must be evaluated scientifically rather than philosophically.

Unfortunately for realists, a moral truth-tracking account appears to have much less scientific plausibility than an 'adaptive link account', as Street describes it. According to this latter type of explanation, tendencies towards certain types of moral beliefs were naturally selected due to their advantages in promoting survival and reproduction. It is called the 'adaptive link' account because the behaviours linked to certain evaluative tendencies are comparable to other biological mechanisms that allow organisms to adapt to their environment and living circumstances. Street gives the example of automatic reflex responses to danger, such as withdrawing one's hand from a hot surface (Street 2006, p. 127). In these non-moral cases involving perceptual beliefs, it is clear that tracking mind-independent natural facts about the world is adaptive; we can detect real danger such as a hot surface, and we respond appropriately. Although mental states are generally more complex than somatic adaptive link mechanisms, Street is nevertheless suggesting that mental adaptations could ultimately serve the same biological purpose. Thus, tendencies towards certain types of moral beliefs may have encouraged early humans to favour actions that promoted survival and reproductive success (Street 2006, p. 127). Importantly,

this adaptive link account does not posit the existence of any independent moral facts. Having described both the moral truth-tracking account and the adaptive link account, Street asserts that the latter is clearly superior by all the usual standards of scientific plausibility. She lists three points in support of this claim: the adaptive link account is more parsimonious, it is clearer, and it has more explanatory power (Street 2006, p. 129). Each of these points will now be examined.

The first point concerns the principle of parsimony, which is the notion that it is more plausible to propose simpler mechanisms in a theory or explanation. Whereas the adaptive link account simply states that certain moral belief tendencies were selected due to their greater promotion of survival and reproductive success, the truth-tracking account states that these tendencies were selected because humans recognised independent moral truths that promoted these same advantages. Thus, the tracking account posits an extra entity—independent moral facts—which makes it less parsimonious.

As for Street's second point that the adaptive link account is clearer, the idea is that it is obscure to conceive of the independent truth of certain moral beliefs having any relevance to evolutionary forces such as natural selection. Compared to moral beliefs, it is much clearer how the mind-independent truth of perceptual beliefs would be relevant to natural selection; for example, it is clearly beneficial to accurately perceive a predator in one's field of vision. The realist is said to be burdened with explaining how the ability to track independent moral truths could have affected our ancestors' chances of survival and reproduction at all (Street 2006, pp. 129-130). However, there are countless widely accepted facts about the natural world that would not have been knowable to early humans, thus making such facts irrelevant to natural selection. Street uses some non-moral facts as an example, namely facts about the existence of low-frequency electromagnetic wavelengths. A capacity to detect such obscure features of the world could even be disadvantageous, since it would use more biological resources without providing any survival benefits (Street 2006, p. 130). Since we have scientific reasons to believe that these sorts of mind-independent properties exist, this raises the question of why moral realists cannot claim that we also have reasons to believe there are mind-independent moral facts. The answer is that Street's denial of this possibility relies on the fact that her argument mainly targets non-natural versions of realism.

Non-naturalist realists cannot easily compare a purported capacity to detect non-natural moral facts with our capacity to discover obscure natural facts through scientific reasoning. Street recognises that naturalistic moral realism may initially appear to be safer from her debunking argument in this regard (Street 2006, p. 131). After all, naturalists

identify moral facts with certain natural facts, which are the kinds of facts that have a causal influence on the world; this makes it possible for us to identify them. However, Street uses the parsimony point again when she opposes the naturalist's potential response to her argument. She demonstrates this point by referring to the ubiquitous moral belief that we are obliged to care for our offspring. On one hand, naturalistic moral realists would claim that we have this obligation as a matter of mind-independent moral fact, such that our belief in this obligation is both independently true and also would have been favoured by natural selection. On the other hand, Street maintains that this posited moral fact is an unnecessary extra entity, as the simpler explanation is that this belief tendency was selected purely because it promoted behaviour that was beneficial to survival and reproduction (Street 2006, pp. 131-132).

Street's third reason for favouring the adaptive link account over the realist's truth-tracking account is that the former has more explanatory power: it explains why there are widespread tendencies towards certain moral beliefs rather than others. These include the six examples listed earlier, such as the belief that our survival is important and worth promoting, that we have greater obligations to our own children than to strangers or distant relatives, and so on. The adaptive link account simply explains the ubiquity of these beliefs as resulting from their promotion of behaviours that were conducive to our ancestors' survival and reproduction. In contrast, the truth-tracking account seems to require an implausible coincidence:

> [H]ow does the tracking account explain the remarkable coincidence that so many of the truths it posits turn out to be exactly the same judgements that forge adaptive links between circumstance and response — the very same judgements we would expect to see if our judgements had been selected on those grounds alone, regardless of their truth? (Street 2006, p. 132)

Thus, Street maintains that the greater scientific plausibility of the adaptive link account makes any truth-tracking account unnecessary, assuming that the realist cannot provide enough examples of plausible non-adaptive beliefs. However, even if realists attempted to provide such examples, Street would likely refer back to one of her claims about the earlier six examples of widespread adaptive beliefs:

> There is, of course, a seemingly unlimited diversity to the evaluative judgements that human beings affirm. Yet even as we note this diversity, we also see deep and striking patterns, across both time and cultures, in many of the most basic evaluative judgements that human beings tend to make (Street 2006, p. 115).

Continuing with her point about explanatory power, Street also suggests that the tracking account is worse off when attempting to explain the prevalence of moral beliefs that most of us do not endorse after critically reflecting on them. For example, we may consider the widespread default tendency to believe that people from 'out-groups' do not deserve to be treated as well as those in our own communities. This belief is not as widely supported now that many societies have collectively reflected upon this tendency. Whereas the adaptive link account would explain the prevalence of this initial tendency in terms of its benefits to our ancestors' reproductive success, it may seem that the tracking account has to maintain that the belief is true.

The issue of moral realism and explanatory power has been examined by philosophers in a general sense as well as in the context of EDAs. For instance, as Geoffrey Sayre-McCord observes: "As many would have it, we have positive reason to believe something only if supposing it true contributes in some way to explaining our experiences" (Sayre-McCord 2007, p. 10). Realists could respond by maintaining that moral facts do play an explanatory role with regard to forming our moral beliefs and justifying moral claims. This option is available to moral naturalists as well as non-naturalists. However, it seems clear that naturalists are in a better position: they can identify moral facts with certain natural facts that are involved in causal explanations of our moral beliefs. For example, our perception of natural facts about pleasure and pain can influence our beliefs about the moral status of these facts. Alternatively, realists could deny that moral facts necessarily have an explanatory role with regard to moral beliefs. Sayre-McCord raises the possibility that the role of moral facts might be to justify moral beliefs, not to explain them (Sayre-McCord 2007, pp. 10-11). This option of denying an explanatory role will not be considered, as the main focus will be on the naturalistic realist's other option of maintaining an explanatory role for natural moral facts. This latter option is examined in Section 2.2, which develops the idea that our evolved capacity for rational reflection is what allows us to identify the moral status of certain natural facts.

As for the current issue of the seemingly superior explanation provided by the adaptive link account, there is another point that Street cites in support of it. She claims that there are countless logically possible moral beliefs that humans could have, yet we only endorse a very limited selection of these possible beliefs (Street 2006, p. 133). The adaptive link account can explain this by stating that other logically possible beliefs would be useless or maladaptive in an evolutionary sense, so a tendency towards such beliefs would not be favoured by natural selection. Street lists some examples of these hypothetical moral judgements: "[F]rom the judgement that infanticide is laudable, to the

judgement that plants are more valuable than human beings, to the judgement that the fact that something is purple is a reason to scream at it" (Street 2006, p. 133). On the other hand, the truth-tracking account can only reassure us that such judgements are false, without offering any explanation of why we think so. As with Street's previous point, it will be argued in Section 2.2 that moral realists can respond by defending the explanatory power of moral facts, which is more viable for naturalists than non-naturalists.

The aforementioned points have focused on the truth-tracking account that realists are expected to endorse if they take the horn of the dilemma in which they affirm a relation between evolutionary influences and independent moral facts. In the next section it will be argued that there is no need for evolutionary forces to have directly granted us a truth-tracking capacity in other domains of knowledge, such as science. Rather, it will turn out that an indirect relation between evolutionary influences and independent facts is sufficient. Importantly, this notion also extends to knowledge of independent moral facts, at least for naturalistic versions of moral realism. The comparison between scientific and moral knowledge will demonstrate that the realist's truth-tracking account is defensible, which undermines Street's supposed dilemma that is central to her argument.

2.2 Evolutionary Influences and Autonomous Rationality

When considering how Street applies scientific methodology to moral philosophy, it is notable that her argument is primarily an epistemological challenge to realist views. This may be because it is easier to challenge a moral truth-tracking capacity on scientific grounds, rather than trying to scientifically disprove the existence of moral facts themselves. The main way in which Street follows scientific practice is in suggesting that we should not assume the existence of something unless it is required for the best explanation of a phenomenon; this relates to her points about parsimony and clarity. However, it is debatable whether moral philosophy should be held to the same standards of explanation as science, even for naturalistic approaches to morality. This issue of explanatory standards across different disciplines is relevant to EDAs in general, since they appeal to the scientific field of evolutionary biology. As such, the main discussion of this issue is saved for Chapter 5, which follows the individual assessment of each prominent EDA.

Street's use of scientific reasoning is important in relation to her points about our evolved mental capacity for rational reflection. One of the central ideas in her argument is

that we cannot completely escape the evolutionary influences on this capacity that we use to rationalise or reject various moral beliefs. Rationality is obviously also used extensively in a vast array of other domains of knowledge, most notably in science. As such, Street's argument will be opposed on the grounds that it denies a plausible comparison between scientific and moral knowledge. Specifically, her argument denies that we can reliably track mind-independent moral facts even though we can discover many obscure and complex natural facts through scientific reasoning. This includes our scientific knowledge of electromagnetic wavelengths, to use the previously mentioned example.

Street's distinction between moral knowledge and other domains of knowledge is connected to the fact that her argument mainly targets non-natural versions of realism. It will be argued that this aspect of her argument is much less effective against naturalistic moral realism. Street is not sceptical of our most basic capacities to track certain natural facts, given that she supports a standard evolutionary explanation of reliable truth-tracking perceptual beliefs. This capacity was presumably selected since it allowed our ancestors to detect real features of the natural world that were relevant to their survival and reproductive success, such as the presence of predators (Street 2006, p. 130). Any heritable trait, including our perceptual capacities, can be susceptible to natural selection if it affects the chances of passing on genes to subsequent generations. Street asserts that our basic evaluative belief tendencies are heritable, which means that our current moral reasoning is ultimately the product of evolutionary influences (Street 2006, pp. 118-119). Unlike the case of perceptual beliefs, it is not clear how a capacity to track non-natural moral facts could be favoured by natural selection. To understand why, we may recall Street's definition of non-natural moral realism: "[E]valuative facts or truths are not reducible to any kind of natural fact, and are not the kinds of things that play a role in causal explanations; instead, they are irreducibly normative facts or truths" (Street 2006, pp. 111-112). The crucial point here is the causal isolation problem for non-natural realism. Although it is beyond the current scope to consider how non-naturalism might be defended in light of this problem, what matters for Street's argument is that this would rule out any truth-tracking relation between our mental capacities and non-natural moral facts. Non-naturalist realists are therefore left with the implausible option of postulating a purely coincidental alignment of our mental capacities with moral facts.

Even if we assume for the sake of argument that non-natural moral facts exist, Street's argument seems to suggest that any mental capacity with evolutionary origins would be unable to reliably track non-natural moral facts due to their causal isolation. Leaving aside the issue of whether even an evolutionarily autonomous mental capacity

could detect such facts, Street focuses on the apparent inescapability of evolutionary influences in order to rule out the possibility of such a capacity. Notably, this inescapability even extends to rational reflection: "[A]ll our reflection over the ages has really just been a process of assessing evaluative judgements that are mostly off the mark in terms of others that are mostly off the mark. And reflection of *this* kind isn't going to get one any closer to evaluative truth" (Street 2006, p. 124). Presumably, moral realists would not want to be limited to positing moral facts that happen to align with naturally selected belief tendencies; for one thing, many contemporary moral beliefs concern issues that are completely unrelated to the living circumstances of our distant ancestors. To allow for the existence of a wide variety of moral facts, including facts that are relevant to modern issues such as climate change and bioethics, realists must oppose Street's claim about the inescapable distorting evolutionary influences on rational reflection.

One way that realists could argue for the autonomy of our mental capacities is by comparing the use of rational thinking in moral philosophy with its use in other fields such as science. William FitzPatrick has defended a position along these lines (2015, 2016). He suggests that Street's debunking argument is invalid, and that in order to become valid it would need the following additional premise:

> If natural selection is responsible for our having certain basic cognitive capacities, which we employ in some domain of thought, then the only way for such exercises of those capacities to be non-accidentally and reliably truth-tracking in that domain would be for natural selection to have made things that way (FitzPatrick 2015, pp. 885-886).

Although this premise would help Street's case against realism, it is unclear how this would be compatible with complex scientific knowledge, including the field of evolutionary biology that Street herself appeals to. Opposing Street's view, FitzPatrick argues that we clearly have transcended the evolutionary function of our mental capacities in some domains of knowledge:

> Truths about abstract mathematics or physics or philosophy had no more role in shaping the cognitive capacities and dispositions of Pleistocene humans than moral truths did, yet we have been able to develop those capacities and dispositions in ways that make for reasonably reliable, truth-tracking exercises of them today, exhibiting significant *autonomy* from the particular evolutionary influences that gave us the raw materials to begin with (FitzPatrick 2016, p. 396).

First, it is important to clarify that the autonomy claim is still compatible with taking the relation-affirming horn of the dilemma, as it only entails that the relation is somewhat

indirect. It is not as though rationality involves a mysterious mental faculty that somehow emerged independently of any evolutionary factors; rather, it is simply that human mental capacities have developed significantly throughout our evolutionary history.

Clearly some aspects of human knowledge and rationality must be autonomous from distorting influences, as Street's argument itself depends on the assumption that her own philosophical reasoning is not distorted. Although one could attempt to undermine Street's argument by arguing that it is self-defeating, this does not appear to be a promising objection. It would be difficult to demonstrate that distorting evolutionary influences apply to her argument but not to philosophical reflection more generally, including the argument of this thesis. Thus, an argument along these lines would risk collapsing into a more general scepticism about philosophical reasoning. FitzPatrick also observes the importance of assuming some degree of autonomy:

> Few would deny the autonomy assumption altogether. To do so in the name of providing alternative evolutionary causal explanations of our beliefs would risk self-defeat: for if we lack the relevant intellectual autonomy across the board, then even the biologist's beliefs about evolutionary biology and its implications would just be attributable to such biological causes, rather than to reasons that provide real warrant for such beliefs within a rational framework with truth-tracking integrity (FitzPatrick 2014, Section 2.4).

Rationality is essential for knowledge and progress in science, just as it is in philosophy. Humans can develop complex philosophical arguments even though natural selection did not directly select for the capacity to think about philosophical topics, such as metaphysics (FitzPatrick 2015, p. 886). Nevertheless, having a capacity for autonomous rationality does not necessarily entail that we can reliably track mind-independent moral facts. After all, there is a gap between ruling out a distorting influence and affirming the truth-tracking reliability of a mental capacity. When faced with this explanatory gap, realists can maintain that their response to EDAs does not require them to provide a full positive account of their view. EDAs simply provide a new scientific challenge to a position that has long been debated on purely philosophical grounds. Given that philosophers already struggle to convince their opponents of the merits of their theories, it would be unreasonable to expect realists to fully vindicate their philosophical position in addition to overcoming evolutionary challenges. Nevertheless, moral realists need to at least establish the possibility of a positive argument that would resist debunkers' claims.

Fortunately for realists, there is at least one clear example to demonstrate that the norms of rationality do not need to have been directly provided by evolutionary forces in

order for us to be capable of tracking mind-independent truths: the example is the case of scientific knowledge. Complex theories about the physical properties of the universe, such as subatomic particles and various principles of astrophysics, clearly extend well beyond the basic capacities and knowledge that our ancestors needed in order to survive. As FitzPatrick argues, a 'parsimonious' evolutionary explanation of our reasons for believing certain scientific claims would hardly be conducive to any progress in science (FitzPatrick 2016, pp. 398-399). Street would not want to claim that we only believe certain scientific theories due to some contingencies of human evolution, yet her appeal to parsimony could undermine scientific knowledge just as she takes it to undermine the realist's moral knowledge. Street tries to maintain a distinction between these domains of knowledge, by referring to the evolutionary advantages of truth-tracking perceptual capacities to exclusively defend science. However, this fails for the aforementioned reason that modern scientific knowledge extends well beyond our evolved sensory capacities. Rationality provides the link between observable phenomena and abstract scientific theories, and the naturalistic realist can argue that this parallels the case of moral knowledge: we can perceive natural properties with our senses, but we must use our rational capacities to recognise the moral status of some of these properties.

These points can allow moral realists to tackle the horn of Street's dilemma in which they affirm a relation between evolution and mind-independent moral facts, by comparing moral knowledge to scientific knowledge. Thus, realists may argue that the capacity to track moral facts through rational reflection is a by-product of our evolved mental capacities that were originally selected for relatively basic reasoning processes. This by-product claim allows realists to not only explain how we can discover obscure mind-independent natural facts in the sciences, but also why the same should be possible for mind-independent moral facts. On this point about comparing scientific and moral knowledge, moral realists can also question Street's treatment of different disciplines: if she is permitted to speculate so freely about our evolutionary history, why must realists be held to a higher standard of explanation when defending their view? Street's explanatory expectations of each discipline appear to be the wrong way around: moral philosophy should not be expected to be as precise as science, particularly the natural sciences. Given the more speculative nature of moral philosophy, it may not be reasonable to expect any method of 'testing' or confirming moral facts. As noted earlier, the issue of different explanatory standards across science and philosophy is relevant to all EDAs due to their appeal to evolutionary explanations.

Although Street does not address the issue of different explanatory standards across disciplines, she does address the 'by-product hypothesis', anticipating it as a possible objection to her argument. She responds by presenting a slightly modified objection to the realist's second option when faced with the dilemma. The modified objection is presented against the scenario in which the realist now affirms a more indirect relation between our evolved evaluative tendencies and the mind-independent moral facts, which we can detect as a by-product of these tendencies' selection for more basic reasoning. The only difference between this and the original second option is the by-product claim, as the realist in the original dilemma was presumed to be claiming a more direct evolutionary cause of our moral truth-tracking capacities. The other option of denying any relation rules out the possibility of a by-product account in the first place, as even the most indirect relation would be more than just a coincidental alignment between beliefs and facts. As such, the status of the denial option is unchanged.

Since the denial option is unaffected by the introduction of the by-product hypothesis, Street focuses on how this hypothesis affects the realist's other option. She denies that realists can compare a moral truth-tracking capacity to the by-product of advanced scientific knowledge that has developed from a more basic capacity to perceive simple facts about the physical world (Street 2006, pp. 143-144). Her main point is the following:

> [T]he realist has to give some account of how this more basic sort of ability to grasp independent evaluative truths arose. And given what has to be the complexity and specialization of even this more basic ability (a point of comparison is the complexity and specialization of the more basic abilities on which the ability to do astrophysics is based), it is implausible to suggest that the emergence of this more basic ability was a mere fluke. The only alternative to saying that the emergence of this ability was a fluke is to claim that we were in some way selected to track the independent evaluative truths posited by the realist (Street 2006, p. 144).

Since Street directly targets non-natural moral realism, it is worth focusing on the non-naturalist's potential line of response first. It is not clear why non-naturalist realists must accept that even a more basic capacity to detect independent moral truths must be highly complex and specialised. Non-naturalists could argue that although morality in general is a complex phenomenon, it has its foundations in relatively simple moral facts. Thus, even though it would be more difficult for non-naturalists to draw a comparison with scientific knowledge, they can at least identify a similar structure of complex knowledge built upon basic foundations: it may be similar to how scientific theories are ultimately built upon

very basic perceptual beliefs. Non-naturalists could also reduce the force of Street's claim by arguing that an imperfect tracking capacity is sufficient; we do not need to always correctly identify moral facts to regard a tracking capacity as being reliable. Moral realists are not necessarily committed to the view that we can know all or even most moral facts; their only definite commitment is to the existence of moral facts in general.

Compared to simply perceiving the external world, morality is a more complex phenomenon to understand. This is precisely why rational reflection is necessary in order to identify moral facts. As such, Street demands too much from moral realists if she holds a moral truth-tracking capacity to the same standard of reliability as our evolved perceptual capacities, which reliably track simple natural facts such as the presence of predators in one's environment. Overall, although non-natural realism may ultimately be more difficult to defend than naturalistic realism, it is clear that non-naturalists do not need to accept the commitments that Street seems to expect from their view. The focus will now turn to the naturalistic perspective that seems to provide a stronger case against Street's debunking argument.

Moral naturalists are on safer ground if they appeal to the by-product view, as they can argue that moral facts are related to or identical with certain natural facts. Whereas it is questionable whether non-natural facts even exist, the naturalist only needs to show that some natural facts are identifiable as moral facts. If this can be shown, then it allows naturalistic moral realists to affirm a relation between our evolved mental capacities and mind-independent natural moral facts. Street's claim about the complexity required of a moral truth-tracking capacity thus appears to have less force against naturalistic realism, since a capacity to identify the moral status of natural facts is presumably less complex than a purported capacity to identify causally inert non-natural facts. However, for the purpose of resisting Street's argument, the more important difference between naturalistic and non-natural realism is that the former view does not posit any causally isolated facts. This matters because the realist's by-product claim is best supported by comparing knowledge of moral facts to complex scientific knowledge, which only involves natural facts. Given the success of science, it is reasonable to assert that even the more abstract scientific facts are causally connected to us through our rational reflection. Just as scientists' knowledge of astrophysics utilises advanced mental capacities that are a by-product of our more basic evolved mental capacities, our knowledge of the moral status of certain natural facts may also be a by-product of the same capacities. Thus, naturalists can escape Street's dilemma by affirming a by-product relation between evolutionary

influences and independent natural moral facts. Although the by-product view involves a somewhat indirect relation, this is still distinct from taking the other horn of the dilemma.

