EVOLUTIONARY DEBUNKING: CAN MORAL REALISM RESIST THE EPISTEMOLOGICAL CHALLENGE?

By

Anders Yuk Pui Lam

Submitted to

Department of Philosophy

Central European University

In partial fulfillment of the requirements for the degree of Masters of Arts

Supervisor: Prof. Simon Rippon

Budapest, Hungary

2017

Abstract

Evolutionary debunking arguments (EDAs) against moral realism are arguments which appeal to the alleged success of the evolutionary explanation of the emergence and the content of human moral thoughts. The goal of EDAs is not to prove that morality can be given such an explanation but to show a certain kind of implications that this explanation would have for moral philosophy. In this thesis, I will firstly give an analysis of the evolutionary approach assumed by antirealist debunkers, explaining what it means to say that the Darwinian account of morality is at least roughly true. Then I shall illustrate how the debunkers infer anti-realist conclusions from two different types of debunking arguments: metaphysical and epistemological. After bringing in the context of the debunking debate, I shall examine prominent realist solutions to the epistemological type of debunking challenge. After giving a critique of each of these solutions, I conclude that none of them has successfully defended moral realism against the Darwinian challenge.

Contents

Introduction

What implications can scientific discoveries have for philosophy? A few philosophers have suggested that the evolutionary perspective on human morality, if empirically confirmed, has a debunking effect on a metaethcial theory, namely moral realism. Realism in general is a sort of theory which insists on the objective or independent existence of certain properties, truths and facts.[1] According to moral realism, there are objective moral truths or moral facts that are independent of our subjective evaluative attitudes.[2] The evolutionary debunker argues that the Darwinian origins of and influences on moral beliefs undermine such a response-independent theory.

The common premise shared by all debunking arguments is that evolutionary forces have given rise to and tremendously influenced our moral beliefs. This scientific claim, though controversial, is adopted as an assumption in the debunking debate. The issue to be addressed in this context is whether moral realism is debunked *if* such a Darwinian account of morality is true or at least roughly correct. Accordingly, moral realists are in

[1] Here I use the terms "objective" and "independent" interchangeably, both meaning "not relying on our subjective responses".

[2] The set of evaluative attitude should include desires, attitudes of approval and disapproval and evaluative judgments that something is good or bad and normative (and moral) judgments that we ought to do certain things. The term "evaluative-attitudes-independent" is equivalent to "stance-independent" and "response-independent". I shall avoid using the term "mind-independent" because it seems to suggest a strong independence from any being which has a mind, and I think this might be too strong for classifying realism.

the face of an empirically grounded challenge but not a theoretic challenge which merely appeals to imaginary or theoretically possible circumstances.

The aim of this thesis is to investigate whether realists can cope with this kind of challenge. In chapter I, I will explain the Darwinian Hypothesis by clarifying what it means to say that the evolutionary approach *explains* morality and Darwinian forces have played a significant role in shaping human moral thoughts. In Chapter II, I will distinguish and illustrate two different kinds of debunking challenges. One is the metaphysical argument which concludes that independent moral truths posited by realists do not exist. The other is the epistemological argument which concludes that moral realism is associated with unsatisfactory skepticism that we do not possess moral knowledge. In Chapter III and IV, I shall examine a number of realist solutions to the epistemological debunking challenge while setting aside metaphysical debunking. The conclusion of this thesis is that the realist's responses fail to resist the epistemological debunking.

Chapter I: The Darwinian Hypothesis of Morality

The Darwinian Hypothesis of morality is the hypothesis that the evolutionary approach explains morality. This hypothesis is adopted by the antirealist debunkers as the first premise of their evolutionary arguments against moral realism. The success of the hypothesis, though debatable at the level of science and philosophy of science, is an assumption of the debunking debate. The whole debate is centered on the conditional: "if the Darwinian Hypothesis that human morality can be explained in evolutionary terms is at least roughly true, then moral realism would be undermined". This debunking strategy can also be applied to undermining other metaethical theories such as constructivism and expressivism, but here I will focus on the debunking arguments against moral realism.[3]

Before entering the debunking debate, I shall summarize and clarify what the Darwinian Hypothesis does and does not say, starting with introducing the historical background of the evolutionary approach which attempts to explain human psychology. As mentioned above, the Darwinian Hypothesis can be summarized as the thesis that evolutionary psychology explains human morality. But this statement is too rough and vague to be a satisfactory summary. So, the key issue which will be addressed in this

[3] For the evolutionary debunking argument against constructivism, see Tropman (2014). For the evolutionary debunking argument against quasi-realist expressivism, see Street (2011).

section is what it literally means to say that evolutionary psychology *explains* human morality in the context of the Darwinian Hypothesis adopted by the debunkers. Before clarifying the hypothesis, I shall introduce the historical background of the development of evolutionary psychology and how the evolutionary psychologists have shed new light on moral psychology.

1(a) The historical background of evolutionary psychology

Evolutionary psychology is a relatively new discipline which has been developed since the 1970s. An important work in the field is *Sociobiology: The New Synthesis*, published in 1975 by Harvard biologist Edward Wilson, in which he gives explanations of animal behaviors from the evolutionary perspective. Nevertheless, his work is not just about social behaviors of non-human animals but also human. In the last chapter, "Man: From Sociobiology to Sociology", Wilson extends the approach to human social behavior and declares the need for the Darwinian explanation of human sociology and psychology. This attempt of digging out the biological basis of social behaviors and extending population biology and evolutionary theory to social organization became the goal of the following trend of evolutionary psychology which was foreseen by Charles Darwin in *The Origin of Species* (1866, 576).