It is worth elaborating on the importance of the non-natural versus natural conception of moral facts in Street's argument. From a moral naturalist's perspective, one of the main problems with her argument is that it is primarily based on the non-naturalist's metaphysical distinction between moral and natural facts. If it is plausible for moral naturalists to compare knowledge of natural moral facts with our knowledge of obscure non-moral facts, Street would then be unable to maintain her own philosophical and scientific claims while simultaneously rejecting naturalists' philosophical claims. Removing the non-naturalist's sharp distinction between natural and moral facts actually creates a dilemma for Street: on one hand, she could accept that rational reflection is reliable in both the moral and scientific domains, thus defeating her argument against moral realism. On the other hand, she could assert that rational reflection is unreliable in general, which would undermine scientific knowledge along with moral realism. It is safe to assume that Street's argument is not intended to encourage any scepticism about science, given her own appeals to evolutionary theory. Knowledge of obscure scientific facts does not require a miraculous coincidence between our evolved mental capacities and the mind-independent natural facts about the universe. After all, natural selection clearly did not directly select for the ability to track facts about astrophysics, evolutionary biology or advanced mathematics. This demonstrates that we can extend well beyond the original function of our basic evolved mental capacities, so we have no need to reject scientific theories or moral realism on evolutionary grounds. Thus, rather than rejecting both scientific and moral knowledge, it is surely more reasonable to reject Street's view that knowledge of independent moral facts cannot occur as an evolutionary by-product. Ultimately, to avoid total scepticism in any given domain of knowledge, we must assume the truth of some of our basic beliefs in that domain.

Although the points in this chapter undermine Street's argument against moral realism, realists obviously still need to develop their positive account of moral knowledge and the nature of moral facts. However, they do not owe a fully developed account to debunkers, as the finer philosophical details of moral realism are fundamentally separate from evolutionary considerations. More generally, defenders of any philosophical position only need to demonstrate that their views are not undermined by any relevant EDAs, as has been done in this chapter. While this section has focused on rational reflection and the associated by-product view as a response to Street's EDA, some of Street's responses to

potential criticisms of her argument are focused more on moral naturalism. As such, the next section examines Street's objections specifically to naturalistic moral realism.

2.3 Street's Opposition to Moral Naturalists

Street has several objections to naturalistic versions of moral realism. She first describes one possible naturalistic account of moral realism: "Given that we have the evaluative attitudes we do, evaluative facts are identical with natural facts N. But if we had possessed a completely different set of evaluative attitudes, the evaluative facts would have been identical with the very different natural facts M" (Street 2006, p. 136). Street gives this example mainly to clarify that she would not regard this as a properly realist view, as it makes the moral facts depend on which beliefs we happen to have. This view would not be subject to the Darwinian dilemma, as it is closer to the anti-realist position that Street endorses. Thus, it only avoids the dilemma by failing to be a properly realist view. A genuinely realist position would not make the moral facts depend in any way on our actual or possible moral beliefs (Street 2006, p. 137).

Street then considers an example of what she regards as a genuine realist view, specifically focusing on the kind of view defended by philosophers such as David Brink and Nicholas Sturgeon (Brink 1989, 2001; Sturgeon 1985). As Street understands them, their views state that we can determine which natural facts are moral facts by following standard practice in moral philosophy. Specifically, this refers to the method in which philosophers begin with certain moral judgements that seem intuitively true, then attempt to incorporate them into a coherent moral theory. In the case of naturalism, the moral judgements must also be compatible with a naturalistic worldview (Street 2006, pp. 139-140). Street opposes this type of realist view by posing essentially the same dilemma as before, since this view begins by taking for granted some of our default moral judgements that are potentially 'saturated' with evolutionary influence. We may recall that the more viable horn of the dilemma for realists to take is the option of affirming a truth-tracking relation between independent moral facts and our evolved evaluative tendencies. However, Street thinks it is even less plausible for realists to posit a truth-tracking relation in this particular naturalistic account of realism: "[I]f the tracking account failed as a scientific explanation when it came to arguing that we were selected to track independent evaluative truths, then it will fail even more seriously when it comes to arguing that we were selected to track independent facts about natural-normative identities" (Street 2006, p. 141). The

term 'natural-normative identities' refers to the identification of moral facts with some particular natural facts, which is an aspect of the naturalistic accounts given by the aforementioned philosophers. Thus, Street is suggesting that it is especially unclear how a capacity to identify moral facts with certain natural facts could be relevant to the survival or reproductive success of our ancestors. Therefore, there is apparently no reason to think that such a capacity would be favoured by natural selection.

However, this notion depends on Street's aforementioned points against the by-product hypothesis, which have been opposed on the grounds that undermining the by-product claim would undermine scientific knowledge. The comparison to science remains relevant here, as the concept of a 'natural-normative identity' hardly seems as complex or obscure as any advanced scientific theory. For instance, knowledge of astrophysics is at least as obscure as knowledge of the moral status of certain natural facts; if anything, it is more likely that moral knowledge would have affected the reproductive success of our ancestors. After all, a great deal of our scientific knowledge is much further removed from the everyday concerns of human beings, compared to moral knowledge. So unless we are willing to be sceptical of science, we cannot simply rule out the possibility that we are capable of reliably tracking natural-normative identities. Importantly, we need not suppose that early humans had the sophisticated concept of 'natural-normative identities' as we understand it now. As with any other philosophical concept, this results from rational reflection, which is a by-product of the more basic mental capacities possessed by our ancestors. These basic capacities would merely need to allow our ancestors to identify certain natural facts as 'calling for' certain actions or meaning that they 'ought' to act in certain ways. It is reasonable to assume that these evaluative tendencies could have influenced our ancestors' behaviour, sometimes in ways that affected their chances of survival and reproduction.

The important point is that basic mental capacities would be sufficient for the by-product hypothesis. Considering that we have the by-product of many highly successful scientific theories (their success is exemplified by modern technology) despite their irrelevance to our ancestors' survival, it seems at least equally likely that our capacity for rational reflection also enables us to track natural moral facts. Thus, a capacity to track natural-normative identities is not obscure in relation to evolutionary pressures, so long as the by-product view is plausible.

Apart from the issue of natural-normative identities, Street also examines another way that moral naturalists may attempt to overcome the Darwinian dilemma. Defenders of

naturalistic moral realism may claim that there is at least one case in which independent moral value can be identified with a certain natural property, without being undermined by the dilemma. Specifically, they could argue that the natural property of pain is intrinsically linked to mind-independent moral value. An argument to this effect could begin with the highly plausible assumption that pain is bad regardless of whether anyone thinks it is. Realists could thus claim that pain is intrinsically morally bad (in the sense of having negative moral value) due to its phenomenological qualities. This claim could then be used to demonstrate a truth-tracking relation between evolutionary influences and mind-independent moral truth. The capacity to experience pain clearly has evolutionary benefits since it discourages actions that could reduce one's chances of survival. However, if the badness of pain is due to its phenomenological character rather than its evolutionary function, then this would suggest that humans have evolved to recognise independent facts about the moral status of pain. Unlike some moral beliefs that change over time due to cultural factors and philosophical developments, the badness of pain remains a constant mind-independent truth if it is intrinsically linked to its negative phenomenology.

One of Street's main objections to a truth-tracking account of pain's negative value is that it is less scientifically plausible than a simpler adaptive link explanation. This is essentially the same line of argument that Street used against the more general truth-tracking account. No matter what natural facts one identifies moral facts with, the adaptive link explanation is always more parsimonious and seemingly better at explaining the phenomenon. In this case, it is simpler to say that we believe pain is bad because this belief tendency promoted behaviour that was conducive to our ancestors' survival and reproduction, thus allowing them to pass down this tendency (Street 2006, p. 151). In Street's view, mind-independent facts about the moral status of pain would have no explanatory role to add to the most scientifically plausible evolutionary account of pain. An analogous argument also applies to the positive mental states of happiness and pleasure, since they also plausibly have an evolutionary function.

It must be admitted that it is philosophically challenging to attempt to demonstrate that pain is intrinsically morally bad. It may even be impossible to 'prove' such a reductive naturalistic identification, since it is a fundamental claim upon which other moral assertions can be based. However, Street's argument does not necessarily change the status of existing philosophical problems of this kind. This is because philosophical debates employ our capacity for rational reflection, and it has been argued that Street cannot undermine this mental capacity without also undermining her own argument. Thus, philosophical reasoning in general is vindicated by the points used against Street. As such,

debates concerning moral value theories can continue in the same way that they proceeded prior to any challenges from EDAs. This is not to suggest that all EDAs are undermined by a defence of rational reflection; as will be seen in later chapters, the other discussed debunking arguments do not specifically try to undermine this capacity. More generally, this thesis does not aim to resolve any philosophical debates that have existed prior to evolutionary considerations from EDAs. Thus, for metaethical views such as moral realism and value theories such as hedonism, the only aim is to provide a response to EDAs in particular.

Overall, it has been argued that Street's argument fails to undermine moral realism on evolutionary grounds, regardless of the philosophical merits of realists' positive arguments for their view. Although it is still a challenge for realists to vindicate their theory in a purely philosophical context, the rejection of any prominent EDAs targeting moral realism will allow this theory to revert to the status it held in philosophical debates prior to EDAs. It will be argued that the same applies to the philosophical views targeted by the other debunking arguments that are discussed in the upcoming chapters. Finally, while it has been suggested that non-naturalist realists can attempt to reduce the force of Street's EDA, her argument provides us with one reason to favour a naturalistic account of moral realism. It has been argued that naturalists have a much stronger defence in the comparison to science and the associated by-product hypothesis. With these points in mind, the next chapter examines whether Richard Joyce's EDA can avoid the problems facing Street's argument.

Chapter 3 – Richard Joyce's Evolutionary Debunking Argument

3.1 Overview of Joyce's Argument

Richard Joyce's argument is another very prominent EDA in the literature (Brosnan 2011; Carruthers & James 2008; Cline 2015; Das 2016; FitzPatrick 2014, 2015; James 2009; Joyce 2006; Kahane 2011; Leibowitz & Sinclair 2017; Shafer-Landau 2013; Toner 2011; Tresan 2010; Wielenberg 2010). As with Street's argument, Joyce draws upon evolutionary biology to support a certain philosophical position. His aim is to undermine the epistemic justification of moral judgements in general, leading to a conclusion of global moral scepticism. As such, his debunking argument also has an epistemological rather than metaphysical focus. Joyce's argument may be considered more ambitious than Street's in terms of its conclusion, as moral scepticism is an unpopular epistemic viewpoint even without the addition of controversial evolutionary claims. However, Joyce actually considers his argument to be less ambitious than Street's, at least in the sense that his debunking target is not moral realism in particular (Joyce 2013, p. 140). Moral realists are nevertheless still affected by Joyce's argument, since they are presumably required to be able to justify their belief in at least some moral facts. Thus, Joyce's broad debunking target of moral belief justification entails that his argument could potentially have far-reaching consequences for moral philosophy if left unchallenged.

The current section of this chapter outlines most of Joyce's argument, with most of the critical analysis occurring in the following sections. This chapter focuses on Joyce's argument as developed in his 2006 book titled *The Evolution of Morality*, as this contains his most detailed and widely discussed presentation of the argument. It should also be noted that the terms 'judgement' and 'belief' will be used interchangeably.

Joyce's argument begins with a detailed evolutionary genealogy of morality. Since his evolutionary account spans several chapters, only the most important details will be covered. To begin with, Joyce notes that psychological mechanisms that were adaptive in our distant evolutionary past are not necessarily adaptive in the present day, as our living circumstances have changed drastically over time (Joyce 2006, p. 5). The next notable point is that Joyce acknowledges the role of non-evolutionary influences on the content of moral beliefs, just as Street admitted in her argument:

Though there is no doubt that the content and the contours of any morality are highly influenced by culture, it may be that the fact that a community has a morality at all is to be explained by reference to dedicated psychological mechanisms forged by biological natural selection. That said, it is perfectly possible that natural selection has taken some interest in the content of morality, perhaps favoring broad and general universals (Joyce 2006, p. 10).

This passage represents one of the closest similarities between Joyce's and Street's arguments. On the possibility that natural selection has favoured some universal types of moral beliefs, Joyce refers to these universal elements of morality as 'fixed' content. This includes the types of judgements that tended to enhance reproductive fitness in the varying conditions of our ancestral environment (Joyce 2006, p. 10). Presumably, this fixed moral content would be along the lines of Street's six examples of widespread adaptive beliefs, as discussed in Chapter 2. This appears to be supported by Joyce's reference to cross-cultural studies that have identified various universal elements in human morality: negative judgements about harming others, values relating to fairness and reciprocity, different expectations depending on one's status in a social hierarchy, and regulations concerning bodily matters (Joyce 2006, p. 65).

Another aspect of Joyce's evolutionary account that is similar to Street's account is the idea that evolutionary forces would have operated on the most basic elements of moral judgements, rather than the specific content of the full-fledged judgements we are now capable of making. This is evident in Joyce's second chapter, in which he considers how the biological theory of kin selection may explain why natural selection would favour certain prosocial emotions such as love for family members. He notes that an evolutionary explanation of prosocial emotions and behaviour is not sufficient to explain the phenomenon of morality as we understand it now (Joyce 2006, p. 49). The reason is that we can imagine a group of humans who cooperate effectively due to their prosocial emotions, without having to ascribe the concept of moral judgements to them:

These imaginary beings have inhibitions against killing, stealing, etc. They wouldn't dream of doing such things; they just don't want to do them. But we need not credit them with a conception of a prohibition: the idea that one shouldn't kill or steal because to do so is wrong. And moral judgments require, among other things, the capacity to understand prohibitions (Joyce 2006, p. 50).

This is similar to Street's point that natural selection would have operated on 'proto' versions of what we now regard as moral judgements. Joyce also speculates that the concept of moral judgements in particular—as opposed to just prosocial emotions—would

have been likely to increase reproductive fitness due to a link between moral judgements and adaptive behaviours:

> [S]elf-directed moral judgment may enhance reproductive fitness so long as it is attached to the appropriate actions. We have already seen that the 'appropriate actions'—that is, the fitness enhancing actions—will in many circumstances include helpful and cooperative behaviors. Therefore it may serve an individual's fitness to judge certain prosocial behaviors—her own prosocial behaviors—in moral terms (Joyce 2006, pp. 108-109).

Although making certain moral judgements does not guarantee that an individual will perform any particular action, these judgements can still serve an evolutionary function if they at least raise the likelihood of adaptive behaviours occurring (Joyce 2006, p. 114). Furthermore, Joyce suggests that the evolutionary benefits of making moral judgements may be due to the idea that only moral rules can justify rules and regulations in a society. To use his example, an emotional state such as anger may cause a desire to punish someone, but only the judgement that they have done something morally wrong can be used to justify their punishment from the perspective of their society. In this way, moral judgements may be more effective than emotions at influencing the social behaviour of groups (Joyce 2006, p. 117).

It is notable that Joyce's argument focuses more on the evolutionary genealogy of moral judgements rather than the question of how they could be connected to moral truth. This is because Joyce intends to debunk the justification of moral judgements in general, rather than a particular metaethical view such as moral realism. To this end, the evolutionary component of Joyce's argument develops the idea that humans have evolved to have an innate 'moral sense' that guides our moral judgements. Joyce uses this broad term to describe several different aspects of morality: our tendency to think in terms of certain moral concepts, our capacity for moral emotions such as guilt, and our inclination to make moral judgements. Joyce goes into much detail while drawing upon various studies in developmental psychology to make the case that the moral sense is innate in human beings (Joyce 2006, pp. 133-139). This chapter will accept the notion of an innate moral sense for the sake of argument, as a proper assessment of this idea would require a level of biological detail that is beyond the philosophical focus of this thesis. The following passage from Joyce highlights his view of the relation between our innate moral sense and our actual moral judgements: "No one would deny that cultural learning plays a central role in determining *the content* of the moral judgments that an individual ends up

making; the claim is that there is a specialized innate mechanism (or series of mechanisms) designed to enable this type of learning" (Joyce 2006, p. 137).

Now that most of the important details of Joyce's evolutionary account of morality have been summarised, the focus will turn to Joyce's philosophical claims in his sixth chapter; this is the point at which he presents his debunking argument. While noting the obvious fact that every belief has a causal history, he observes that in some cases our knowledge of a belief's genealogy can undermine our confidence in that belief (Joyce 2006, p. 179). The concept of genealogical debunking is demonstrated by a thought experiment that plays an important role in Joyce's argument:

> [P]retend there were such things as belief pills, such that taking one would inevitably lead to the forming of a certain particular belief (while at the same time invoking amnesia about the taking of the pill and, to be on the safe side, amnesia about the existence of such pills in general). Suppose that there were a pill that makes you believe that Napoleon won Waterloo, and another one that makes you believe that he lost. Suppose also that there were an antidote that can be taken for either pill. Now imagine that you are proceeding through life happily believing that Napoleon lost Waterloo (as, indeed, you are), and then you discover that at some point in your past someone slipped you a 'Napoleon lost Waterloo' belief pill. It is not a matter of your learning of the existence of such pills and having no way of knowing whether you have ever taken one; rather, we are imagining that you somehow discover beyond any shred of doubt that your belief is the product of such a pill. Should this undermine your faith in your belief that Napoleon lost Waterloo? Of course it should (Joyce 2006, p. 179).

The belief pills in this thought experiment are supposed to be analogous to the evolutionary process of natural selection, while the Napoleon belief is an analogy for moral beliefs in general (Joyce 2006, p. 181). There are problems with this thought experiment's representation of our epistemic position, as will be discussed in Section 3.2. For now, it is important to note Joyce's intended message: although knowing that one has taken a belief pill is not sufficient to show that the resulting belief is false, it does undermine one's justification for holding that belief. It is possible to obtain knowledge through unreliable methods and causes; this is why the thought experiment can only target the justification of beliefs, not their truth. In the belief pill case, it is clearly assumed that we have a reason to think the pill (representing natural selection) is an unreliable source of knowledge. The thought experiment would not work without this assumption of

unreliability; it will later be shown that this assumption is a problematic aspect of Joyce's argument.

Joyce thinks that in the belief pill case, discovering the unreliable cause of the belief should lead us to scepticism about that belief, at least until we obtain reliable independent evidence for or against it (Joyce 2006, p. 180). Importantly, suspending one's judgement is not the same as assuming that it is probably false. Thus, Joyce's conclusion that we should be sceptical of pill-induced beliefs applies to the analogous real case as follows: our discovery of the evolutionary causes of our moral judgements should lead us to suspend those judgements. Despite Joyce's recognition of non-evolutionary influences such as culture, it appears that the notion of an innate moral sense underlying all our moral concepts provides his reason for global scepticism, rather than selectively undermining only moral beliefs that have clear evolutionary explanations. So although the belief pill thought experiment only involves one particular belief, it is just an illustration of how global evolutionary influences would apply to one particular case.

Joyce anticipates the objection that natural selection does not affect our psychological traits in the same way that the belief pills directly cause certain beliefs. For instance, it may be objected that our evolved mental traits do not inevitably develop in a certain way without being affected by external factors. However, Joyce maintains that a belief requiring environmental input should not automatically be considered a learned belief, in the sense of contrasting it with an innate belief (Joyce 2006, p. 180). This is due to his claim that there is an innate psychological mechanism that enables the acquisition of moral beliefs; as such, all moral beliefs are fundamentally a product of this evolved mechanism (Joyce 2006, pp. 180-181). In order to demonstrate this point that the evolutionary influences on moral beliefs occur at the fundamental level of general moral concepts, Joyce presents a modified version of the belief pill scenario:

> Suppose that the imaginary belief pills do not generate particular propositional beliefs but, rather, dispose you to form beliefs involving a particular concept—a concept that otherwise wouldn't figure in your beliefs. Thus, rather than a pill that makes you believe that Napoleon lost Waterloo, it's just a 'Napoleon pill' that makes you form beliefs about Napoleon in general. Without this pill you would never have formed any beliefs about Napoleon at all. We needn't worry too much about what other factors determine the precise content of these Napoleon beliefs; perhaps it is determined randomly, or perhaps there are certain environmental triggers… (Joyce 2006, p. 181)

As with the first version, Joyce asks us to suppose that we discover with absolute certainty that we have taken such a pill. He maintains that this new scenario would undermine any beliefs we have about Napoleon, which represents the justification of all moral judgements being undermined. This follows from Joyce's view that all it takes to undermine a belief is for any concept involved in that belief to be undermined (Joyce 2006, p. 181). As such, the genealogy of a belief is said to have a debunking effect even if the specific content of the belief is also shaped by other factors. Joyce's global debunking of moral judgements clearly requires that our innate moral concepts must be very general. As such, he lists a few examples of fundamental moral concepts, such as obligation, fairness and virtue (Joyce 2006, p. 181). Joyce asserts that it is only due to our particular evolutionary history that we have certain moral concepts such as these. As a result, Joyce thinks we should be agnostic with respect to any beliefs that involve our innate moral concepts. He does not seem to consider it a possibility that any widespread moral beliefs could be sufficiently detached from the concepts granted by our innate moral sense.

Joyce is aware of the possible objection that his belief pill thought experiment misrepresents the real case of natural selection and moral beliefs: "It may be objected that in the case of the belief pills the story has been carefully stipulated such that forming a belief as the result of taking a pill is entirely independent of whether or not the state of affairs necessary to render the belief true obtains in the world" (Joyce 2006, p. 182). However, Joyce maintains that we have no reason to think that natural selection would have led us to reliably form true moral beliefs. To this end, he follows a similar strategy to Street in arguing that moral knowledge is unlike other domains of knowledge, in terms of the evolutionary advantages of reliably tracking truths. Joyce refers to the case of mathematical knowledge to demonstrate this point. He begins by assuming that our most basic beliefs concerning arithmetic are innate, such as the belief that $1 + 1 = 2$. An innate capacity for this type of reasoning would apparently be favoured by natural selection since we can conceive of its relevance to survival in prehistoric times: "False mathematical beliefs just aren't going to be very useful. Suppose you are being chased by three lions, you observe two quit the chase, and you conclude that it is now safe to slow down" (Joyce 2006, p. 182). Joyce thus thinks that a genealogical explanation of mathematical beliefs does not debunk them, although this raises the question of whether our highly developed capacities for abstract mathematical reasoning can be traced back to the evolution of much more basic capacities. Clearly, something like the by-product view discussed in the previous chapter would be necessary in order to vindicate abstract mathematical knowledge as with scientific knowledge.

Joyce also parallels Street's argument by claiming that unlike mathematical beliefs, the evolutionary benefits of making moral judgements can be explained without having to suppose that any of them are true. He refers back to his evolutionary account of how judging actions in moral terms could promote prosocial behaviours and thus be adaptive, due to the benefits of living in a social group with regulated behaviour. Unlike mathematical beliefs, the truth of any moral belief seems unlikely to have ever had effects on adaptive behaviour (Joyce 2006, p. 183). Even if it is necessary for humans to believe there are objective moral truths in order for moral beliefs to sufficiently motivate adaptive behaviour, believing something is obviously not enough to make it true. Joyce's points about truth detection only having evolutionary benefits for certain types of beliefs leads to the possibility of his argument being self-defeating, as he recognises. Clearly Joyce's argument requires that the evolved mental faculties used in scientific reasoning—such as the evolutionary speculations that he draws upon—must be capable of reliably tracking the truth. Although he grants that these mental faculties are the product of natural selection, Joyce defends scientific reasoning by comparing it his aforementioned example of basic arithmetic. Presumably, the more advanced levels of scientific reasoning are ultimately built upon the basic reasoning capacities that were adaptive due to their truth-tracking reliability. These points allow Joyce to protect his argument from his own debunking claims (Joyce 2006, p. 183).

Joyce recognises that merely defending the mental capacities that he uses to formulate his debunking argument is not sufficient to undermine the justification of moral judgements. His points that have been discussed so far do not rule out the possibility of some other relation between moral facts and natural evolutionary facts; there could be a relation that makes the notion of reliable moral judgements compatible with Joyce's evolutionary genealogy of our moral sense (Joyce 2006, p. 184). Since a naturalistic account of morality could posit such a relation, Joyce devotes much attention to the task of rejecting moral naturalism. The issue of moral naturalism will be covered in Section 3.3, as it is essentially a separate component of Joyce's overall debunking argument.

After rejecting moral naturalism, Joyce very briefly considers two other metaethical views: moral supernaturalism and moral non-naturalism. The former view holds that some supernatural entity or entities are the only source of moral truths; this is central to many religious perspectives on morality. In contrast, non-naturalists do not need to posit any 'divine' entities, but they still regard moral facts as being distinct from natural facts. Both non-naturalists and supernaturalists hold that moral properties exist beyond the natural world that can be described by the sciences. Whereas Joyce required a separate argument

against naturalism, he quickly dismisses these other metaethical views by referring back to his evolutionary genealogy of morality and combining it with Ockham's Razor: both non-naturalism and supernaturalism explain morality by positing extra entities, but these entities can be discarded since they add nothing to the explanatory power of the evolutionary account (Joyce 2006, pp. 209-210). Joyce extends this point to dismiss other ways that one might attempt to justify moral beliefs; in particular, he suggests that epistemological theories such as coherentism and foundationalism cannot compete with the explanatory power of his evolutionary account: "[A]n 'explaining away' strategy is preferable to persisting with the problematic project of trying to show that our moral beliefs are epistemically justified, since it recognizes no remaining mysterious phenomena or unanswered questions; it is explanatorily complete" (Joyce 2006, p. 219).

Since Joyce rejects moral naturalism, non-naturalism and supernaturalism, it may seem that he is only left with moral nihilism: roughly the view that all moral judgements are false. However, he maintains that lacking any reason to believe in moral facts does not necessarily entail that we should disbelieve that they exist (Joyce 2006, p. 210). There are many cases in which we are not justified in holding a certain belief, but this does not automatically mean we are justified in believing the opposite. To use Joyce's example, a person might be unjustified in believing they have an odd number of hairs on their head, but this does not mean they are justified in disbelieving this, as this would falsely suggest that they are justified in believing they have an even number of hairs (Joyce 2006, pp. 210-211). As with Joyce's questionable belief pill analogy, it is debatable whether this example is similar enough to the case of evolution and moral beliefs. Regardless, Joyce combines this general point about justification with his evolutionary genealogy of morality, leading him to conclude that although moral beliefs could potentially be true or false, we are not justified in holding any such beliefs until we have more reliable grounds for holding them (Joyce 2006, p. 211). With the main points of Joyce's evolutionary argument now summarised, Section 3.2 critically examines the belief pill analogy and the notion of the moral sense, then Section 3.3 focuses on Joyce's rejection of moral naturalism.

3.2 Belief Pills and the Moral Sense

One of the main problematic aspects of Joyce's argument is his 'belief pill' thought experiment, in which the imagined belief pills represent evolutionary influences (specifically natural selection) and the resulting beliefs about Napoleon represent our

moral beliefs. As Russ Shafer-Landau points out, the belief pills in Joyce's analogy seem to override any other factors that could potentially affect the formation of beliefs: "The doxastic influence of a belief pill is 100 percent. Its operation is direct and guaranteed. You take the pill, you have the programmed belief, no matter whether the belief is true or not" (Shafer-Landau 2013, p. 18). Although this applies most directly to Joyce's original formulation of the belief pill story, it also applies to his modified version: even if the pills only directly determine the concepts involved in our beliefs, it is not clear that thinking in terms of these concepts would necessarily lead us towards moral truths. In other words, the belief pills still directly influence the range of beliefs we form, regardless of whether that range includes any true beliefs. If this interpretation is correct, then the belief pill story does not accurately reflect the nature of evolutionary influences on our moral beliefs. As with Street's claims about our belief tendencies being 'thoroughly saturated' with evolutionary influences, Joyce's story underestimates the impact of other factors such as rational reflection. As mentioned earlier, Joyce at least recognises other belief-influencing factors such as culture. Although when it comes to rational reflection, he only implicitly defends this capacity in the context of scientific and mathematical reasoning. Thus, it appears that the only reason he disregards this mental capacity in the context of morality is because of his claim that the moral sense does not necessarily require any connection to truth to serve its evolutionary function.

However, just as it has been argued in Chapter 2 that rational reflection allows us to attain knowledge in both the scientific and moral domains, this point can be extended to reject even Joyce's modified belief pill scenario. After all, even if we grant that the general concepts in our reflective moral judgements are traceable to evolutionary factors, this does not undermine the fact that we can apply rational reflection to moral issues, just as we use this capacity in science. It does not seem that debunkers such as Joyce could object to this point, as one does not need to hold any particular metaethical view to recognise the complexity of philosophical reasoning that humans are now capable of; this complexity is even demonstrated by Joyce's own argument. Thus, any possible differences in the evolutionary origins of moral reasoning as opposed to other types of reasoning become less relevant now that our reasoning capacities in general have extended so far beyond their original functions.

Apart from the issue of ignoring other potential influences such as rational reflection, the belief pill analogy is also questionable in terms of how much it overestimates the direct influence of evolution on particular moral beliefs. As discussed in Chapter 2, natural selection could only indirectly affect the kinds of beliefs that humans

have in the present day. The degree of influence could vary significantly depending on the type of belief; for instance, we would expect greater influence on the kinds of beliefs that would have been salient in our evolutionary past, such as Street's examples relating to survival and reciprocity. These types of beliefs might require more extensive rational reflection to eliminate any indefensible evolutionary biases. Regardless, so long as we have the capacity to reflect upon any of our moral beliefs and potentially reject or alter them, this weakens the force of evolutionary influences as compared to the belief pills in Joyce's analogy. This line of reasoning follows Shafer-Landau's objections to Joyce's EDA, as he refers to the potential for other 'doxastic operations' to correct for any distorting evolutionary influences (Shafer-Landau 2013, p. 19).

Shafer-Landau raises another related point: evolutionary debunkers have not been able to demonstrate that evolutionary influences are different from cultural and historical influences, in terms of our ability to correct for biases (Shafer-Landau 2013, p. 19). Given that Joyce and Street do not present cultural debunking arguments, they do not seem to regard these other factors as distorting influences. As such, we should question why the case of evolution might be different. One possible explanation that debunkers might favour is the idea that only biological factors can 'hardwire' certain tendencies in humans, whereas cultural factors can only cause us to acquire certain belief-forming tendencies through learning; this presumably makes learned beliefs more susceptible to the effects of rational reflection. The concept of an innate biological mechanism is important for Joyce's argument in particular, as he depends on the idea of an innate moral sense that explains why we have our particular moral concepts. Joyce tries to maintain a distinction between this moral capacity and our other reasoning capacities (as used in science and mathematics) by disregarding a possible objection to the notion of an innate moral sense:

> First, it might be pointed out, no human trait is 'hard-wired' in the sense of developing inevitably, irrespective of environmental factors; an innate belief would, like any other phenotypic adaptation, require environmental input in order to become manifest. This observation does not, however, compromise the very notion of an innate belief. A belief that requires environmental input is not thereby a learned belief, and not thereby a belief that is formed in a way sensitive to the evidence (Joyce 2006, p. 180).

Joyce notes that although cultural learning plays an important role in forming the content of our moral beliefs, his evolutionary story only claims that we have a specialised innate mechanism for acquiring general moral concepts. Thus, moral concepts may be innate even if the content of specific moral beliefs is influenced by external factors (Joyce 2006,

pp. 180-181). This point leads Joyce to argue that even our reflective moral beliefs are not justified since they draw upon the concepts granted by our moral sense, which evolved with no apparent need for any connection to truth. As he puts it, "We have no reason to think in the case of the moral sense that natural selection is likely to have produced true beliefs" (Joyce 2006, p. 182). In this way, both Joyce and Street attempt to debunk morality by identifying reflective moral beliefs as the product of a mental faculty that presumably did not evolve in such a way as to reliably track moral truths.

Since the concept of an innate moral sense is central to Joyce's argument, it is important to scrutinise this aspect of the argument when responding to his evolutionary challenge. One problem concerns the way he considers the relation between the moral sense and the content of our reflective moral beliefs. Christopher Toner focuses on this issue, pointing out that Joyce does not go into detail regarding the genealogy of any particular moral judgement, instead just listing a few general moral concepts:

> Joyce sometimes slides without warrant from talk of a genealogy of the moral sense
> to a genealogy of moral beliefs. Although he gestures in this direction, I am pretty
> sure he has not actually given us a genealogy of a single moral belief; I am certain
> he has not given us one of all of them. Thus it remains possible that moral facts that
> played no role in the genealogy of the moral sense still play a role in the genealogy
> of moral beliefs (Toner 2011, p. 530).

This is especially problematic for Joyce when considering his claim that all moral judgements are unjustified by default, not just a subset of moral judgements that have an obvious evolutionary explanation. As such, it could be argued that Joyce faces a burden of proof, rather than defenders of moral beliefs. Joyce's radical conclusion of global moral scepticism needs to be supported with enough specific examples to convince his opponents that any conceivable moral judgement can be traced back to the concepts granted by our evolved moral sense. In the absence of such examples, Joyce's opponents can maintain that although the evolution of our moral sense may not have required any reliable connection to moral truths, our reflective moral judgements can still align with such truths (Toner 2011, p. 530). As with the similar point against Street's argument, this depends on the idea that rational reflection allows us to track moral truths as a by-product of this mental capacity's selection for relatively basic reasoning capabilities.

Furthermore, just as the by-product view was defended by comparing moral knowledge to scientific knowledge, Joyce's notion of an innate moral sense undermining moral judgements can be opposed by comparing the moral sense to other evolved traits

that are reliable despite their selection for other functions. This point is highlighted by Toner:

> Consider vision, which depends upon eyes possessing a certain transparency. The transparency of eyes stems from their being composed largely of crystallins in close-packed arrays. But these crystallins were originally selected for purposes (such as heat resistance) that had nothing to do with perception. Are we thereby to conclude that vision is unreliable? Certainly not, for we can preserve our confidence in its reliability by noting that these crystallins, originally selected for one purpose, were later co-opted for another that had everything to do with perception (Toner 2011, p. 533).

Similarly, although the moral sense may have been naturally selected due to its promotion of behaviours that supported our ancestors' reproductive fitness, it can now be utilised in the very different domain of modern moral philosophy. In this different context, judgements that are based in our moral sense are subject to reflection and gradual refinement; it can be argued that this allows us to align our moral judgements towards moral truths.

Toner's example of an evolved physical trait also provides a response to debunkers' claims that moral beliefs are unlike perceptual beliefs, in that the former would not need to reliably track features of the world in order to be favoured by natural selection. The vision example shows that even our perceptual capacities need not have been directly selected for truth-tracking functions in order for us to now have reliable perceptual beliefs. This point clearly reduces any apparent distinction between the perception case and the case of the evolved moral sense. By reducing any apparent distinction that debunkers try to maintain between our moral and non-moral faculties, it can thus be argued that reliable moral faculties are like our perceptual faculties; both are by-products of traits that were originally selected for much more basic functions.

Much of the discussion so far has focused on Joyce's concept of the evolved moral sense, as the distinction between this innate mental capacity and the content of our actual moral beliefs is important to keep in mind when examining his belief pill analogy. For reasons that were also discussed in Chapter 2, it should now be clear that even if our moral sense is a 'distorting' influence with regard to moral truth, the content of our reflective moral judgements can still be shaped by more reliable processes such as our rational capacities. Thus, we can reject the idea that evolutionary forces such as natural selection (as represented by Joyce's belief pills) would inevitably lead us to form certain types of

beliefs. As Shafer-Landau observes, evolutionary forces would only lead to sceptical concerns about morality if they actually worked like the belief pills, by forcing certain beliefs or tendencies upon us without the possibility of reflecting upon them and changing them: "[D]ebunkers have said nothing on behalf of this special status for evolutionary forces, nothing to show that evolutionary forces possess an immunity to correction that is unique among all possibly distorting doxastic influences" (Shafer-Landau 2013, p. 19).

As mentioned earlier, Joyce presents an adjusted version of the belief pill analogy in which we only form beliefs with certain moral concepts because of pills that induce these types of beliefs. The analogy was suggesting that without natural selection we would not have the particular set of moral concepts that we have actually ended up with. This is a scientific claim about human evolution, so questions about its plausibility cannot be settled with philosophical speculation. It has been noted that debunkers often rely on untestable scientific speculations to support their philosophical conclusions, while demanding a much higher degree of certainty from philosophers who oppose their debunking arguments. In Joyce's case, he claims to have shifted the burden of proof over to those who wish to reinstate the justification of any moral beliefs, as he argues that the evolutionary genealogy of morality makes our beliefs unjustified by default. Joyce reaffirms this point in his later work, in which he maintains that the widespread acceptance of many moral beliefs is not enough to put the burden of proof on moral sceptics (Joyce 2016). Thus, even though many moral beliefs seem initially well-justified, Joyce asserts that this intuitive feeling of justification should not just be taken for granted: "The fact that many people believe their moral beliefs to be well-justified clinches nothing; people believe all sorts of silly things" (Joyce 2016, p. 139).

In addition to the points discussed so far, another issue is that there is a discrepancy between Joyce's belief pill story and our actual epistemic position: Joyce's analogy seems to assume that we definitely know the pill is an unreliable source of knowledge. In this way the analogy begs the question against moral justification, since it only works for Joyce's argument if we already know or assume that some of the resulting moral beliefs are false. After all, we cannot be sure that a process is unreliable unless we know what proportion of the beliefs it causes are true. This point is highlighted by Erik Wielenberg, who describes Joyce's reasoning in the following terms: "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" (Wielenberg 2010, p. 462). He notes that Joyce seems to assume this is obvious, considering the lack of argument for this point. Wielenberg discusses this in the context of one of Joyce's examples that is supposed to show that holding certain false

beliefs could have evolutionary benefits; this extends to the idea that evolutionarily beneficial types of beliefs are not any more likely to be true. Joyce's example is about a belief that a certain plant has magical properties; it is assumed that we know this belief is false, although the plant does have potent medicinal properties. The false belief that such a plant is magical could promote the survival of people whose environment contains this plant. The following passage demonstrates how Joyce applies this hypothetical example to the case of moral beliefs:

> If there were an innate belief that certain plants have magical powers, and this belief were the product of a dedicated psychological mechanism with a distinct evolutionary history … since in this case we would have an empirically confirmed hypothesis of how this belief-formation mechanism works which does not require that any of the beliefs be even approximately true, we would have to conclude that any such innate beliefs are products of an *unreliable* process. … We can now apply this lesson to the case of innate moral beliefs. What seems clear is that in the crucial respect such beliefs are like the imaginary belief about the plant's magical properties. We have seen that nowhere does the evolutionary hypothesis outlined in earlier chapters assume that moral beliefs are or were true. … Thus I conclude that by the process reliabilist's own lights a certain plausible view of how innate moral beliefs may have evolved leads naturally to the conclusion that such beliefs are epistemically unjustified (Joyce 2006, p. 215).

For the current purposes, we do not need to consider the details of the epistemological theory of reliabilism that Joyce attempts to undermine. The important point here is the difference between the hypothetical plant case and the actual case of moral beliefs: while we do have prior knowledge that the plant belief is false, we cannot just assume in advance that moral beliefs are probably false. Comparing the case of moral beliefs to 'magical' beliefs would misrepresent our actual epistemic situation. We only know that the plant belief is formed by an unreliable process because it is assumed that we already know that magic-related beliefs must be false. As Wielenberg puts it, "That the mechanism routinely produces false beliefs is what makes it clear that the mechanism in question is unreliable, not the fact that the beliefs it produces can be explained without appealing to their truth" (Wielenberg 2010, p. 462). Our reasons for being sceptical of any magic-related beliefs clearly go beyond just explaining various phenomena without assuming their truth. In particular, theories involving the concept of 'magic' cannot compete with scientific theories when attempting to explain and understand any phenomenon, as evidenced by the much greater success of the latter.

Overall, it is reasonable to hold that Joyce's belief pill analogy misrepresents our actual epistemic situation regarding moral knowledge. This concludes the main objections to two of the central elements of Joyce's argument, specifically the concept of the innate moral sense and the belief pill thought experiment. Since this covers the main evolutionary aspects of Joyce's EDA, the remainder of this chapter focuses more on the metaethical aspect of his argument in which he rejects moral naturalism.

3.3 Moral Naturalism and Practical Clout

Earlier in this chapter it was noted that Joyce recognises that moral scepticism cannot be established simply by explaining moral judgements without presupposing their truth; he must also be able to rule out any relation between his genealogical account of morality and the truth of our resulting moral judgements. Moral naturalism's identification of moral facts with particular natural facts could accommodate such a relation, which is why Joyce devotes much attention to rejecting this metaethical view. Specifically, one could defend a naturalistic account of moral realism in which the moral facts are compatible with Joyce's evolutionary explanation of our moral sense. Joyce's points against moral naturalism are essentially a supplement to his main debunking argument, as he relies on metaethical claims that are fundamentally separate from his evolutionary considerations: "[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" (Joyce 2013, p. 143). He begins by drawing upon some much earlier work by Gilbert Harman, in which Harman rejected naturalistic moral realism on explanatory grounds (Harman 1977, 1986). Joyce follows a similar line of reasoning and combines it with his own evolutionary genealogy of morality to complete his debunking argument. In particular, Joyce focuses on Harman's idea that positing moral facts does not seem to add anything to a naturalistic explanation of morality. Joyce's evolutionary story is clearly such an explanation, as it does not posit any non-natural or supernatural phenomena.

One of the general ideas that Joyce follows is that naturalistic moral realism cannot be defended by merely suggesting how moral facts would fit within a naturalistic worldview if we assume that such facts could exist. Joyce illustrates this point by using the concept of ghosts in place of moral facts: "[O]ne might claim that talk of ghosts could be vindicated if it could be shown that ghostly properties may be comfortably integrated

within a naturalistic worldview. Quite so. However, this observation alone gives not a glimmer of a reason to believe in ghosts" (Joyce 2006, p. 189). As such, Joyce maintains that the burden is on moral naturalists to provide a plausible reductive account of moral facts that not only convinces us of their existence but is also compatible with his evolutionary genealogy of the moral sense.

Although Joyce considers a full argument against moral naturalism to be beyond the scope of his book, he aims to provide a preliminary challenge to this metaethical view in order to raise the difficulty of overcoming its apparent explanatory burden. Thus, he argues that naturalists cannot account for certain aspects of morality that are typically thought to be essential to it (Joyce 2006, pp. 190-191). Specifically, Joyce uses the term 'practical clout' to describe what he regards as two necessary features of moral claims and rules: they involve both a sense of inescapability and a sense of authority. The point about inescapability is that moral claims are usually taken to override apathy or any desires that conflict with the moral prescription (Joyce 2006, pp. 191-192). In other words, moral reasons take priority over selfish motivations. As for the point about authority, this notion is used to contrast morality with social etiquette: whereas a violation of etiquette is merely considered rude, a violation of morality is taken to be much more serious and 'wrong'. The authority of morality also means that anyone automatically has a reason to comply with moral rules, whether or not they want to. Supposedly, moral reasons are distinct from our reasons to comply with etiquette, since the latter can be overridden under certain circumstances (Joyce 2006, pp. 192-193).

In response to Joyce's assertion that practical clout is a necessary feature of morality, moral naturalists could either deny its necessity or try to accommodate it. However, Joyce has objections to both of these options. When examining the naturalist's option of denying this requirement, Joyce considers what naturalistic morality would be like without practical clout. He claims that such a conception of morality would not provide any grounds for caring about moral reasons:

> To be sure, there are *moral* reasons for caring about moral wrongness—that is, there are reasons pertaining to what can be justified from the point of view of that particular normative framework. But these aren't the kind of reasons you are after, for what you are really asking is why, on this occasion and pertaining to this action, you should care about that justificatory framework at all (Joyce 2006, p. 205).

If a person were to realise that they have no reason to care about the notion of moral wrongness, it would be rational for them to ignore any moral reasons that conflict with their motivations and desires. This would clearly not be conducive to any effective

conception of morality. Overall, Joyce's point is that moral naturalism without practical clout fails to capture what we expect from the concept of morality, particularly in terms of its regulation of certain behaviours. Joyce does grant that moral naturalists can identify moral properties with certain natural properties in a way that is compatible with mainstream moral views, such as the view that theft is wrong because it causes distress to others. However, he maintains that this would still fail to accommodate practical clout:

> The problem is that thinking in moral terms seems entirely superfluous to such a person's reasons and motivations. If the person wonders why she should not steal the newspaper, all the answer she is ever going to get can come from reflecting on the fact that it will cause unnecessary distress. Her identifying that causal property with moral wrongness seems to add nothing to her reasons or motivations (Joyce 2006, p. 207).

In Joyce's view, his evolutionary account therefore suggests that a naturalistic conception of morality could not explain how moral judgements tended to promote prosocial behaviours by allowing early humans to overcome their conflicting desires or their lack of willpower to follow moral prescriptions. Importantly, Joyce thinks that morality still requires practical clout in order to have this desire-overriding effect in the present day, not just in our distant evolutionary past:

> Moral naturalism without clout, first of all, seems to enfeeble our capacity to morally criticize wrongdoers; second, it might actually encourage wrongdoing for certain persons; and third, it renders moral language and moral thinking entirely redundant. Such a value system is (to recycle a phrase used earlier) surely too wimpy to be mistaken for morality. Moral thinking has a function, I have argued—both evolutionarily and contemporarily—and deliberations in terms simply of what we want and need will not suffice (Joyce 2006, p. 208).

Thus, Joyce concludes that moral naturalism without practical clout barely counts as a concept of morality, and that it does not seem possible for naturalism to accommodate practical clout in any case. Since moral naturalists can apparently neither accommodate practical clout nor deny its necessity, Joyce uses this point to reject naturalism.

One way to vindicate moral naturalism in light of these issues would be to argue that practical clout is not a necessary feature of morality, at least not in the way Joyce conceptualises it. This aspect of Joyce's argument has been examined by Jon Tresan. His interpretation of Joyce's notion of practical clout is as follows: first, the 'inescapability' of moral obligations means that they don't depend on our having certain motives; we have certain obligations even if our desires conflict with them. As for the 'authority' aspect of

practical clout, this means that moral obligations necessarily entail reasons for certain actions. Thus, moral obligations always count as reasons for or against certain actions, regardless of one's motives that may conflict with these moral reasons (Tresan 2010, p. 222).

Tresan specifically opposes Joyce's claim that moral naturalism without practical clout would make moral terms superfluous. In Joyce's view, if thinking of an action as being 'morally wrong' provides some extra meaning that is not present when only considering the natural properties involved in that action, then this counts as a point against reductive moral naturalism (Joyce 2006, p. 207). Tresan objects that this begs the question against naturalists, since it assumes that an adequate explanation of our need for moral terms must involve the idea that they add some substantial meaning that is lacking in the relevant natural terms. Naturalists have no reason to accept such a standard of explanation: "Good explanations of our need for some type of discourse aren't in general required to include a clause to the effect that the discourse does something more than facilitate our thought and talk about some part of the natural world." (Tresan 2010, pp. 235-236). In other words, the use of moral terms is acceptable when describing certain aspects or properties of the natural world that are taken to be normatively relevant. This does not undermine the naturalist's claim that the moral relevance of a particular natural property is intrinsically linked to that property, even before the introduction of moral terms to describe such a property. Although the normative relevance of any given natural property (such as pain or pleasure) can be contested, what matters is that describing such properties in moral terms does not require us to posit any non-natural properties; as such, moral terms can be included in a naturalistic worldview. In the end, any kind of terminology is just a way for humans to describe various aspects of the world in different contexts, so the use of specifically 'moral' terms does not necessarily imply the existence of any non-natural properties.

Apart from rejecting the idea that naturalism without practical clout would make moral terms redundant, moral naturalists could also maintain that their account of morality is still functional in terms of regulating behaviour, due to contingent facts about human psychology. For instance, it could be argued that the moral aspect of certain natural facts is a relational property, in that the moral significance depends not only on the intrinsic features of particular natural facts but also on the nature of humans who stand in some relation to these facts. While it is beyond the current scope to properly defend any particular account of moral naturalism, it is worth raising this possible line of argument as a response to Joyce's objections to 'cloutless' moral naturalism. Thus, even if it turns out

that 'moral' reasons only motivate us due to the necessity of prosocial behaviour for a fulfilling life, this could still provide a way to resist Joyce's claim that moral naturalism would enfeeble our ability to criticise wrongdoers. After all, we could at least criticise their behaviour on the grounds of it being irrational due to the negative consequences for themselves and for society overall; such consequences can be identified in terms of natural properties such as pain and suffering.

Joyce's other objection was a similar point, namely that naturalism without practical clout might encourage wrongdoing in some individuals. However, there will always be some individuals who are not psychologically disposed to care about morality, regardless of how philosophers conceive of it. We cannot expect everyone's behaviour to be successfully regulated by Joyce's notion of practical clout.

Overall, these points suggest that moral naturalists have various ways to resist Joyce's non-evolutionary claims about the necessity of practical clout in any system of morality, which he used to support the main evolutionary component of his argument. When combined with the previous chapter's points about moral naturalism being on safer ground than non-naturalism with regard to Street's argument, it should now be clear that naturalists are able to resist a variety of different debunking strategies.

So far, the concept of practical clout has been discussed in relation to moral naturalism. However, it is also worth considering how it relates specifically to moral realism. To examine this relation, it is useful to first recall how Joyce conceives of the evolutionary benefits of believing in a traditional conception of morality. This is one way to describe the default attitude towards morality that would be expected of most people; it refers to the conviction that morality is 'real' in some way, thus contrasting with Joyce's moral scepticism. Although Joyce does not specifically state that we have evolved to unreflectively tend towards moral realism, this is a possible interpretation of his aforementioned claim that a specifically 'moral' sense would be more effective than non-moralised prosocial emotions in terms of promoting reproductive fitness (Joyce 2006, p. 117).

While Joyce clarifies in later work that even a 'subjective' conception of morality could work in his debunking argument so long as it were still based in innate moral concepts (Joyce 2013, p. 140), it is hard to conceive of this having as many evolutionary benefits as a realist's conception of morality. Our ancestors would not need to recognise the abstract philosophical concept of 'moral realism' in order to gain evolutionary advantages; rather, they would merely need to share a common belief that there are certain

objective moral truths. Joyce's argument seems to entail that this default evolutionary tendency towards realism should lose its power over us now that we are capable of speculating about our own evolutionary history, including the genealogy of moral concepts. Without the 'illusion' of objective morality, Joyce's argument suggests that we should be moral sceptics rather than moral realists. Nevertheless, he maintains that moral thinking still has practical benefits despite our awareness of its evolutionary origins: "The question of what we ought to do, once we have come to see that our moral discourse is a philosophically indefensible illusion, is a practical question. A neglected answer is that the discourse may be maintained, accepted, but not believed – that it may have the role of a fiction" (Joyce 2000, p. 730). In other words, there can be purely pragmatic reasons to take some of our unjustified moral beliefs seriously. This theme continues in his book that has been the main focus of this chapter, as his conclusion includes some reassurance that moral scepticism does not actually change much in practical terms:

> Moral skepticism amounts to the recognition that there is, or may be, nothing distinctively morally wrong with stealing, but it is absolutely not to be identified with the proposal that ordinary people have no reason at all to refrain from stealing … To claim otherwise is to admit that the only thing standing between us and a life of savagery and rampant spoon stealing is a sense of moral duty, which is a truly depressing thought. To cast into doubt one particular kind of normative framework is not to imply that 'anything goes' (Joyce 2006, p. 224).

In other words, it may seem that our evolved tendency to favour the default idea of objective morality is no longer necessary in order to achieve morality's original function of regulating social behaviour. This demonstrates how Joyce defends his position of moral scepticism, by suggesting that the practical aspects of morality can be maintained even if all our moral judgements are epistemically unjustified.

However, Joyce's argument fails to debunk the notion of objective moral truth for essentially the same reasons that his argument fails to undermine the justification of moral beliefs in general: first of all, it was noted that our capacity for rational reflection cannot be undermined by his misleading belief pill analogy. This allows moral realists to present a by-product defence of our capacity to track moral facts, as discussed in Chapter 2. Moral realists can also focus on Wielenberg's aforementioned point that we can only know a mental faculty is unreliable if we already know or assume the truth or falsity of some of the resulting beliefs. Overall, if one takes a moral realist's perspective when examining Joyce's EDA, many of the same general strategies used against Street's argument can also be applied to his argument.

With Joyce's debunking argument rejected, this leaves our moral beliefs with the same epistemic status that they held prior to evolutionary challenges: it is a philosophical question whether our particular moral beliefs are justified, not a matter that can be settled with evolutionary considerations. This thesis has not made any commitments regarding which particular moral beliefs are true, it has merely defended naturalistic moral realism against evolutionary challenges. Such a defence only needs to show that evolutionary theories cannot rule out the possibility of natural moral facts existing and of humans being able to identify such facts. Now that Street's and Joyce's arguments have been examined, the next chapter focuses on Joshua Greene's debunking argument which takes a rather different approach.

Chapter 4 – Joshua Greene's Debunking Argument

4.1 Overview of Greene's Argument

Joshua Greene presents a debunking argument that differs from Street's and Joyce's arguments in several ways. Whereas Street aimed to debunk moral realism and Joyce targeted the justification of moral beliefs, Greene's debunking target is a particular class of normative ethical theories. Specifically, his goal is to debunk deontological moral theories by developing an argument based on the results of neuroscientific experiments along with a certain theory of moral psychology (Greene 2008, 2014). The appeal to neuroscientific evidence is one of the most notable differences from the other discussed EDAs, as this is a case of drawing upon testable empirical claims rather than speculative evolutionary theories. Nevertheless, there is still an evolutionary component in Greene's argument, so it can also be considered an EDA.

Greene's debunking target is selective, as he maintains that consequentialist moral theories are safe from his argument. Despite the selective nature of his debunking target, it is still broad in the sense that there are a wide variety of deontological theories, each with their own complexities. However, for the sake of simplicity, the term 'deontology' will be used to describe this broad approach to moral philosophy. The remainder of this section summarises Greene's overall argument, with the later sections focusing on particular problems that should ultimately lead us to reject his selective debunking of deontology.

Since Greene's argument targets deontology while apparently leaving consequentialism unaffected, it is important to begin by defining these competing normative frameworks. Deontology and consequentialism are two different classes of normative moral theories that often come into conflict in particular moral cases: consequentialists hold that only consequences are of fundamental moral importance, whereas deontologists often allow or require actions that would not produce the best overall outcomes, due to various moral duties and prohibitions (Greene 2008, p. 37). In terms of moral rightness and wrongness, consequentialists judge actions purely in terms of their outcomes; depending on one's preferred variety of consequentialism, this can either be the actual, possible or expected outcomes. The consequentialist approach to ethics is exemplified by utilitarian theories, which judge actions in accordance with their promotion

of the greatest overall good; the definition of goodness depends on one's theory of value. For example, one might identify happiness as the type of value that should be maximised.

On the other hand, deontological theories are less concerned with consequences, instead defining moral rightness and wrongness in terms of an action's alignment with concepts such as duties and principles. Thus, deontology focuses more on the motives underlying an action, rather than the outcome. Different deontological theories can vary in terms of their lenience regarding exceptionally bad consequences that may arise from strictly following duties or prohibitions. Some examples of prohibitions are moral rules against intentionally harming others, lying, or using other people as a means to an end. Duties may include keeping promises, helping others, or perfecting one's own personal abilities, just to name a few.

Greene's argument is similar to Street's and Joyce's EDAs in the sense that it depends on claims about the nature of moral judgements. In these previously discussed arguments, one of the main points was that moral beliefs can apparently be 'explained away' by appealing to evolutionary theory. In Greene's selective debunking argument, he only intends to undermine deontological moral judgements, thus leaving consequentialist judgements unaffected. This is due to his view that the existence of competing moral theories can be explained by considering different types of psychological processes: "I believe that consequentialist and deontological views of philosophy are not so much philosophical inventions as they are philosophical manifestations of two dissociable psychological patterns, two different ways of moral thinking, that have been part of the human repertoire for thousands of years" (Greene 2008, pp. 37-38). The last part of this quote hints at the presence of evolutionary factors in Greene's argument, which will become clear when examining a particular moral dilemma later on. Since Greene's debunking argument depends on the application of a psychological theory to people's responses to a certain moral dilemma, it is important to first examine the details of this theory.

The theory of moral psychology that Greene endorses is the dual-process theory of moral judgements. The dual-process aspect of the theory states that the human brain operates in two distinct modes: on the one hand, there is the 'automatic' and intuitive type of mental processing that often involves emotional engagement. Since this type of processing occurs automatically without any conscious effort, we apparently rely on it most of the time (Greene 2014, p. 696). The other type of mental processing is the more effortful, conscious reasoning that we 'manually' engage in, rather than it simply being triggered by external factors such as our environment. This type of processing enables

highly complex and abstract reasoning, including moral reasoning. It also allows us to override our default 'automatic' processes in many cases; Greene gives the example of ignoring one's immediate craving for chocolate cake in order to be healthier in the long term (Greene 2014, pp. 696-697). Greene has sometimes referred to this second type of process as 'cognitive' processing (Greene 2008, p. 36). As such, this term will frequently be used throughout the chapter. According to Greene, the way the dual-process theory applies to moral judgements is straightforward: "[M]oral psychology looks much like the rest of judgment and decision making. Moral judgment is influenced by both automatic emotional responses (automatic settings) and controlled, conscious reasoning (manual mode)" (Greene 2014, p. 698).

While there may be debates regarding the details of the dual-process theory in general, the more relevant point of contention is Greene's further claim that different normative theories are generally linked to different mental processes. Greene describes this as the 'central tension' principle: "Characteristically deontological judgments are preferentially supported by automatic emotional responses, while characteristically consequentialist judgments are preferentially supported by conscious reasoning and allied processes of cognitive control" (Greene 2014, p. 699). In order for this claim to support a selective debunking argument against deontology, Greene must be able to demonstrate two points. The first is the empirical matter of whether there is in fact such a connection between particular types of moral judgements and particular mental processes. The second point is the theoretical issue of whether emotional processing should be considered inferior in some way to cognitive processing, at least in the context of moral judgements. Greene requires an affirmative answer in order to argue that deontological theories are primarily based on an inferior type of moral thinking (emotion-based moral judgements), in contrast with consequentialism being based on cognitive moral judgements. Each of these two main points will now be examined in turn.

To support the first point, Greene draws upon the results of neuroscientific experiments that he and his colleagues conducted (Greene et al. 2001). Their study examined the connections between mental processes and moral judgements. The participants were regular people (non-philosophers) who were tasked with reading various hypothetical moral dilemmas and then indicating whether they considered each proposed action to be "appropriate" or "inappropriate" (Greene et al. 2001, p. 2106). This research was notable for its use of functional magnetic resonance imaging (fMRI), as this exemplified an empirical approach to the traditionally theoretical field of philosophy. The fMRI was used to examine the neural activity in participants' brains during their

contemplation of the moral dilemmas. Participants' levels of emotional engagement were measured through certain patterns of neural activity, based on prior studies of the neural correlates of emotion. A greater degree of brain activity in certain 'emotional' brain regions was interpreted as evidence of more emotional engagement (Greene et al. 2001, pp. 2106-2107).

The investigation involved participants responding to 60 moral dilemmas while their brains were scanned. These 60 cases were divided into 'personal' and 'impersonal' categories, depending on how "up close and personal" the potential moral violations were (Greene et al. 2001, p. 2106). Some examples from the 'personal' category included the decision of whether to steal one person's organs to help five other people, and whether to throw people out of a sinking lifeboat. Compared to the other category, these cases were expected to produce more emotional engagement in the participants. The 'impersonal' category included situations such as deciding whether to keep money obtained from a lost wallet, or whether to support a policy that would be expected to cause more deaths than its alternatives. Among the many hypothetical moral dilemmas were two well-known 'trolley' cases: the switch case and the footbridge case. Greene describes the switch case in the following terms:

> A runaway trolley is headed for five people who will be killed if it proceeds on its present course. The only way to save them is to hit a switch that will turn the trolley onto an alternate set of tracks where it will kill one person instead of five. Ought you to turn the trolley in order to save five people at the expense of one? Most people say yes (Greene et al. 2001, p. 2105).

This was included in the 'impersonal' category since it involves using a switch to indirectly cause death from a distance. Notably, the common intuitive 'yes' answer can be considered a consequentialist judgement. On the other side, the 'personal' category included the similar footbridge case:

> [A] trolley threatens to kill five people. You are standing next to a large stranger on a footbridge that spans the tracks, in between the oncoming trolley and the five people. In this scenario, the only way to save the five people is to push this stranger off the bridge, onto the tracks below. He will die if you do this, but his body will stop the trolley from reaching the others. Ought you to save the five others by pushing this stranger to his death? Most people say no (Greene et al. 2001, p. 2105).

Unlike the switch case, the most common intuitive response (a negative response in this case) aligns more with deontology than consequentialism. These two cases are notable because when considered together, they lead to the 'trolley problem' that has caused a

great deal of debate among philosophers. The problem lies with attempts to accommodate the different intuitive responses to both of the cases; these conflicting judgements occur despite the only notable difference being the method of sacrificing one person to save five. Neither consequentialists nor deontologists have been able to accommodate both common intuitions while maintaining a consistent moral theory. The trolley problem has been important in normative ethics since it appears to capture the main point of conflict between deontological and consequentialist reasoning. As Greene puts it, "We philosophers have puzzled over trolley dilemmas for decades because they capture a central—if not *the* central—tension in normative ethics, and the myriad scientific results these dilemmas have generated implies that they tap something deep—revealing the hidden tectonics of the moral mind" (Greene 2014, p. 705).

Many philosophers have attempted to find a moral principle that would be consistent with both of the intuitive responses (Berker 2009, p. 297). One of the common assumptions underlying the trolley problem is the notion that there is some morally significant difference between the two cases. In particular, some philosophers think that the problem demonstrates a moral difference between different methods of inflicting harm. This idea is represented by the 'doctrine of double effect', which applies as follows: in the footbridge case, pushing the man off the bridge is impermissible since the harm is intentional; conversely, killing the one person in the switch case is permissible as the harm is merely foreseen, not intended (Kumar & Campbell 2012, p. 312). This view is just one of many contentious approaches to the trolley problem.

Since the traditional philosophical debates have not led to any widely accepted solution to this problem, Greene and his colleagues hoped to make progress by taking their empirical approach, thus investigating why these two cases produce different intuitive judgements. Specifically, they considered whether the cause of the different judgements might involve the footbridge case triggering people's emotions in a way that does not occur in the switch case. When considering the 60 moral dilemmas that participants responded to, Greene's research team found that variations in emotional engagement during moral judgements correlated with a certain factor that differs between the trolley and footbridge cases. Due to methodological constraints, the relevant factor was represented as a clear distinction between 'personal' and 'impersonal' actions (Greene et al. 2001, p. 2107). The main finding was that there was greater emotional engagement in personal dilemmas; this was the category that included the footbridge case.

The researchers acknowledged that this is merely a psychological finding, such that these experimental results cannot directly inform us about what is morally right or wrong.

However, these studies raised the question of how a better understanding of the psychological processes underlying moral judgements could potentially affect the philosophical status of these judgements (Greene et al. 2001, p. 2107). This extends to the idea that if certain classes of normative theories (such as deontological theories) are mainly associated with moral judgements based on one particular mental process, then this point could be used to assess the plausibility of these normative theories. Greene thus describes the connection between the empirical study of moral judgements and the philosophical study of normative theories along these lines: "Science can advance ethics by revealing the hidden inner workings of our moral judgments, especially the ones we make intuitively. Once those inner workings are revealed we may have less confidence in some of our judgments and the ethical theories that are (explicitly or implicitly) based on them" (Greene 2014, pp. 695-696).

Since Greene's experimental results are supposed to support his 'central tension' principle that associates deontological and consequentialist judgements with emotions and cognition respectively, it is important to consider how the switch and footbridge cases might represent the differences between these competing approaches to ethics. First, Greene clarifies that the most intuitive responses to these cases are 'characteristically' consequentialist and deontological judgements respectively. What this means is that one can make a judgement that is characteristic of a certain theory without knowing anything about that theory; all that matters is that one's judgement happens to align with the theory. A similar point applies to people whose intuitive responses conflict with the normative theory that they explicitly prefer. For instance, one might have a 'characteristically' deontological intuition despite identifying as a consequentialist.

Greene seeks to explain the tendencies towards these characteristic moral judgements in the switch and footbridge cases, but he does not follow the assumption that the most common intuitive judgements are necessarily correct or reasonable. Thus, rather than searching for a principle that could justify the different responses to these very similar cases, Greene presents what he considers to be a 'purely descriptive' partial solution to the trolley problem, which aims to explain the cause of the different judgements (Greene 2008, p. 42). Greene's proposed explanation draws upon the aforementioned studies that identified the switch case as involving 'impersonal' harm while the footbridge case involves 'personal' harm. Here it is important to recall the details: the switch case involves an indirect method of killing someone by using a switch to remotely redirect a trolley towards them, while the footbridge case involves directly pushing a person off a footbridge to their death. Greene considers this difference to be significant purely from an

evolutionary perspective, rather than from a moral perspective: "Given that personal violence is evolutionarily ancient, predating our recently evolved human capacities for complex abstract reasoning, it should come as no surprise if we have innate responses to personal violence that are powerful but rather primitive" (Greene 2008, p. 43). These primitive emotional reactions are often described as 'alarmlike' emotions throughout Greene's work, as they occur suddenly and cannot simply be ignored.

As for cases of impersonal harm, Greene asserts that they would not trigger our evolved emotional alarm responses, thus allowing us to carefully contemplate these cases with our cognitive capacities (Greene 2008, p. 43). This is because early humans presumably lacked the means to intentionally cause long-distance indirect harm; devices such as switches had not yet been invented. Greene notes that his earlier experimental results (Greene et al. 2004; Greene et al. 2001) confirm predictions that follow from the evolutionary considerations about personal and impersonal harm:

> The contemplation of personal moral dilemmas like the footbridge case should produce increased neural activity in brain regions associated with emotional response and social cognition, while the contemplation of impersonal moral dilemmas like the trolley [switch] case should produce relatively greater activity in brain regions associated with 'higher cognition.' This is exactly what was observed (Greene 2008, p. 43).

Furthermore, Greene states that his experiments also confirmed predictions about response times: "Trials in which the subject judged in favor of personal moral violations took significantly longer than trials in which the subject judged against them, but there was no comparable reaction time effect observed in response to impersonal moral violations" (Greene 2008, p. 44). Since emotional responses are quick and automatic, the slower cognitive responses would require some time to override this initial reaction when a person ultimately decides to approve of personal harm in cases that involve this type of harm. For example, the consequentialist judgement that one should push the man off the footbridge would have to override an emotion-based deontological judgement that opposes this instance of personal harm.

Greene provides further support for the evolutionary explanation of certain types of moral judgements responding to certain factors in moral dilemmas, by examining another case in which most people perceive an apparent moral distinction. The case he refers to is the 'drowning child' scenario from Peter Singer:

> [I]f one notices a small child drowning in a shallow pond, one is morally obliged to wade in and save that child, even if it means muddying one's clothes. … Why, Singer

asks, do we have a strict obligation to save a nearby drowning child but no comparable obligation to save faraway sick and starving children through charitable donations to organizations like Oxfam? (Greene 2008, pp. 46-47)

Greene observes that one's interaction with the drowning child is up close and personal, as with the types of situations that were possible for early humans to encounter. In contrast, helping distant people from across the world would not have been possible in our evolutionary past, which may explain why we have not evolved with any strong emotional inclinations to help others in such impersonal cases (Greene 2008, p. 47). Greene thinks that this explains why most people (unlike Singer) do not have a consequentialist intuition in this situation, thus spending money on luxuries rather than promoting better overall consequences by relieving the suffering of distant strangers. Along with the trolley cases, this exemplifies how evolutionary factors may explain tendencies towards or against certain types of moral judgements.

In order to demonstrate how evolutionary considerations link to Greene's debunking of deontology, there are some further points regarding emotion-based judgements in response to personal moral dilemmas. Greene speculates that personal harm triggers automatic emotional responses that evolved in our vastly different ancestral environment. Moral judgements based on these emotional responses may have been an effective natural solution to the types of social cooperation challenges that early humans encountered (Greene 2008, p. 59). Here one may recall Joyce's concept of the moral sense, which offered a similar account of moral judgements serving the function of promoting prosocial behaviour. Thus, it is not difficult to conceive of negative emotional responses to personal harm having evolutionary benefits, by discouraging unnecessary violence among early humans. In Greene's argument, the idea is that the quick and automatic nature of emotional responses makes them well suited for their original evolutionary function. In contrast, cognitive processing is slower and more mentally demanding, so it presumably would have been less reliable in the circumstances of our evolutionary past (Greene 2008, pp. 59-60). The cognitive capacities that are uniquely developed in humans are presumed to have evolved relatively recently in comparison to our emotional capacities.

Although the living circumstances for humans have changed significantly over the time since our emotional capacities evolved, Greene argues that our primitive emotional responses have now been codified by philosophers into a certain class of normative moral theories: specifically, deontological theories. Greene thus puts forth the controversial idea that deontological philosophy is a rationalisation of our evolved emotion-based moral judgements. One point that he raises in support of this bold claim is that humans have a

strong tendency to attempt to explain and justify their behaviour. He refers to psychological studies that have found that people invent seemingly plausible stories to explain their own behaviour, even when they do not know why they are actually behaving in a certain way (Greene 2008, pp. 60-61). However, even if we grant that humans have this tendency, this would not be sufficient to 'explain away' deontological philosophy, which represents a significant portion of normative ethics. Greene thus attempts to provide more support for his point about deontological rationalisations by focusing again on the footbridge case. Pushing the man off the bridge involves direct personal harm, which is a type of harm that was possible to inflict in our evolutionary past. Due to our evolved alarmlike emotional reactions to personal harm, this action seems intuitively wrong to most people. According to Greene, this sense of wrongness is perfectly captured by deontology: "[W]hat better way to express that feeling of non-negotiable absolute wrongness than via the most central of deontological concepts, the concept of a right: You can't push him to his death because that would be a violation of his rights. Likewise, you can't let that baby drown because you have a duty to save it" (Greene 2008, p. 63). Although these are particular cases, Greene takes this to support his view that general deontological concepts such as 'rights' and 'duties' can be explained away as a rationalisation of our emotional responses.

Greene recognises that his selective debunking of deontology would be incomplete without an explanation of why consequentialism is immune to his claims about rationalisation. He maintains that consequentialist theories do not merely rationalise intuitive responses to moral issues: although such theories may ultimately be based on consequentialist intuitions, these theories do not aim to match our moral intuitions on a case-by-case basis. Rather, consequentialists are supposed to accept any unintuitive prescriptions of their theory. According to Greene, consequentialists can determine their response to a moral problem without even considering their intuitions about it: "An act consequentialist can know what she thinks about a case without knowing anything other than the answer to this question: Which choice produces better consequences?" (Greene 2014, p. 724). Thus, unlike deontologists, consequentialists apparently cannot be accused of simply rationalising or 'chasing' intuitions (Greene 2014, p. 724).

Greene also supports his selective debunking argument by claiming that consequentialism is superior to deontology in terms of recognising morally relevant factors in moral problems. To make this point, he first refers back to the purported link between different normative theories and different mental processes:

I argued earlier that there is a natural mapping between the content of deontological philosophy and the functional properties of alarmlike emotions. Likewise, I believe that there is a natural mapping between the content of consequentialist philosophy and the functional properties of 'cognitive' processes. Indeed, I believe that consequentialism is inherently 'cognitive,' that it couldn't be implemented any other way (Greene 2008, p. 63).

This forms the basis of his claim that consequentialist reasoning makes use of our cognitive capacities to properly account for relevant factors in a moral dilemma; the consequentialist ideal is to impartially weigh up every competing moral concern (Greene 2008, p. 64). This is supposed to stand in contrast with a notion of deontologists excessively focusing on certain factors that our emotional responses draw our attention to.

Greene substantiates this picture of superior consequentialist reasoning by asserting that cognitive processes are behaviourally neutral, unlike emotional processes. By this he means that cognitive processing does not automatically trigger certain behaviours; cognition is flexible enough to allow for a wide variety of responses to a situation (Greene 2008, p. 40). This flexibility gives cognitive processing a theoretical advantage over emotional processing in terms of decision-making based on factors in moral dilemmas. Greene gives the following example of a morally relevant consideration in the footbridge case:

'Is it okay to push the guy off the bridge if he's about to cure cancer?' … Deontologists can dismiss these sorts of complicated, situation-specific questions, but consequentialists cannot, which is why, I argue, that consequentialism is inescapably 'cognitive' (Greene 2008, p. 64).

In this way, Greene's view suggests that deontologists would generally stick with their intuitions about 'rights' and 'duties' even though this may lead them to disregard various morally relevant factors. Greene recognises that his normative assumptions about moral relevance are made independently of the empirical findings from his experiments. As such, he proposes the following relation between these aspects of his argument:

Such experiments identify factors to which our moral judgments are sensitive. This information may be combined with independent normative assumptions concerning the kinds of things to which our judgments ought to be sensitive. This combination can lead us to new, substantive moral conclusions. In other words, scientific information can allow us to trade in difficult 'ought' questions for easier 'ought' questions, and thus advance ethics (Greene 2014, p. 711).

Greene specifically relies on the normative assumption that the 'personalness' or proximity of inflicted harm is morally irrelevant. To clarify, the claim is not that harm itself is morally irrelevant, but rather that the particular method of inflicting harm is irrelevant when the outcome would be the same from another method. This would entail that our evolved tendency for emotion-based disapproval of personal harm must be considered a distorting influence on our moral judgements, since these emotional responses are not also triggered by considering impersonal harm. If we grant Greene's empirical claims about the association of emotional processing with deontology, this then leads to his conclusion that deontological philosophy is debunked, as it merely rationalises our misguided emotion-based judgements.

Greene is aware that the issue of different responses to personal versus impersonal moral dilemmas does not provide sufficient grounds for debunking an entire class of normative theories. He therefore supplements his debunking argument with some further points, which concern the relative suitability of different mental processes in different situations. We may recall that Greene draws upon the dual-process theory of moral judgements, which distinguishes between 'automatic' emotional processing and 'manual' cognitive processing. He clarifies that he does not consider emotion-based moral judgements to be worse than cognitive moral judgements by default; rather, the usefulness of each type of mental process depends on the context. The general idea is that emotion-based processing is efficient but inflexible, while cognitive processing is less efficient but more flexible (Greene 2014, p. 714).

Using this empirical claim as a foundation, Greene makes the normative claim that manual cognitive reasoning is therefore better suited to assessing complex moral problems, due to its greater flexibility. Specifically, he suggests that we should rely more on manual cognitive reasoning than automatic emotion-based responses when faced with 'unfamiliar' moral problems (Greene 2014, p. 715). This refers to situations that would not have been encountered by early humans, so it is an evolutionary consideration. Greene provides some examples: "[M]oral problems that arise from recent cultural developments, most notably the rise of modern technology and the intersection of disparate cultures, are especially likely to be unfamiliar. Think climate change, global terrorism, global poverty, bioethics, etc" (Greene 2014, p. 716). It seems unlikely that natural selection would have favoured a capacity for automatic emotional responses that happen to produce reliable moral judgements regarding these modern issues. This appears to be similar to the idea in Street's argument that evolutionary influences on our moral judgements may have no relation to purported moral facts. However, Greene's argument only applies to moral judgements

based on emotional processing, as he argues that cognitive processing is flexible enough to properly account for morally relevant factors in unfamiliar moral problems.

To emphasise how these claims favour consequentialism over deontology, Greene applies what he describes as a 'no cognitive miracles' principle: we should rely more on manual cognitive reasoning when faced with unfamiliar moral problems (Greene 2014, p. 715). The idea is that it would be a cognitive miracle if emotional processing allowed us to properly account for morally relevant factors in these cases. This is based on a further claim that Greene makes: that only three mechanisms are known to provide our automatic mental processes with the information required for them to function reliably. These mechanisms are cultural transmission, genetic transmission, and learning from personal experiences (Greene 2014, pp. 714-715). Since we lack a sufficient degree of prior experience with unfamiliar moral problems, this rules out these mechanisms for reliability. Thus, consequentialism should apparently be favoured over deontology at least in unfamiliar cases, due to consequentialism's association with the more flexible cognitive processes that allow us to properly account for morally relevant factors. This claim provides the final component of Greene's selective debunking of deontology, as it demonstrates some apparent advantages of consequentialist reasoning; this can be combined with the various negative points about deontology that have been noted. The negative points mainly concerned the apparent rationalisation of emotional responses, such as responses to morally irrelevant factors.

The discussion so far has covered the main points in Greene's argument. The remaining sections turn the focus to various problems with his overall argument. The main problem covered in Section 4.2 is the difficulty of selectively debunking particular theories without collapsing into a broader debunking argument. Section 4.3 then focuses on the issue of Greene's reliance on undefended normative assumptions, such as the notion of morally relevant or irrelevant factors. These points will lead into the conclusion that Greene's EDA must be rejected along with the arguments from Street and Joyce.

4.2 Moral Judgements and Selective Debunking

One of the problematic aspects of Greene's argument is the selectiveness of the debunking target: he attempts to undermine deontology while maintaining that consequentialism is safe from his debunking claims. Whereas 'global' EDAs may face the challenge of justifying ambitious debunking claims about moral judgements in general, one difficulty

for selective EDAs is to avoid collapsing into a global argument. Selim Berker considers the selective aspect of Greene's argument, noting that Greene has not adequately defended his consequentialist intuitions from his evolutionary debunking of deontological intuitions: "Presumably consequentialist intuitions are just as much a product of evolution—whether directly or indirectly—as deontological intuitions are, so an appeal to evolutionary history gives us no reason to privilege consequentialist intuitions over deontological ones" (Berker 2009, p. 319). For example, it seems plausible that the consequentialist intuition that one should maximise welfare in one's society would have had evolutionary benefits. This type of intuition could have encouraged our ancestors to focus on the survival of the majority of their social group, which would clearly provide survival advantages over other groups that refrain from sacrificing a few individuals for the sake of the majority. Although Greene combines his evolutionary claims against deontology with his empirical findings, the evolutionary story about personal harm is just as speculative as any evolutionary account of consequentialist intuitions. After all, experiments cannot directly tell us why certain mental capacities and behavioural tendencies evolved.

One of Greene's reasons for favouring manual cognitive reasoning (and thus the associated consequentialism) was that it is supposedly more likely to be reliable when faced with unfamiliar moral problems. However, Victor Kumar and Richmond Campbell oppose the idea that problems like the trolley cases are unfamiliar in evolutionary terms (Kumar & Campbell 2012). Although such cases involve modern inventions such as switches and trolleys, the necessity of harming some people in order to save or benefit a greater number of people is surely a very old type of dilemma, and thus likely to have been relevant during human evolution. As such, they assert that the burden is on Greene to demonstrate that there is something new and unfamiliar about such cases (Kumar & Campbell 2012, p. 321). Although it seems obvious that indirectly causing death via a vehicle or other similar means could not have happened during early human evolution, any further claims regarding our evolutionary familiarity with these types of problems would be purely speculative. Since we lack the ability to empirically test evolutionary hypotheses about the types of situations faced by early humans, we cannot simply assume that Greene's evolutionary speculations are accurate. As such, claims about familiarity do not adequately support the selective debunking of deontology. Debunking an entire class of normative theories is a very ambitious project, so relying heavily on speculations is not sufficient for this purpose.

Furthermore, even if we were to grant that situations such as the trolley cases involve unfamiliar factors, this point would only support Greene's 'no cognitive miracles'

principle that is limited to favouring consequentialism in unfamiliar cases. Favouring consequentialism in these cases would not be the same as debunking deontology, as it still leaves open the possibility of deontological reasoning being reliable for familiar types of moral problems. Greene attempted to debunk deontology in familiar cases by arguing that it is merely a rationalisation of our evolved emotion-based responses, some of which are triggered by morally irrelevant factors. The notion of moral relevance will be examined in Section 4.3; the current focus is on the rationalisation claim. We cannot simply assume that any deontological judgement involving familiar factors such as personal harm is merely a rationalisation of evolved emotional responses. To recall one of the main points from earlier in the thesis, the relatively recent evolution of advanced rational capacities has enabled philosophers to think about moral principles in the abstract. This type of complex reasoning surely involves cognitive processing on Greene's dual-process model. These points can be used against the notion that deontology merely rationalises emotion-based intuitions about particular moral issues: even when we consider the types of moral issues that early humans could have encountered, the abstract moral concepts that we now apply to such cases have been developed in our modern living circumstances, using our more recently evolved rational capacities. Therefore, it is plausible to maintain that deontological concepts that apply to familiar moral problems are not merely rationalisations of evolved emotional responses.

On the issue of abstract moral concepts and rationalisation, it is worth examining some of Greene's points against Kantian deontology in particular, which have not yet been discussed in this chapter. Greene takes Kant to be a prime example of a deontologist. He notes that many Kantians focus on constructing seemingly rational moral principles, rather than attempting to identify moral principles that may be true in themselves. Kantian moral philosophy thus appears to represent an anthropocentric approach to morality, as it attempts to rationalise intuitions rather than supporting principles that could challenge them (Greene 2008, p. 75). In response to Greene's characterisation of Kantian deontology, Kantians could argue that it is difficult to conceive of their theory as a mere rationalisation of emotion-based intuitions. Along these lines, Richard Dean notes that Kant's moral theory is based on abstract intuitions about the nature of morality, not intuitions about particular moral issues (Dean 2010, p. 52). He also points out that Kant in his *Metaphysics of Morals* explicitly disapproves of constructing moral theories from the starting point of intuitions about particular cases (Dean 2010, p. 52).

However, from Greene's perspective the problem with Kantian deontology is not so much its starting point, but rather its end point: Greene's view targets any normative

theory that is primarily concerned with justifying emotion-based intuitive responses to moral issues, even if it purports to begin from a rational conception of morality itself. Nevertheless, Greene cannot simply assume that the main function of deontological principles is to rationalise intuitive responses to particular issues. Presumably, any sufficiently general moral principles would be applicable to a variety of both intuitive and counterintuitive judgements, and such principles can remain relevant even as dominant intuitions change with regard to specific moral issues.

Furthermore, one could challenge Greene's claims about deontological rationalisation of intuitions by undermining the 'central tension' principle that follows from his dual-process theory. Thus, if it can be demonstrated that deontological judgements also involve a significant degree of cognitive processing, then this would weaken Greene's argument since he mainly associates intuitions with emotional processing (Greene 2014, p. 696). To exemplify how one could criticise Greene's argument along these lines, we may consider how Guy Kahane and Nicholas Shackel have addressed this aspect of the argument (Kahane & Shackel 2010). They point out that even if one type of mental process (either cognitive or emotional) were more closely associated with a particular class of moral judgements, the same process would surely also be involved in contrary moral judgements (Kahane & Shackel 2010, p. 579). For example, it is highly unlikely that subjects who make a characteristically deontological judgement in footbridge cases are ignoring facts about the total expected harm to all individuals. A more reasonable interpretation is that most subjects simply judge that pushing a man to his death is worse than saving more lives through this action; this judgement clearly involves weighing up the total harm. As Kahane has noted, "Deontologists are not numerically challenged. It is near certain that all subjects considering Footbridge and similar dilemmas make this simple calculation, whether or not they reach a 'utilitarian' conclusion." (Kahane 2012, pp. 529-530).

Thus, when attempting to determine what kind of neural processing leads to deontological judgements, what really matters is the process that causes deontological considerations to be prioritised over consequentialist considerations. Kahane and Shackel consider this to be important because this processing could turn out to be a significant component of deontological reasoning, such as weighing up conflicting duties or deciding whether overall expected harm outweighs a deontological prohibition. As they summarise, "[S]uch deontological reasoning might underlie many common decisions merely conforming to utilitarianism, a point utterly obscured when we describe such choices as 'utilitarian'. Hence, what are reported as the uniquely cognitive processes underlying

utilitarian judgment might, for all we know, reflect deontological reasoning" (Kahane & Shackel 2010, p. 579).

Kahane and his colleagues have also provided empirical support for criticisms of Greene's dual-process theory and the associated 'central tension' principle (Kahane et al. 2012). Their aim was to investigate what type of connection really exists between intuitive moral judgements and different normative theories. Using fMRI brain scanning as in Greene's experiments, Kahane's study presented subjects with a range of moral dilemmas including the usual switch and footbridge cases. Notably, they included some previously unstudied moral dilemmas in which a 'deontological' judgement would be counterintuitive, such as following a moral duty against lying even though the refusal to lie would result in considerable harm (Kahane 2012, p. 538). To see how this is significant, first we may recall that Greene's studies typically compared counterintuitive utilitarian judgements—such as pushing the man off the footbridge—with intuitive deontological judgements, such as refusing to push the man. Greene interpreted the differences in neural activity (emotional versus cognitive processing) as being characteristic of the difference between utilitarian and deontological judgements. However, Kahane presents an alternative interpretation: the differences in processing may actually be linked to how intuitive certain moral judgements are, rather than their specific content. He describes the implications of this different interpretation: "If this is correct, then, again, the apparent tie between process and content is really just an artefact of the kinds of scenarios that researchers have studied, reflecting nothing deep about utilitarian and deontological judgments" (Kahane 2012, p. 536). This was supported by results from Kahane's study, which provided much stronger evidence for his interpretation than Greene's. In particular, it was found that 'utilitarian' judgements—such as approving of pushing the man off the footbridge—involved very similar neural activity to 'ultra-deontological' judgements such as refusing to lie even to prevent significant harm. This demonstrated that moral judgements with vastly different content could be associated with similar neural processes (Kahane 2012, p. 538). Ultimately, Kahane concludes with a more modest dual-process model of moral judgements as compared to Greene's model: intuitive moral judgements are more associated with automatic mental processes, while counterintuitive moral judgements involve more controlled cognitive processing (Kahane 2012, pp. 539-540).

Since Kahane's points undermine Greene's association of deontology with emotion-based intuitive judgements, this provides further evidence that deontological philosophy does not merely rationalise automatic emotional responses. Nevertheless, even if both deontological and consequentialist principles can support a wide variety of intuitive

and counterintuitive moral judgements, the actual plausibility of these principles is a separate philosophical issue. As such, it remains unclear whether questions concerning the plausibility of normative theories themselves can be guided by empirical findings.

The discussion up to this point has focused on how deontologists may resist Greene's selective debunking claims that were based on the notion of evolved intuitive responses to moral issues. However, it is also worth considering how consequentialists may defend some of their own intuitions that appear to have clear evolutionary explanations. As noted earlier, Greene recognises that such a defence is necessary in order to avoid his evolutionary claims causing a collapse into a broader (potentially global) debunking argument. His claim was that consequentialism is not a mere rationalisation since it does not attempt to follow moral intuitions on a case-by-case basis; instead, consequentialists must accept any unintuitive prescriptions that follow from their theory. However, apart from some notable examples such as Peter Singer, it is not apparent that many consequentialists are always willing to defend highly unintuitive moral prescriptions for the sake of maintaining consistency in their reasoning.

Regardless of whether consequentialism is supposed to ignore intuitions in theory, consequentialist prescriptions can at least coincidentally align with many common intuitions that happen to be amenable to evolutionary explanations. Although consequentialism and deontology may each support a different subset of all the widespread intuitions that people have, the important point is that at least some consequentialist claims can align with such intuitions. One example of a ubiquitous moral intuition held by philosophers and non-philosophers alike is the belief that we are obliged to care for our own children more than the children of others; the widespread acceptance of this view is noted by Street (2006, p. 115). For the sake of argument, it will be assumed that beliefs such as this one can indeed be attributed to evolutionary factors. Consequentialists might attempt to defend such inflexible evolved intuitions by asserting that our inability to overcome them entails a practical reason to maintain them. For instance, it could be noted that more overall well-being (a better consequence) would be promoted by maintaining such beliefs, rather than hopelessly attempting to resist them. In this way, consequentialists could argue that these types of practical considerations override any concerns about possible evolutionary explanations of certain intuitions. This reasoning would also align with Greene's view that consequentialism is based on highly flexible cognitive processes, as this flexibility can allow us to rationally consider moral issues regardless of evolutionary influences.

However, this still leaves the question of whether this line of reasoning would only selectively support consequentialism and not deontology. In order to rule out the possibility of deontologists defending their theory in a similar manner, Greene would have to rely on his central tension principle along with the idea that any cognitive deontological reasoning is a rationalisation of emotion-based intuitions. Unfortunately for Greene, this aspect of his argument has already been argued against, thus leaving deontologists in a similar situation as consequentialists. Thus, defenders of either type of theory can make similar points about their theory's general principles being unaffected by evolutionary considerations.

Greene's apparent underestimation of the versatility of deontological principles is also noted by Richard Dean, who argues that Greene's neuroscientific findings fail to establish the conclusion that consequentialism should be favoured over deontology (Dean 2010). He observes that deontological philosophy encompasses many moral duties, rights and obligations that do not appear to be based on evolved automatic emotional responses to certain factors in moral dilemmas. For example, we would not expect obligations relating to lying and promise-keeping to be based on alarmlike emotional responses that may have carried over from our evolutionary past (Dean 2010, p. 49). These are just some examples of moral issues that do not necessarily have to involve immediate danger or any other sense of urgency, thus allowing enough time for cognitive processing to influence one's eventual moral judgement.

Furthermore, Dean notes that even if we were to accept Greene's dual-process model as it applies to the trolley problem, the neuroscientific experiments have not shown that this model can account for moral judgements in general (Dean 2010, p. 50). Due to Greene's primary focus on the trolley problem and related personal and impersonal dilemmas, the scope of his studies is too narrow to account for the complexity of deontological philosophy, which could be highlighted by considering a wider variety of moral issues. As Dean observes, Greene is well aware that the brain areas he identifies with cognitive processing are involved in much more than just quantitative calculations when faced with moral dilemmas. Since these brain regions are also implicated in tasks such as general abstract reasoning, deductive reasoning and planning, it should not be surprising that cognitive processes would also have significant involvement in non-consequentialist moral judgements (Dean 2010, p. 52).

Selim Berker also observes that deontology is more multifaceted than its depiction in Greene's examples of 'characteristically' deontological judgements (Berker 2009). He points out that the distinction between deontological and consequentialist judgements is

surely based on more than just personal versus impersonal means of inflicting harm (Berker 2009, p. 311). As such, Greene's studies seem to involve an overly narrow conception of what is distinct about consequentialist and deontological theories, which may also be attributed to his excessive focus on the trolley problem and related dilemmas.

Guy Kahane has also criticised Greene's excessive focus on the trolley problem and similar cases: he notes that this is problematic for Greene's dual-process theory of moral judgements, as it is supposed to be applicable to these judgements in general, not just those about trolley problems and similar scenarios (Kahane 2012, p. 521). However, even if a more general dual-process model turned out to be accurate, this would still leave an explanatory gap between the basis of different moral judgements and the nature of particular normative theories. So far we have little more than speculation regarding the connections between moral judgements and moral theories (Kahane 2012, p. 522). This relates to a previously discussed criticism of Street's and Joyce's evolutionary claims: for appeals to science to have any chance of helping us resolve philosophical disputes, we require strong evidence rather than a heavy reliance on speculation. Speculation is more in line with traditional philosophical methodology, which conflicts with Greene's intention of advancing philosophical debates by utilising scientific methods.

One final methodological issue with Greene's studies is that presenting subjects with descriptions of hypothetical situations is not ideal for testing mental processes involved in moral judgements, particularly emotional processing. Normative ethical theories are supposed to be directly applicable to real situations, not just to imagined situations such as thought experiments. In the neuroscientific experiments that have been discussed, neural activity was examined while subjects merely read about and contemplated moral dilemmas in which their own involvement was merely imagined. As such, the findings may not necessarily indicate what kind of mental processing would occur when people are actually involved in such situations. It seems reasonable to expect that there would be more emotional engagement if subjects actually found themselves in the described situations, rather than just being informed about them. It is also possible that the required use of imagination when considering these moral dilemmas may affect the results, depending on what kind of mental processing is involved in imagination. These points raise broader questions about the use of thought experiments in philosophy, which are unfortunately beyond the current scope.

In summary, this section has focused on various problems with the selective aspects of Greene's debunking of deontology, along with issues relating to moral judgements and

the dual-process theory. Now that these points have been covered, the next section examines the issue of Greene's reliance on certain normative assumptions.

4.3 Normative Assumptions and Moral Relevance

So far, most of the discussed criticisms of Greene's argument have related to empirical issues such as the association between certain mental processes and certain types of moral judgements, as well as Greene's rationalisation claims against deontology. However, the discussion has yet to directly address whether emotion-based moral judgements are in fact less reliable or in any way inferior to cognition-based moral judgements. The notion of 'reliability' is normative in the sense that its definition depends on what one considers to be a desirable or appropriate outcome of a certain process. In the case of moral judgements, reliability may refer to the frequency with which these judgements align with moral truths, but this would obviously depend on the assumption of particular moral facts.

Greene's assumptions about reliability are criticised by Kumar and Campbell, who point to his unsubstantiated claims about the comparative reliability of emotional versus cognitive moral judgements (Kumar & Campbell 2012). They observe that Greene provides three main reasons for disfavouring moral judgements based on the 'automatic' side of the dual-process theory: this automatic type of processing is emotion-based, it is an evolutionary adaptation, and it is relatively simple and inflexible compared to 'manual' cognitive processing. The problem is that Greene has not adequately explained how each of these claims is supposed to undermine the reliability of emotion-based moral judgements (Kumar & Campbell 2012, pp. 319-320). Before addressing Greene's three main points, it is important to recall that he does not think emotional processing is inherently worse than cognitive processing (Greene 2014, p. 714). Rather, his view is that manual cognitive reasoning is at least better suited to dealing with unfamiliar moral problems.

Regarding the first of Greene's three points, he has not presented a compelling argument that the emotional component of automatic processing makes it worse than cognitive processing. He seems to rely on other questionable assumptions to support this view, which is evident in his discussion of the footbridge case. In this case, he suggested that our alarmlike emotional responses that underlie deontological judgements about this dilemma are responding to a morally irrelevant factor: the 'personalness' or proximity involved in pushing the man off the bridge. His conclusion that emotion-based

deontological judgements are unreliable thus depends on this particular assumption of moral irrelevance. While this assumption could turn out to be plausible, it needs to be adequately defended even if we grant Greene's point that our responses to such factors are primarily emotion-based. Berker focuses on this aspect of Greene's argument, particularly how the notion of moral relevance is supposed to fit in with the overall debunking of deontology (Berker 2009). Berker describes this component of Greene's overall debunking argument as the 'argument from morally irrelevant factors', which he interprets as follows:

P1. The emotional processing that gives rise to deontological intuitions responds to factors that make a dilemma personal rather than impersonal.

P2. The factors that make a dilemma personal rather than impersonal are morally irrelevant.

C1. So, the emotional processing that gives rise to deontological intuitions responds to factors that are morally irrelevant.

C2. So, deontological intuitions, unlike consequentialist intuitions, do not have any genuine normative force (Berker 2009, p. 321).

Greene recognises that the notion of moral relevance cannot be directly derived from empirical facts; it is simply a matter that must be assumed. Berker thus considers the second premise to be a positive aspect of the argument, as it avoids any controversial leaps from purely empirical claims to substantive normative claims (Berker 2009, p. 322). However, in Berker's view, this reliance on independent assumptions also highlights the normative insignificance of neuroscience (Berker 2009, p. 325). If this empirical component of Greene's argument is indeed irrelevant, then the argument may have no significant advantage over more traditional philosophical arguments against deontological theories.

Berker's point about normative insignificance still seems to apply to Greene's updated, more modest view of the significance of neuroscientific studies, which he presented subsequent to Berker's article and other criticisms. As we may recall, Greene simply claimed that scientific information can allow us to exchange difficult 'ought' questions for somewhat easier 'ought' questions, thus facilitating progress in ethics (Greene 2014, p. 711). However, moral philosophers tend to be more concerned with the most fundamental issues that do not appear to be resolvable by an appeal to empirical evidence. Thus, philosophers generally focus more on the plausibility of normative assumptions themselves, rather than the processes underlying moral judgements. Rather than making normative issues easier to resolve, Greene's work might just shift the focus to a somewhat different set of difficult questions. No matter what philosophers choose to

focus on, there will always be untestable assumptions since philosophy deals with the most fundamental questions, including questions about knowledge itself.

In addition to these points, Greene's view of the unreliability of emotion-based moral judgements can also be challenged by making a positive case for emotions in this context. On this note, Folke Tersman considers how moral judgements may actually be enhanced by emotional processing: unless we adopt a radical conception of morality that has no relation to people's needs and interests, we should grant that emotional capacities such as empathy could be important in evaluating moral issues (Tersman 2008, p. 393). We cannot expect to properly account for factors such as people's interests and well-being if we are too emotionally detached from a situation involving such people. Tersman thus suggests that moral reasoning may need to be treated differently from other types of reasoning: even if empirical studies show that emotional processing interferes with judgements in other domains (such as science), this would not necessarily extend to the moral domain. In his words, "[E]thics is in many ways different, and it could be argued that the reliability of moral judgements, as contrasted with judgements in other areas, is in fact enhanced by certain kinds of emotional involvement" (Tersman 2008, p. 393).

Admittedly, this objection to Greene requires its own set of normative assumptions, specifically about the nature of morality and how it differs from 'non-moral' domains. However, the general subject matter of morality is surely less controversial than Greene's criticisms of emotion-based judgements. For instance, deontologists and consequentialists may disagree about the importance of factors such as well-being, duties and rights, but most would at least agree that such factors are all within the scope of morality. Thus, given that Greene's normative assumptions are much more contentious, and that emotions could enhance moral judgements as suggested by Tersman, it can be argued that Greene faces the burden of proof. So, even if Greene were to successfully demonstrate that deontology is mainly based on emotions while consequentialism is more associated with cognition, such an empirical finding would be irrelevant without normative assumptions about the relative merits of cognitive versus emotional moral judgements. Berker also draws attention to this point, noting that Greene's argument is incomplete without such assumptions:

> We need a substantive reason for thinking that intuitions based in emotion are less likely to be reliable than those based in 'reasoning' for this argument to be at all convincing. After all, there is a venerable tradition that sees emotions as an important way of discerning normative truths. One might disagree with this tradition, but showing that it rests on a mistake requires more than mere name-calling (Berker 2009, p. 316).

The points discussed so far demonstrate some of the ways in which one could argue that the burden is on Greene to demonstrate the unreliability of emotion-based moral judgements. At this point, Greene might fall back on his second point that was mentioned by Kumar and Campbell, which concerned the evolutionary adaptiveness of alarmlike emotional responses to certain factors. After all, Greene's claims about the unreliability of emotions were linked to his speculations about the evolutionary advantages of automatic emotional responses to personal harm. The general idea was that the alarmlike nature of these responses is not conducive to careful consideration of moral dilemmas. However, as with the first point about emotions in general, the significance of evolutionary explanations of emotion-based judgements also depends on normative assumptions: specifically, the apparent moral irrelevance of the factors to which our evolved emotional processes are sensitive.

Furthermore, even if we were to grant Greene's evolutionary speculations about this apparent adaptation, he has neglected to defend the more general notion that evolutionary adaptations are unreliable in terms of aligning with moral truth. Greene could at least attempt to draw upon ideas from Street and Joyce, but then he would encounter the problems discussed in the previous chapters. In opposing these types of evolutionary arguments, it has been argued that rational reflection is a by-product of our adaptive mental faculties, which may nevertheless allow us to identify moral facts. Nothing about this concept of rational reflection limits it exclusively to the 'cognitive' type of processing in Greene's dual-process model, at least in the context of morality. Although rational reflection was earlier defended by comparing its use in moral philosophy to its use in science, this does not necessarily mean that emotional and cognitive processes would play the same role across all domains. Following Tersman's suggestion, emotional processing may be more relevant and useful in some areas than in others.

Having rejected two of Greene's main points against emotion-based moral judgements, we may now consider the last of these points: the claim that emotional processes are relatively inflexible and simplistic compared to cognitive processes. This was supposed to favour cognition-based consequentialist judgements at least in the case of unfamiliar moral problems, as the greater flexibility would apparently allow us to properly consider the morally relevant factors involved. However, it has been noted that Greene's normative assumptions about moral relevance have not been sufficiently defended. More generally, even leaving aside the issue of moral relevance, it is not obvious that the more complex cognitive moral judgements would necessarily be superior to efficient but

inflexible emotion-based judgements. Since Greene considers these different types of moral judgements to underlie different respective moral theories, this point can be demonstrated by considering how complexity may affect the plausibility of moral theories themselves. Whereas too much simplicity in a theory may cause it to be unclear in its applications to certain issues, excessive complexity in any type of theory—not only in the moral domain—can introduce its own problems. Due to the limitations of human knowledge, more detail in a theory can raise more epistemic issues regarding the certainty of particular claims and their implications. As such, a more complex and detailed theory will generally depend on many other background theories and assumptions, thus opening up more ways to indirectly challenge and potentially undermine the theory. Therefore, even if flexible cognitive processes enable us to consider more factors in moral dilemmas, it does not necessarily follow that a theory more associated with such processes would be superior to a simpler theory that is more associated with emotion-based judgements.

Additionally, even if we were to grant Greene's point that automatic and manual processes are respectively better or worse in different situations, it would still be unclear how we could determine which kinds of moral issues are best suited to which type of processing. Greene's idea that we should favour manual cognitive processes in unfamiliar cases is weakened by the earlier points against his notion of evolutionary familiarity. If we cannot rely on evolutionary speculations to determine which cases favour which type of processing, then this further demonstrates the limited impact of Greene's empirical evidence: at most, all it could show is that our brains process different types of moral dilemmas in different ways.

Although more could be said about the reliability of different types of moral judgements, the discussion will now move to one last topic to conclude the assessment of Greene's argument. The issue concerns the notion of 'debunking explanations' in the context of Greene's argument, which also leads into some of the next chapter's points. To begin with, we may recall from Berker's points that Greene's evolutionary and empirical claims do not affect the argument from morally irrelevant factors. This was because Greene's debunking of deontology depended on his view that deontologists, due to their rationalisation of emotion-based judgements, focus more on morally irrelevant factors in moral dilemmas. Although the notion of moral relevance is clearly independent of Greene's empirical and evolutionary claims, he considered these aspects of the argument to be important in making his case. As noted by Tersman, Greene's evolutionary speculations are supposed to have a debunking effect because they can explain the different moral judgements between

footbridge and switch cases without supposing that there actually is a morally relevant difference (Tersman 2008, p. 395). In this sense, Greene's argument uses the same method as in Street's and Joyce's debunking arguments: they all claim to provide a non-moral, scientific explanation of phenomena that are traditionally regarded as being distinctly moral. On this note, Tersman describes the concept of debunking explanations as follows:

> Consider a fact F that is offered as evidence for a theory T. A debunking explanation of F is an explanation that does not entail that T is true or significantly likely. To provide such explanations is a common way to question the relevance of considerations offered as evidence (Tersman 2008, p. 395).

However, it is debatable whether normative ethical theories require any 'evidence' beyond a sufficient level of rational argumentation in their favour. On Greene's view, it seems that deontologists are expected to prove that considerations against pushing the man off the footbridge are based on the plausibility of their theory, rather than being 'explained away' by evolutionary speculations.

On this point, it is worth noting that because normative ethical theories have existed since long before neuroscientific experiments and EDAs were possible, normative claims have long been provided by traditional 'armchair' philosophy. This leads into the broader issue of whether moral theories now need to be supported by empirical evidence in order to be taken seriously. An affirmative answer would radically change moral philosophy, as it would require a significant integration of current scientific and philosophical methodology (Berker 2009, p. 295). Greene's experiments are just one example of an attempt at such integration. It is currently uncertain whether we could test much beyond the prevalence of certain intuitions, the neural activity during moral judgements, and the factors that people respond to in moral dilemmas. As such, it currently seems unlikely that scientific methods could enable us to test or prove any kind of 'first principle' in moral philosophy, such as the consequentialist view that only consequences have any direct moral relevance. It is also unclear how the notion of morally relevant factors could be empirically supported. This is because such principles and assumptions concern the fundamental nature of morality. The situation may be comparable to the relation between metaphysics—representing the most fundamental assumptions about reality—and empirical science, which focuses more on verifiable claims and theories that depend on background assumptions.

Ultimately, these points suggest that Greene's argument fails partly because his empirical findings lack any significant debunking role in the argument. If debates between consequentialists and deontologists depended on empirical questions such as which factors

our moral judgements respond to, then the types of experiments conducted by Greene could certainly be useful. However, the debate has always been much broader than this, as it traces back to the different fundamental principles underlying each opposing theory. Greene thinks that his empirical findings can help us resolve philosophical problems when combined with independent normative assumptions, but the experiments cannot tell us what assumptions to make. This is the case even if we already assume that there must be some kind of independent moral truth about the trolley and footbridge cases, and even if we think that only one normative theory or framework can accommodate the truth.

Overall, it has been shown that Greene's debunking argument against deontology does not fare much better than Street's or Joyce's arguments. Now that three of the most prominent EDAs have been rejected, the final chapter examines some issues that apply to evolutionary debunking arguments more generally.

Chapter 5 – General Issues for EDAs: Evolution and Explanation

Much of the EDA literature has passed over the evolutionary premises in these arguments, often taking these speculative accounts for granted at least for the sake of argument. Since this aspect of EDAs has often been overlooked, this chapter now focuses on general issues that pertain to evolutionary claims in these types of arguments. The first point to be examined is whether the kinds of evolutionary claims from debunkers are actually plausible. This general question arises from the issue of whether evolutionary premises in particular—as opposed to non-evolutionary debunking accounts of morality—are essential to the discussed EDAs. Various perspectives on this matter will be considered, beginning with the view that specifically evolutionary claims have essential importance in EDAs. In Section 5.1 it will be argued that any EDAs that specifically rely on evolutionary accounts are weakened by some general problems with these types of explanations.

In Section 5.2, the next perspective to be considered is the contrasting view that evolutionary claims in EDAs may be interchangeable with non-evolutionary accounts of morality, such as sociological or historical explanations of moral beliefs. The discussion focuses on the implications of the limited relevance of specifically evolutionary explanations in debunking arguments. This then leads into another view that reduces the apparent significance of evolutionary claims: the notion that any explanation of morality—whether evolutionary or otherwise—will have a limited role compared to the philosophical assumptions in debunking arguments. For instance, EDAs may primarily depend on epistemic claims about the reliability of certain belief-influencing factors with regard to moral knowledge. Although these types of claims are to be expected in a philosophical argument, this point will be used to suggest that debunking arguments lack any inherent advantage over certain positive arguments in moral philosophy, such as defences of moral realism.

Since one of the main apparent advantages of EDAs is their appeal to scientific explanations of morality, the focus in Sections 5.3 and 5.4 then turns to explanatory issues that are relevant to EDAs at a more general level, such as the relation between scientific and philosophical accounts of morality. The discussion will particularly focus on ways in which scientific explanations—such as evolutionary explanations—can be compatible with philosophical accounts. This compatibility view will be used to oppose the general

'explaining away' strategy of EDAs, which is partly based on the point that evolutionary explanations make no reference to philosophical concepts, such as moral facts.

Finally, this chapter will conclude that future EDAs and discussions surrounding them should pay more attention to these general explanatory issues and the points concerning evolutionary premises, especially if future debunking arguments hope to overcome the problems with the current EDAs in the literature.

5.1 The Plausibility of Evolutionary Premises

When assessing EDAs in general, one of the main points to consider is the status of their evolutionary premises. One possible view is that specifically evolutionary premises play an essential role in EDAs, such that they could not be substituted with non-evolutionary debunking explanations. If one holds this view and aims to oppose EDAs by targeting their evolutionary premises, there are various points that can be used to challenge the general plausibility of evolutionary debunking accounts of morality. Some particularly relevant points are found in David Buller's assessment of various ways in which the popular understanding of evolutionary psychology tends to be limited and misguided. According to Buller, one problematic assumption is that we can understand how our minds have been 'designed' by selection by speculating about possible adaptive problems faced by our early human ancestors (Buller 2009, p. 76). Street speculated about the types of widespread moral beliefs that would have been adaptive, Joyce speculated about the adaptiveness of our supposedly innate moral sense, and Greene's conjectures focused on the adaptiveness of moral beliefs that rationalise an aversion to personal violations such as directly inflicting harm. These evolutionary accounts are all problematic, as Buller notes that we do not know nearly enough about the living circumstances of our distant ancestors to identify the adaptiveness of certain types of mental traits. Unfortunately, the only way to provide a more detailed account of mental adaptations is to rely more on speculation (Buller 2009, p. 76).

Debunkers' reliance on speculation has been noted in previous chapters; this will not help their case that evolutionary accounts provide the most plausible explanations of moral beliefs, as opposed to truth-conducive explanations such as the by-product account of rational reflection. We cannot expect many more evolutionary details to be filled in, as most of the evidence has been lost over time. Thus, if a specifically evolutionary account

of morality is required for EDAs, then these arguments will be troubled by the lack of evidence, which compromises the explanatory power of such accounts.

Despite the limited evidence for evolutionary explanations, their appeal may be partly due to their potential to support global rather than selective debunking arguments (O'Neill 2015, p. 1078). After all, evolutionary forces have undoubtedly influenced our general mental capacities to some extent. In contrast, certain non-evolutionary explanations of moral beliefs may be more amenable to empirical confirmation, at the cost of a more limited scope for debunking claims. For instance, a debunking argument could be based on the notion that certain measurable emotional responses are a 'distorting' influence on moral judgements: one example is Daniel Kelly's view that we should be sceptical of moral judgements that involve the emotion of disgust (Kelly 2011, p. 140). Clearly, emotional responses can be empirically tested, as demonstrated by Greene's experiments.

The choice between selective versus global debunking arguments appears to be a choice between explanatory scope or certainty, assuming that only evolutionary accounts are sufficiently broad to support global arguments. Whereas the more speculative evolutionary factors may influence a vast array of adaptive moral beliefs, direct causes that can be ascertained through empirical methods only seem applicable to much narrower debunking arguments. In other words, we would not expect any particular type of direct cause (such as a certain emotion) to be involved in all or most moral judgements. Thus, in Street's global debunking argument against moral realism, her appeal to an evolutionary account may be partly motivated by the idea that it provides a particularly broad non-truth-tracking explanation of moral belief tendencies, as compared to narrower debunking explanations based on more immediate influences. In this way, global debunkers might hold that evolutionary influences on moral beliefs have more philosophical significance than the more immediate influences such as emotions, since they ultimately explain why we have particular emotional responses to moral issues.

On this note, Greene's argument is an interesting case as he uses evolutionary premises for a selective debunking argument, exclusively targeting deontology. However, he still faces the problems of global debunking arguments, as it cannot be empirically confirmed that the examined emotional influences on certain moral beliefs are ultimately due to distorting evolutionary factors. Thus, Greene's argument still relies on evolutionary speculations as with global EDAs. Overall, the inability to empirically confirm speculations about evolutionary influences on our moral beliefs is problematic for any

EDAs that cannot replace their evolutionary premises with a testable, non-evolutionary explanation of the moral views that they intend to debunk.

There are further difficulties facing debunking arguments that specifically depend on evolutionary claims: apart from the lack of evidence that would support evolutionary debunkers' explanations of morality, there is also counter-evidence against their explanations. This can be seen by considering debunkers' assumptions regarding the extent to which humans have retained primitive psychological tendencies, as opposed to the by-product view of rational reflection that allows us to overcome such tendencies. The notion that humans essentially still have a 'Stone Age' mind would support EDAs by suggesting that our mental capacities have not significantly evolved beyond the adaptive functions that they granted in our evolutionary past, at least in the case of the moral reasoning underlying philosophical views that debunkers oppose. Debunkers have defended their own moral views from this point by arguing that they cannot be explained away as adaptations; for instance, Greene claimed that only deontology is a rationalisation of evolved alarmlike emotions, whereas consequentialism is apparently unaffected.

On this point about adaptations, Buller opposes the view that our minds are still primarily adapted to the distant past. He notes that popular conceptions of evolutionary psychology tend to underestimate the rate of evolutionary change, particularly in the context of rapidly changing human environments:

> Any Pleistocene-selected genes we possess will interact with these new environments to produce psychological traits that may differ in important ways from those of our Pleistocene ancestors. So there is no good reason to think that all of our evolved psychological characteristics remain adapted to the lifestyle of Pleistocene hunter-gatherers (Buller 2009, p. 79).

This supports the view that our more recently evolved capacity for complex rational reflection can overcome distorting evolutionary influences relating to adaptiveness in our ancestral environment, as argued in Chapter 2. Buller notes that although the timeframe since the Pleistocene epoch is a relatively short evolutionary period of around 10,000 years, it is not as though the human brain would have needed to evolve completely new complex designs over this period in order to significantly alter our psychological tendencies. Thus, in as few as 400 generations our previously Pleistocene-adaptive psychological tendencies could be substantially modified by selection (Buller 2006, p. 206).

Buller substantiates these points with specific examples of rapid evolutionary change in humans, which can be used to support the case that humans have had enough

time to develop a moral reasoning capacity that is sufficiently autonomous from the evolutionary influences of our past environment. For instance, there are clear cases of relatively recent physiological adaptations in humans, such as the increased prevalence of lactose tolerance being driven by the domestication of cattle, which enabled dairy farming (Buller 2006, p. 207). In terms of human behaviour and morality, Buller cites the agricultural and industrial revolutions as notable examples of major changes in the organisation of societies, which have changed the selection pressures that affect interpersonal behaviours. In particular, living among much larger groups of people in modern cities has changed the challenges related to mating and forming alliances (Buller 2006, p. 206).

Partly as a result of changing social structures, prevailing moral attitudes throughout the world have changed significantly in recent history; this point will be examined shortly (Huemer 2016, pp. 1988-1994). These points about rapid evolutionary change—which has enabled humans to reflect on moral beliefs and reconsider them— clearly conflict with the notion of our moral capacities still being 'saturated' with evolutionary influences relating to the adaptive challenges of our distant past. Buller's points thus suggest that we humans have had sufficient time to adjust our psychological tendencies away from primarily adaptive functions and towards truth-seeking reasoning capacities. The capacity for rational reflection is able to thrive in our modern environment where the prospects of survival and reproduction have been greatly enhanced by our changing social structures and medical advances. Since most of us do not need to worry so much about survival in these circumstances, we can devote our mental capacities to the investigation of highly abstract and complex issues, as demonstrated by science and philosophy. However, it is a separate question whether our capacity for non-adaptive moral reasoning has actually enabled us to identify moral truths. Nevertheless, even if it turns out that the prevailing moral beliefs of our time are still not closely aligned with moral truths, these points at least suggest that our moral reasoning capacities have advanced due to changing circumstances; our capacities are not significantly constrained by the circumstances of our distant evolutionary past.

This view is further supported by Allen Buchanan and Russell Powell, who argue that evolutionary explanations of morality cannot account for contemporary moral attitudes (Buchanan & Powell 2015). They focus on recent moral developments, which serve as evidence of Buller's claims about rapid evolutionary processes enabling significant psychological changes: "[C]ontemporary morality, as experienced and exhibited by significant numbers of people, is strikingly more inclusive than one would expect if

selectionist explanations were the whole story, or even most of it" (Buchanan & Powell 2015, p. 48). To substantiate this claim, they refer to various 'inclusivist' elements of contemporary moral views that have become increasingly prevalent in recent history. For instance, we have seen the emergence of the view that non-human animals are subjects of proper moral consideration, not merely for instrumental reasons regarding human interests. Another example is the notion that moral norms can be universalised rather than varying based on an individual's group membership, such as one's ethnicity or religion. There is also the idea of human rights applying regardless of an individual's status and contributions in society (Buchanan & Powell 2015, pp. 48-50). These are just some elements of contemporary morality that do not seem amenable to debunkers' evolutionary accounts. Rather, these points support the idea that human rational capacities have become sufficiently autonomous from evolutionary influences, including in the moral domain.

Having challenged the evolutionary accounts offered by debunkers, this makes way for the development of a positive account of moral knowledge. This is exemplified by Michael Huemer's account of moral realism (Huemer 2016), which will be briefly examined to illustrate how evolutionary explanations of morality can be compatible with the views that debunkers have targeted, such as moral realism. Huemer's argument includes his own account of how humans have gradually headed towards moral truths from an evolutionary starting point. He argues that although early humans started off with mostly distorted moral beliefs under the influence of evolutionary forces, many different societies have gradually converged towards moral truth over time. Thus, he describes the initial primitive state of human moral beliefs as follows:

> In primitive times, human beings begin with badly misguided moral beliefs. This parallels the widespread and severe error that primitive societies begin with in all other areas of inquiry. In the case of morals in particular, we have non-rational emotions and desires influencing our beliefs and hence leading us astray—the very sort of influences that the debunking skeptics advert to in their effort to impugn all moral beliefs (Huemer 2016, p. 2004).

He notes that despite these primitive origins, in recent history various moral beliefs that seem to indicate evolutionary influences have fallen out of favour while liberal moral values have become more prevalent. For instance, the belief that it is right or permissible to conquer other societies for their territory and resources, and the different standards of sexual morality for males and females have been challenged (Huemer 2016, p. 2001). Huemer uses this point about gradual convergence towards liberal moral values to argue that these values reflect moral truth. It is not necessary to assess the plausibility of this

particular account of realism, as it merely serves to illustrate how opponents of debunking arguments—such as moral realists—may account for evolutionary factors while developing a positive argument. This demonstrates that an acknowledgement of our evolutionary history need not be limited to debunking arguments:

> In one trivial sense, an evolutionary account of ethics must be correct: human beings evolved; therefore, however our capacity for moral judgment works, that capacity is "a product of evolution," in the same sense that our capacity for any sort of judgments is a product of evolution. This thesis of "an evolutionary origin for ethics" poses no threat to moral realism (Huemer 2016, p. 1994).

As such, the philosophical significance of evolutionary accounts of morality crucially depends on certain assumptions, particularly epistemic claims about the reliability or unreliability of our moral reasoning capacities that have initially been shaped by evolutionary processes. Thus, a scientific explanation of human morality can be combined with either positive or negative epistemic claims about the effects of evolutionary influences, leading to either a positive account of moral knowledge such as Huemer's, or a negative account as exemplified by EDAs.

The points that have been discussed so far in this chapter are most relevant to any EDAs that specifically depend on evolutionary accounts of morality, as it is clear that we should not simply take their evolutionary claims for granted. Unless debunkers can account for modern developments in moral reasoning, their evolutionary premises are questionable since they overstate the continuing influence of our distant evolutionary past. As it stands, the current debunking explanations of morality do not seem any more plausible than accounts offered by those who resist EDAs, such as moral realists' accounts of moral knowledge. Overall, this suggests that proponents of EDAs may be better off developing non-evolutionary debunking arguments. However, the following discussion will show that even this option presents some difficulties for debunkers.

5.2 The Relevance of Evolutionary Premises

Although debunkers' evolutionary explanations of morality are problematic, the general structure of debunking arguments could be preserved if it turns out that evolutionary premises in particular are not essential to EDAs. Thus, it might be the case that any explanation of morality could serve the same function in these arguments, as long as such an account does not seem conducive to the truth or plausibility of the targeted

philosophical position, such as moral realism or deontology. To use Joyce's argument as an example, although it may appear to depend on specifically evolutionary claims since he devotes a major portion of his work to developing an evolutionary genealogy of morality, Joyce has made the following statement about his position:

> [T]he 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 judgements are true, the same epistemological conclusion could be drawn (Joyce 2016, p. 125).

Thus, if there is nothing special that distinguishes evolutionary explanations of morality from non-evolutionary debunking explanations, then EDAs may simply be treated as one variety of genealogical arguments that target moral beliefs; this view clearly reduces the apparent relevance of evolutionary premises. On this point, Russ Shafer-Landau notes the similarities between different kinds of genealogical critiques of moral beliefs: "Despite the variety of specific forms that such critiques have taken, all instances of the form share a common structure. They first allege an empirical claim about the causal origins of our moral beliefs, and then proceed to raise doubts about the reliability of beliefs formed on that basis" (Shafer-Landau 2017, p. 175). This seems to entail that debunking explanations are essentially about the causation of moral beliefs, as moral beliefs are being questioned specifically due to their causes.

However, considerations about belief causation cannot have a debunking effect unless supplemented by some account of the relation between certain belief causes and moral truth. This relation must also be accounted for by proponents of positive views of moral knowledge, even in the absence of any debunking challenges. In this sense, debunkers are essentially attempting to shift the explanatory burden to their opponents, by suggesting that a debunking account of this relation is more plausible than any positive account. This is noted by Jeff Behrends in a discussion of debunking arguments against moral realism: "[T]he challenge that the realist must meet is that of explaining how, in light of a non-normative genealogy of our normative judgments, those judgments could come to be correctly correlated with the normative facts, realistically construed" (Behrends 2013, p. 487). The challenge of explaining a connection between moral beliefs and facts is nothing new for moral realists; debunkers simply provide their own allegedly more plausible accounts of morality that are supposed to rule out the possibility of any reliable connection. It has been argued throughout the thesis that EDAs are unsuccessful in both regards: the evolutionary accounts in EDAs have been criticised, while Chapter 2

presented the by-product view of rational reflection that enables positive accounts of moral knowledge.

Thus, given that both debunkers and their opponents need an account of the relation between moral beliefs and moral truth, the relevance of particular debunking accounts of morality may depend on their plausibility as compared to that of positive accounts. Although evolutionary accounts are problematic, debunkers who seek alternative debunking explanations should aim to present accounts of morality that are at least as detailed as the philosophical accounts they intend to undermine. On this note, Shafer-Landau's criticisms of insufficiently detailed evolutionary debunking accounts are also relevant to other kinds of debunking explanations: "Absent a specific story about how selective pressures actually did work to form our moral dispositions and beliefs, the debunker is left with an unsubstantiated allegation of unspecified influence, which cannot be enough to warrant a skeptical attitude toward our moral beliefs" (Shafer-Landau 2017, p. 179). Similarly, William FitzPatrick maintains that evolutionary debunking explanations should not just be regarded as the most plausible accounts by default:

> The debunkers could of course turn out to be correct. They do not, however, win by default, and they cannot legitimately use science to provide decisive leverage in favor of their position, since the science itself fails to provide adequate support for the very strong and generalized explanatory claims on which they rely (FitzPatrick 2017, p. 201).

In other words, strong claims about morality require strong explanations, rather than the highly speculative and problematic accounts in the discussed EDAs. These points also apply to any kind of debunking explanation of morality, so although other explanations could potentially avoid the problems with evolutionary accounts, it would still be difficult to develop a plausible debunking explanation with a significant scope.

Following on from the idea that any genealogical account of morality could be used in debunking arguments, this may seem to suggest that any given explanation of morality will be relatively unimportant as compared to the philosophical aspects of such arguments, such as the application of certain epistemic principles to particular causes of moral beliefs. Epistemic principles are certainly required in order to demonstrate the significance of certain claims about causes of moral beliefs. For example, debunkers may use the notion of 'sensitivity' to argue that a complete explanation of the causes of our moral beliefs reveals them to be insensitive to moral truth, as we would hold the same beliefs whether or not they are true (Leibowitz & Sinclair 2017, p. 217). This idea seems applicable to each of the three main EDAs that have been discussed: they all identify evolutionary adaptiveness as

the main influencing factor, which presumably requires no connection to moral truth. Another relevant epistemic principle is that of 'safety', which can challenge moral beliefs by suggesting that even if some of them are true, this would not indicate reliability unless we can rule out the possibility that their truth is purely a matter of luck (Leibowitz & Sinclair 2017, p. 218).

Regardless of which epistemic principles are used, it appears that such claims will do most of the important work in debunking arguments, rather than the debunking explanations of morality (Leibowitz & Sinclair 2017, p. 220). It is therefore important for debunkers to recognise the limited role of their purported explanations; this point has at least been acknowledged by Joyce: "No one, though, thinks that genealogical empirical data alone can secure a sceptical victory; at most it battles alongside sceptical arguments of an a priori metaethical nature" (Joyce 2016, p. 136). As such, debunkers should not overstate the importance of their explanations as compared to the philosophical aspects of their arguments.

This view of debunking explanations can be demonstrated by briefly examining Michael Klenk's assessment of Street's EDA (Klenk 2017). As Street mainly focuses on debunking non-natural versions of moral realism, Klenk argues that her evolutionary argument actually depends on a more general epistemic challenge: specifically the Benacerraf-Field challenge of explaining how we can have knowledge of any kind of causally inert, mind-independent properties (Klenk 2017, p. 786). The causal isolation problem can be avoided by naturalistic moral realism, since it identifies moral facts with certain natural facts that are accessible to us. Thus, the idea is that Street's argument partly depends on a problem inherent to non-natural moral realism, which is just as problematic in the absence of any evolutionary considerations. As such, Klenk asserts that evolutionary premises are not merely insufficient to reach Street's conclusion, they are actually redundant since the causal isolation problem does the debunking work (Klenk 2017, pp. 787-788). For any purported non-truth-conducive causes of moral beliefs—whether evolutionary or otherwise—such causes would be irrelevant if it turns out that no belief-forming mechanism could allow us to reliably identify non-natural moral facts.

Similar points have been made by Daniel Crow, although his overall view contrasts with Klenk's in some ways (Crow 2016). Crow distinguishes between causation-based objections and evolutionary objections to non-natural moral realism. He observes that the former type of objection depends on the claim that non-natural moral properties cannot influence our moral beliefs, whereas the latter objection is that our beliefs have been influenced by evolutionary processes that are presumed to not involve any such moral

properties (Crow 2016, p. 380). Thus, even if non-natural moral facts are potentially accessible to us, the point of the evolutionary challenge is that our beliefs have been influenced by evolutionary factors rather than these facts. As such, Crow's view is that evolutionary objections to non-natural moral realism are not dependent on or reducible to a causal isolation claim; this clearly differs from Klenk's perspective (Crow 2016, p. 381). However, even this position suggests that debunking explanations have a limited role, as all that matters is the assumption that the causes of our moral beliefs are something other than moral facts; the actual method of causation is not particularly important. Since debunking arguments can target views other than non-natural moral realism, the more general point here is that both Klenk's and Crow's perspectives downplay the role of debunking explanations as compared to the epistemic assumptions in debunking arguments.

Overall, these points indicate that debunking arguments are not so different from more traditional philosophical arguments in which the premises are not based on scientific claims. This can be taken as a negative assessment in the case of EDAs, since they appear to draw much of their initial appeal from the assumed scientific credibility of evolutionary explanations. Reducing the importance of evolutionary premises thus suggests that EDAs lack any inherent advantage over non-evolutionary arguments in moral philosophy. At the least, this challenges the notion that defenders of positive moral views face the burden of proof in light of evolutionary explanations, as opposed to the debunkers who are sceptical of these moral views.

5.3 Explanatory Issues for EDAs: Incompatible Accounts?

Since EDAs appeal to the apparent credibility of their scientific explanations of morality to challenge certain philosophical accounts of morality, the relation between scientific and philosophical accounts warrants further discussion. Scientific investigations of morality are concerned with explaining why human moral thought has developed in certain ways, whereas moral philosophy is primarily focused on the nature of moral truth and moral knowledge. Debunking arguments attempt to draw a link between explanatory scientific premises and philosophical conclusions, in such a way that seems conducive to 'explaining away' certain philosophical views. On this point, the role of debunking explanations could potentially be expanded if debunkers could show that their scientific explanations are incompatible with the philosophical views that they aim to debunk, in such a way that

favours their own scientific accounts and allows them to 'explain away' these particular philosophical views.

Debunkers such as Street, Joyce and Greene present evolutionary explanations of morality that are supposed to override or challenge certain philosophical views that seem incompatible with these explanations. However, it should not simply be taken for granted that scientific explanations of morality necessarily conflict with certain philosophical moral views such as those targeted by debunkers. Before discussing ways in which scientific and philosophical accounts may be compatible, the first point to address is the incompatibility view that appears to be represented by the 'explaining away' approach of EDAs. This incompatibility view will be rejected, thus reducing the philosophical significance of debunkers' explanations which are supposed to challenge certain moral views.

The incompatibility view is the notion that we must identify just one type of explanation as correct, or at least much more plausible than any alternatives. For instance, Joyce considered his evolutionary explanation of morality to be plausible despite the fact that it does not presume any moral judgements to be true. Therefore, the incompatibility view would suggest that the apparent completeness of this scientific explanation makes it more plausible than positive philosophical accounts of moral knowledge. As such, the philosophical significance of Joyce's evolutionary account partly depends on the notion of scientific completeness, which does not seem attainable by non-sceptical philosophical accounts of morality.

With this example in mind, it is worth considering a principle that is used when assessing multiple proposed explanations: inference to the best explanation (IBE). IBE is often used in epistemology, metaphysics and the philosophy of science in general, and is considered by some philosophers to be essential and fundamental (Day & Kincaid 1994, pp. 271-272). Since the term 'best' requires a point of comparison, IBE is only applicable when there are multiple proposed explanations. Thus, if several proposed explanations seem plausible when considered on their own merits, their relative plausibility must then be assessed according to some general criteria. Some examples of explanatory criteria may be recalled from Chapter 2: Street cited parsimony, clarity and overall explanatory power as points in favour of her debunking explanation, as opposed to a truth-tracking account of moral beliefs.

Notably, debunkers such as Street have applied explanatory criteria to moral philosophy in the same way that these criteria would normally be applied in science. As such, the philosophical significance that they ascribe to their evolutionary explanations

seems to draw upon the notion that certain philosophical views cannot be compatible with their scientific perspective on morality, thus leading them to 'explain away' these views. This highlights a general issue with EDAs: debunkers have not elaborated on why they hold metaethical and normative moral theories to the same explanatory standards as scientific theories, such as evolutionary theories. Debunkers need to adequately defend their view that scientific methods can be used to derive philosophical conclusions, as the use of such methods in philosophy remains contentious. On this note, FitzPatrick asserts that we cannot simply take it for granted that science provides the best explanations of moral beliefs, as this assumption largely depends on debunkers' own philosophical views that cannot be directly supported by appealing to science (FitzPatrick 2016, pp. 397-398). Thus, to determine whether science and philosophy require different explanatory standards or frameworks, it will be helpful to briefly consider some aspects of the relation between these disciplines, in order to identify some of the relevant points of difference. As this is a very broad topic, only some of the most relevant aspects can be discussed here.

The differences between philosophy and science have been examined by Robert Hartman, who observes that scientific theories are generally more precise and complex than theories in moral philosophy, particularly in the natural sciences such as physics (Hartman 1963, p. 353). Although some philosophers such as Kant have developed highly complex and rigorous moral theories, these still do not compare with the level of precision and detail in modern physics (Hartman 1963, p. 354). This can be partly attributed to the nature of philosophy, which deals with issues at the most abstract and general level, such as what kinds of assumptions about the world are reasonable to hold. In contrast, each field of science focuses on specific subject matter and investigates particular phenomena in as much detail as possible, under the guidance of various underlying assumptions. Thus, it is an important point that scientific inquiry is situated within the broader philosophical framework of fundamental assumptions about knowledge and reality; these underlying assumptions are usually examined separately in the philosophical domains of epistemology and metaphysics. As such, although science encompasses physics—which investigates the most fundamental components of the universe—philosophy is still more fundamental as it is concerned with such abstract and general topics as truth and knowledge itself.

As topics such as the nature of moral truth and moral knowledge are highly abstract compared to scientific topics such as evolution, it may therefore be unreasonable to expect a scientific standard of precision or parsimony in philosophical accounts that involve concepts such as moral facts. Thus, the scientific completeness of an evolutionary account that makes no reference to abstract philosophical concepts does not necessarily undermine

certain philosophical views. At the least, it cannot be assumed without argument that moral theories should be held to the same standards as theories in the natural sciences. This is partly because the abstract nature of moral philosophy presents difficulties for any attempt to verify moral theories, or particular claims about moral truth. As Hartman observes, "Ethics is still a philosophy and not a method. It lacks the precision, and hence the complexity, to be applicable to definite human situations" (Hartman 1963, p. 354). If it is not possible to develop or assess philosophical theories with the rigorous processes used in science, then there is no reason to hold moral claims and theories to unattainable scientific standards; we may simply have to accept the limitations of moral philosophy. This raises many issues regarding methodological naturalism in moral philosophy, which are beyond the current scope: the important point here is that it is controversial to apply scientific methods to philosophy. As such, the standing of EDAs in general partly depends on contentious methodological issues that will have to be resolved separately.

Although EDAs have attempted to overcome some of the limitations of traditional philosophical arguments by appealing to science, they will always be constrained by the abstract nature of moral philosophy, as debunkers are still concerned with matters such as moral knowledge. Since this point about abstract subject matter clearly applies to any argument in moral philosophy, it cannot be used to exclusively criticise EDAs. Rather, the point is that EDAs must not only justify their own methodology and their adoption of a scientific perspective; they must also justify holding their targeted philosophical views to these standards. For instance, the principle of parsimony has been used in EDAs to undermine moral realism, but the application of this principle in philosophical contexts is more controversial than its role in science (FitzPatrick 2016, p. 398).

Overall, it appears that debunkers cannot rely on an incompatibility view as the basis for deriving negative conclusions about certain philosophical views from their scientific explanations of morality. As will be discussed in the next section, opponents of EDAs could favour a compatibility view of evolutionary explanations with their own views, which will be demonstrated by focusing on non-conflicting differences between scientific and philosophical accounts of morality.

5.4 Explanatory Issues for EDAs: Compatibility

Before considering the non-conflicting differences between scientific and philosophical views, it is important to note that emphasising the differences between these disciplines

does not necessarily undermine the by-product rational reflection account of moral knowledge, which has been presented as an objection to EDAs such as Street's. This view involved a comparison between scientific and moral knowledge, but this does not contradict the present chapter's discussion: the comparison was purely in terms of our evolved mental capacities enabling both kinds of complex knowledge. The view does not entail that our mental capacities are limited to knowledge of precise scientific subject matter. As such, the relatively abstract nature of moral concepts does not necessarily support evolutionary debunkers' view that certain moral beliefs are hopelessly misguided.

Despite the differences between scientific and philosophical accounts of morality, one way to defend a compatibility view is to argue that these may just represent different levels or aspects of explanation. The notion of multiple viable accounts of a phenomenon—morality in this case—can be used to counter the incompatibility view that would favour EDAs. This general idea of different but compatible levels of explanation is exemplified by Brad Majors in his defence of moral realism against scientific explanatory challenges. His argument essentially supports the view that focusing exclusively on the most basic causes or aspects of a phenomenon does not necessarily provide the most complete and useful explanation of it (Majors 2003). Specifically, he uses the example of biological and psychological explanations as compared to purely physical explanations. In this context of different descriptive theories, physics provides the most fundamental level of scientific explanation; physical explanations only refer to the basic physical properties of the universe. In contrast, biological and psychological explanations are at a different level, as they refer to non-basic properties and relations that are not mentioned in purely physical explanations.

However, the fact that a phenomenon can be explained without appeal to certain levels of explanation does not entail that these other explanations are inaccurate or redundant. As Majors expresses this point, "[E]xplanations at different levels do not generally compete with one another. The fact that a certain phenomenon has a physiological explanation does not prevent it from having a true and valuable psychological explanation" (Majors 2003, p. 126). Just as Majors rejects the view that physical explanations make other accounts redundant, this can be extended to the idea that moral theories may be viable even if there are plausible parsimonious explanations of morality. On this note, the role of non-basic levels of explanation is highlighted by David Owens:

> [T]here is a hierarchy made up of different levels of explanation. Physics is at the base of this hierarchy and the rest of the structure depends upon it. But the higher

reaches of the scientific edifice have explanatory features which could not be discerned by someone who confined himself to exploring the ground floor (Owens 1989, p. 59).

This suggests that certain levels of scientific explanation—such as physical explanations—do not necessarily provide useful accounts of all phenomena, at least not to the extent of making other types of explanations redundant. This idea is supported by the fact that physicists have been unable to 'translate' other kinds of explanations—such as economic explanations—into the language of physics (Owens 1989, p. 59). The same point may be applicable to parsimonious evolutionary explanations of moral beliefs: even if such explanations are complete from a scientific perspective, their lack of reference to moral facts does not clearly entail the non-existence of such facts. This could be considered analogous to the view that a physical explanation's lack of reference to any biological properties does not entail that there are no biological facts. However, this analogy raises issues concerning the relation between descriptive and normative theories, as opposed to different levels of description such as physical versus biological explanations. Some philosophers have nevertheless argued along these lines; for instance, Nicholas Sturgeon makes the following point in the context of moral and non-moral explanations:

> [T]he question of how we decide when nonmoral explanations conflict with, or support or amplify moral ones. The answer, I believe, is that abstractly the procedure is no different from that used in answering similar questions about chemical and biological explanations, or about psychological and sociological ones. Conflict between explanations is virtually never a matter of outright contradiction (Sturgeon 1992, p. 100).

David Copp has also suggested that different types of explanations can be compatible in the case of moral and non-moral explanations:

> [W]e would not want to say that the existence of a neurophysiological alternative to every psychological explanation would show automatically that no psychological explanation is evidence of any psychological postulate. A scientifically viable alternative to a moral explanation is not necessarily in competition with the moral explanation, and its mere existence does not show that the moral explanation is evidentially impotent (Copp 1990, p. 243).

Although there is insufficient space to assess these views, the point is that explanatory issues in the context of EDAs can be traced back to more general issues regarding moral and non-moral explanations. If it turns out that philosophers such as Sturgeon and Copp are correct that moral explanations can be compatible with non-moral explanations, this

could undermine EDAs at a more general level. For instance, Street would no longer be able to argue against moral realism on the grounds that her evolutionary explanation of morality does not refer to moral facts; this may just be a different level of explanation from that offered by moral realists. Overall, these points suggest that the relation between moral and non-moral explanations may warrant further attention in discussions of EDAs.

Even if it turns out that debunkers cannot undermine particular moral views by highlighting the apparent problems with moral explanations, there are other ways in which they might argue that their scientific explanations of morality have implications for certain philosophical views. For instance, they might criticise certain moral theories on the grounds that they lack any method of confirmation: it is not clear how we could verify moral facts if they lack any observational consequences. In response, it could be argued that science is not completely different from philosophy in this regard. It can thus be observed that the 'a priori' methodological aspects of moral theories are not exclusive to philosophy: science also depends on assumptions that may be untestable or involve unobservable phenomena. The essential role of such assumptions is noted by Robert Audi:

> A priori explanations are non-scientific; but they need in no way be such that philosophers or others who are scientifically oriented should reject them. Logic and pure mathematics, after all, are essential for science—and for scientific explanations—but contain a priori explanations (Audi 2014, p. 161).

This commonality between science and philosophy seems to be downplayed or overlooked by EDAs that criticise particular moral views on the basis of scientific standards of explanation. As such, debunkers should take note of the following point: "[W]e must not take the prominence of scientific methods and explanations as paradigms to give the impression that they are the only instances of theoretical methods and of genuine explanations" (Audi 2014, p. 146).

The issue of confirmation in moral theories has also been examined by Geoffrey Sayre-McCord (1988, p. 433). He makes the similar point that moral theories cannot be challenged on the grounds that they posit unobservable entities, as scientific theories do the same (Sayre-McCord 1988, p. 435). Following on from Audi's point that a priori assumptions are essential in both science and philosophy, Sayre-McCord also notes that theories in both disciplines require background assumptions in order to be testable; as such, we should not reject a scientific or moral principle just because it is untestable when considered in isolation (Sayre-McCord 1988, pp. 436-437). It can thus be asserted that the acceptance of background assumptions in science should also be extended to fundamental assumptions in moral philosophy. In this case, debunkers must either grant their opponents

some assumptions as with their own scientific accounts, or they must convince us that philosophical theories are entitled to fewer untestable assumptions as compared to scientific theories. Considering the abstract nature of philosophy, this latter option would surely be an unreasonable constraint to place on moral theories. This leaves debunkers with the other option of granting their opponents some basic assumptions, which could provide the necessary foundation for developing positive moral theories and resisting EDAs. By allowing such assumptions, debunkers would have to concede that their scientific explanations of morality do not necessarily put pressure on certain philosophical accounts. Overall, these points support a compatibility view of evolutionary theories and moral theories by highlighting important similarities between science and philosophy, along with the possibility that they represent the non-conflicting aspects of different levels of explanation.

5.5 Conclusion

To summarise the overall perspective of this thesis, it is now worth recalling some of the main points from each chapter. Chapter 1 introduced the concept of evolutionary debunking arguments, along with a brief account of how such arguments rose to prominence. The next three chapters each focused on a particular prominent EDA. Chapter 2 criticised Sharon Street's argument, which primarily targets non-naturalistic versions of moral realism. One of the main points against her argument was the by-product view that could be used to vindicate moral realism, while also showing that her opposition to this view would undermine scientific theories, including her own evolutionary claims. Chapter 3 examined Richard Joyce's argument against the justification of moral beliefs in general, in which his conclusion was a form of global moral scepticism. One of the main issues was his reliance on a separate, non-evolutionary argument against moral naturalism, as well as his problematic 'belief pill' analogy which does not accurately represent our epistemic situation.

Chapter 4 focused on Joshua Greene's selective debunking argument against deontological moral theories. His argument had the advantage of drawing upon more empirical evidence rather than depending so heavily on untestable evolutionary speculations. However, the evidence was still insufficient to support his criticisms of deontology in general: Greene relied on a narrow conception of deontological philosophy along with various normative assumptions that he had not defended, such as the apparent

moral irrelevance of personal versus impersonal harm. Finally, the present chapter has examined issues that affect the standing of EDAs in general, specifically the plausibility and relevance of evolutionary premises, along with explanatory issues such as the relation between philosophical and scientific accounts of morality.

Despite all the problems with the discussed debunking arguments, it is important to reiterate what was mentioned at the outset in Chapter 1: the aim is not to rule out the possibility of any plausible EDAs being developed in the future. Rather, the main purpose has been to examine the flaws in the debunking arguments from Street, Joyce and Greene, which may allow future debunkers to learn from their mistakes. Any forthcoming EDAs and discussions of them should also take note of the more general issues surrounding these types of arguments, as discussed in this final chapter. After all, these general points could ultimately determine the prospects of EDAs in the long term.

References:

Audi, R 2014, *Naturalism, Normativity & Explanation*, Copernicus Center Press, Chicago.

Behrends, J 2013, 'Meta-Normative Realism, Evolution, and Our Reasons to Survive', *Pacific Philosophical Quarterly*, vol. 94, pp. 486-502.

Berker, S 2009, 'The Normative Insignificance of Neuroscience', *Philosophy & Public Affairs*, vol. 37, no. 4, pp. 293-329.

Braddock, M 2016, 'Evolutionary debunking: Can moral realists explain the reliability of our moral judgments?', *Philosophical Psychology*, vol. 29, no. 6, pp. 844-857.

Brink, DO 1989, *Moral Realism and the Foundations of Ethics*, Cambridge University Press, Cambridge.

Brink, DO 2001, 'Realism, Naturalism, and Moral Semantics', *Social Philosophy and Policy*, vol. 18, no. 2, pp. 154-176.

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

Bruni, T, Mameli, M & Rini, RA 2014, 'The Science of Morality and its Normative Implications', *Neuroethics*, vol. 7, pp. 159-172.

Buchanan, A & Powell, R 2015, 'The Limits of Evolutionary Explanations of Morality and Their Implications for Moral Progress', *Ethics*, vol. 126, no. 1, pp. 37-67.

Buller, D 2006, 'Evolutionary Psychology: A Critique', in E Sober (ed.), *Conceptual Issues in Evolutionary Biology*, MIT Press, Cambridge, MA, pp. 197-214.

Buller, D 2009, 'Four Fallacies of Pop Evolutionary Psychology', *Scientific American*, vol. 300, no. 1, pp. 74-81.

Carruthers, P & James, SM 2008, 'Evolution and the Possibility of Moral Realism', *Philosophy and Phenomenological Research*, vol. 77, no. 1, pp. 237-244.

Clarke-Doane, J 2012, 'Morality and Mathematics: The Evolutionary Challenge', *Ethics*, vol. 122, pp. 313-340.

Cline, B 2015, 'Nativism and the Evolutionary Debunking of Morality', *Review of Philosophy and Psychology*, vol. 6, no. 2, pp. 231-253.

Copp, D 1990, 'Explanation and Justification in Ethics', *Ethics*, vol. 100, no. 2, pp. 237-258.

Copp, D 2008, 'Darwinian Skepticism About Moral Realism', *Philosophical Issues*, vol. 18, pp. 186-206.

Crow, D 2016, 'Causal Impotence and Evolutionary Influence: Epistemological Challenges for Non-Naturalism', *Ethical Theory and Moral Practice*, vol. 19, no. 2, pp. 379-395.

Das, R 2016, 'Evolutionary debunking of morality: epistemological or metaphysical?', *Philosophical Studies*, vol. 173, no. 2, pp. 417-435.

Day, T & Kincaid, H 1994, 'Putting Inference to the Best Explanation in its Place', *Synthese*, vol. 98, no. 2, pp. 271-295.

Dean, R 2010, 'Does Neuroscience Undermine Deontological Theory?', *Neuroethics*, vol. 3, no. 1, pp. 43-60.

Deem, MJ 2016, 'Dehorning the Darwinian dilemma for normative realism', *Biology & Philosophy*, vol. 31, no. 5, pp. 727-746.

Enoch, D 2010, 'The epistemological challenge to metanormative realism: how best to understand it, and how to cope with it', *Philosophical Studies*, vol. 148, no. 3, pp. 413-438.

FitzPatrick, WJ 2014, *Morality and Evolutionary Biology*, The Stanford Encyclopedia of Philosophy, <https://plato.stanford.edu/entries/morality-biology/>.

FitzPatrick, WJ 2015, 'Debunking evolutionary debunking of ethical realism', *Philosophical Studies*, vol. 172, pp. 883-904.

FitzPatrick, WJ 2016, 'Scientific Naturalism and the Explanation of Moral Beliefs: Challenging Evolutionary Debunking', in KJ Clark (ed.), *The Blackwell Companion to Naturalism*, 1st edn, Wiley-Blackwell, Hoboken, pp. 386-400.

FitzPatrick, WJ 2017, 'Why Darwinism Does Not Debunk Objective Morality', in M Ruse & RJ Richards (eds), *The Cambridge Handbook of Evolutionary Ethics*, Cambridge University Press, Cambridge, pp. 188-201.

Greene, J 2008, 'The Secret Joke of Kant's Soul', in W Sinnott-Armstrong (ed.), *Moral Psychology, Volume 3: The Neuroscience of Morality: Emotion, Brain Disorders, and Development*, MIT Press, Cambridge, MA, pp. 35-79.

Greene, J 2014, 'Beyond Point-and-Shoot Morality: Why Cognitive (Neuro)Science Matters for Ethics', *Ethics*, vol. 124, no. 4, pp. 695-726.

Greene, J, Nystrom, LE, Engell, AD, Darley, JM & Cohen, JD 2004, 'The neural bases of cognitive conflict and control in moral judgment', *Neuron*, vol. 44, no. 2, pp. 389-400.

Greene, J, Sommerville, RB, Nystrom, LE, Darley, JM & Cohen, JD 2001, 'An fMRI Investigation of Emotional Engagement in Moral Judgment', *Science*, vol. 293, no. 5537, pp. 2105-2108.

Hales, SD 2007, 'Moral relativism and evolutionary psychology', *Synthese*, vol. 166, no. 2, pp. 431-447.

Hanson, L 2017, 'The Real Problem with Evolutionary Debunking Arguments', *The Philosophical Quarterly*, vol. 67, no. 268, pp. 508-533.

Harman, G 1977, *The Nature of Morality: An Introduction to Ethics*, Oxford University Press, New York.

Harman, G 1986, 'Moral Explanations of Natural Facts—Can Moral Claims Be Tested Against Moral Reality?', *The Southern Journal of Philosophy*, vol. 24, no. 5, pp. 57-68.

Hartman, R 1963, 'The Logical Difference Between Philosophy and Science', *Philosophy and Phenomenological Research*, vol. 23, no. 3, pp. 353-379.

Hopster, J 2018, 'Evolutionary arguments against moral realism: Why the empirical details matter (and which ones do)', *Biology & Philosophy*, vol. 33, no. 5, pp. 1-19.

Huemer, M 2016, 'A liberal realist answer to debunking skeptics: the empirical case for realism', *Philosophical Studies*, vol. 173, no. 7, pp. 1983-2010.

James, SM 2009, 'The Caveman's Conscience: Evolution and Moral Realism', *Australasian Journal of Philosophy*, vol. 87, no. 2, pp. 215-233.

Joyce, R 2000, 'Darwinian Ethics and Error', *Biology and Philosophy*, vol. 15, pp. 713-732.

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

Joyce, R 2013, 'Ethics and Evolution', in H LaFollette & I Persson (eds), *The Blackwell Guide to Ethical Theory*, Wiley-Blackwell, Hoboken, pp. 123-147.

Joyce, R 2014, 'The origins of moral judgment', *Behaviour*, vol. 151, pp. 261-278.

Joyce, R 2016, 'Reply: Confessions of a Modest Debunker', in UD Leibowitz & N Sinclair (eds), *Explanation in Ethics and Mathematics: Debunking and Dispensability*, Oxford University Press, Oxford, pp. 125-145.

Kahane, G 2011, 'Evolutionary Debunking Arguments', *Noûs*, vol. 45, no. 1, pp. 103-125.

Kahane, G 2012, 'On the Wrong Track: Process and Content in Moral Psychology', *Mind & Language*, vol. 27, no. 5, pp. 519-545.

Kahane, G & Shackel, N 2010, 'Methodological Issues in the Neuroscience of Moral Judgement', *Mind & Language*, vol. 25, no. 5, pp. 561-582.

Kahane, G, Wiech, K, Shackel, N, Farias, M, Savulescu, J & Tracey, I 2012, 'The neural basis of intuitive and counterintuitive moral judgment', *Social Cognitive and Affective Neuroscience*, vol. 7, no. 4, pp. 393-402.

Kelly, D 2011, *Yuck!: the Nature and Moral Significance of Disgust*, MIT Press, Cambridge, MA.

Kelly, D 2017, 'Moral Cheesecake, Evolved Psychology, and the Debunking Impulse', in R Joyce (ed.), *The Routledge Handbook of Evolution and Philosophy*, Routledge Press, New York, pp. 342-358.

Klenk, M 2017, 'Old Wine in New Bottles: Evolutionary Debunking Arguments and the Benacerraf-Field Challenge', *Ethical Theory and Moral Practice*, vol. 20, no. 4, pp. 781-795.

Kumar, V & Campbell, R 2012, 'On the normative significance of experimental moral psychology', *Philosophical Psychology*, vol. 25, no. 3, pp. 311-330.

Leibowitz, UD & Sinclair, N 2017, 'Evolution and the Missing Link (in Debunking Arguments)', in M Ruse & RJ Richards (eds), *The Cambridge Handbook of Evolutionary Ethics*, Cambridge University Press, Cambridge, pp. 210-225.

Majors, B 2003, 'Moral Explanation and the Special Sciences', *Philosophical Studies*, vol. 113, pp. 121-152.

Mason, K 2010, 'Debunking Arguments and the Genealogy of Religion and Morality', *Philosophy Compass*, vol. 5, no. 9, pp. 770-778.

Mogensen, A 2016, 'Do evolutionary debunking arguments rest on a mistake about evolutionary explanations?', *Philosophical Studies*, vol. 173, no. 7, pp. 1799-1817.

O'Neill, E 2015, 'Which Causes of Moral Beliefs Matter?', *Philosophy of Science*, vol. 82, pp. 1070-1080.

Owens, D 1989, 'Levels of Explanation', *Mind*, vol. 98, no. 389, pp. 59-79.

Ruse, M & Wilson, EO 1986, 'Moral Philosophy as Applied Science', *Philosophy*, vol. 61, no. 236, pp. 173-192.

Sayre-McCord, G 1988, 'Moral Theory and Explanatory Impotence', *Midwest Studies in Philosophy*, vol. 12, no. 1, pp. 433-457.

Sayre-McCord, G 2007, 'Moral Realism', in D Copp (ed.), *The Oxford Handbook of Ethical Theory*, Oxford University Press, Oxford.

Shafer-Landau, R 2013, 'Evolutionary Debunking, Moral Realism and Moral Knowledge', *Journal of Ethics & Social Philosophy*, vol. 7, no. 1, pp. 1-37.

Shafer-Landau, R 2017, 'Moral Realism and Evolutionary Debunking Arguments', in M Ruse & RJ Richards (eds), *The Cambridge Handbook of Evolutionary Ethics*, Cambridge University Press, Cambridge, pp. 175-187.

Singer, P 2005, 'Ethics and Intuitions', *The Journal of Ethics*, vol. 9, no. 3, pp. 331-352.

Singer, P 2011, *The Expanding Circle*, Princeton University Press, Princeton, NJ.

Skarsaune, KO 2011, 'Darwin and moral realism: survival of the iffiest', *Philosophical Studies*, vol. 152, no. 2, pp. 229-243.

Street, S 2006, 'A Darwinian Dilemma for Realist Theories of Value', *Philosophical Studies*, vol. 127, no. 1, pp. 109-166.

Sturgeon, N 1985, 'Moral Explanations', in D Copp & D Zimmerman (eds), *Morality, Reason and Truth: New Essays on the Foundations of Ethics*, Rowman and Allanheld, Totowa, NJ, pp. 49-78.

Sturgeon, N 1992, 'Nonmoral Explanations', *Philosophical Perspectives*, vol. 6, pp. 97-117.

Tersman, F 2008, 'The Reliability of Moral Intuitions: A Challenge From Neuroscience', *Australasian Journal of Philosophy*, vol. 86, no. 3, pp. 389-405.

Toner, C 2011, 'Evolution, Naturalism, and the Worthwhile: A Critique of Richard Joyce's Evolutionary Debunking of Morality', *Metaphilosophy*, vol. 42, no. 4, pp. 520-546.

Tresan, J 2010, 'Question authority: in defense of moral naturalism without clout', *Philosophical Studies*, vol. 150, no. 2, pp. 221-238.

Tropman, E 2014, 'Evolutionary debunking arguments: moral realism, constructivism, and explaining moral knowledge', *Philosophical Explorations*, vol. 17, no. 2, pp. 126-140.

Vavova, K 2015, 'Evolutionary Debunking of Moral Realism', *Philosophy Compass*, vol. 10, no. 2, pp. 104-116.

White, R 2010, 'You Just Believe That Because…', *Philosophical Perspectives*, vol. 24, pp. 573-615.

Wielenberg, E 2010, 'On the Evolutionary Debunking of Morality', *Ethics*, vol. 120, pp. 441-464.