The scholars involved in the project, for example the best representatives Leda Cosmides, John Tooby and Steven Pinker, have attempted to throw new light on human psychology by adopting an evolutionary perspective. The main task of evolutionary psychology is to identify which human psychological traits are the functional products of natural selection. As Cosmides and Tooby (1997, 85) claim, "Our modern skulls house a stone age mind". Evolutionary psychologists trace human psychological traits back to their evolutionary history. The principal areas of research have targeted many crucial elements of human nature including consciousness, sensation, learning, emotion, motive, language, mating, parenting, reciprocity and altruism. Moral psychology, as an important part of human nature which lies in the intersections of above parts, has also become an explanandum of the project in more recent times.

1(b) The evolutionary explanation of moral psychology

A psychological trait being an adaptation means that having such trait or a tendency to have such trait gave its bearers, namely our ancestors, a selective advantage in their living environment, namely environment of evolutionary adaptedness (EEA). Selective advantage brought by a psychological trait, roughly, means that because of that trait the bearers had more offspring in successive generations than those who did not have this trait, so that the fitness of such bearers is higher than their competitors. If such

biological history can be spelled out with adequate scientific evidence, then we can say that the psychological trait in question is *explained* by the evolutionary approach. For example, fear of snake might be reasonably regarded as an adaptive trait because such a psychological trait prevents dangerous contact with animals like King Cobra and thus provides a selective advantage. As for human moral psychology, the evolutionary psychologist attempts to prove that not necessarily all but quite a few moral psychological traits are biological adaptations. The targeted explananda have ranged from moral emotion, disposition and attitude to descriptive judgment. The main task of the project is to identify which of these psychological traits are the functional products of natural selection.

When the thinking of evolutionary psychology is applied to morality, it comes to *the Darwinian Hypothesis* of the emergence and development of human moral psychology. Accordingly, the evolutionary approach greatly explains the origin and phenomena of moral psychology including what we morally feel, think, judge, act and react. Altruism is one of the examples often discussed by the relevant theorists. Psychological altruism is defined by the intention to benefit any other individuals as an ultimate goal at the expense of one's own interests. The adaptive explanation of altruistic behavior and corresponding evaluative judgment is that the selective forces, which might include

mutualism, reciprocity, kin selection and group selection, resulted in the phenomena that people perform and praise altruistic acts.[4] Likewise, similar moral practices can be explained by the very same Darwinian mechanisms. If the whole moral system can be given such evolutionary explanation, then it could be said that the emergence and development of morality are biological adaptations.

Some of the moral thoughts in the moral system we have can be pretty straightforwardly and plausibly explained by the Darwinian method. For instance, it seems to be a universal phenomenon that people regard taking good care of one's own children as an important virtue and harming them as a great evil. The evolutionary explanation of the psychological tendency to do good to offspring is that our ancestors who had this psychological trait ended up with having more babies and more flourishing clans in successive generations than their competitors who did not have the same trait. Moreover, we normally regard incest as immoral and feel disgusted with such behavior. The evolutionary account smoothly provides a story of why we think and feel so given that incest is highly disadvantageous in terms of biological fitness. More generally, it seems

[4] Notice that apart from psychological altruism, altruism can be defined differently as evolutionary altruism according to which A is acting altruistically if and only if the effect of A's behavior is an increase of some other organism's fitness at the expense of its own fitness. Here we can an evolutionary paradox of altruism simply because psychological altruism seems to overlap with or lead to evolutionary altruism but evolutionary altruism is selectively disadvantageous. So, it would be problematic to say that altruism is a product of natural selection. Various solutions have been proposed to solve this paradox. I shall not invoke the debate here and I shall assume that the paradox is solved and altruism can be a product of natural selection.

that many psychological features in human morality could have certain connection with selective pressure and history of natural selection. Thus, explaining the whole moral system in Darwinian terms appears to be a feasible project.

However, there is a difficulty if the psychologist's task is to explain the *whole* moral psychology. Although quite a number of moral thoughts appear to be evolutionary-account-friendly, not all of them are. It might be hard to see the connection between quite a few of moral thoughts and selective pressure, especially thoughts which are relatively complicated and generated through many years of rational reflection. For example, nowadays many people think that it is wrong to restrict freedom of speech or academic freedom, perhaps by making unjust new law to shut down academic institutions. Nonetheless, it is very unobvious how such belief would promote selective advantage. Nor does it seem possible to trace the belief back to our ancestral evolution because people who lived in the Stone Age and before modern times did not even have such a concept. Apart from unobvious connection, some moral thoughts even seem to conflict with the evolutionary principle – "survival of the fittest".[5] Sometimes behaviors which are normally regarded as immoral, such as rape, could appear to be selectively advantageous in certain circumstances. Cases of moral practices which

[5] "Survival of the fittest" is a phrase used by British economist Herbert Spencer to describe the mechanism of natural selection in the Darwinian framework, see Spencer (1864, 444)

appear to have nothing to do with evolution and other practices which appear to go against selective forces should not be neglected. So how could advocates of the Darwinian Hypothesis, given such difficulty, say that this approach can explain human morality as a whole?

1(c) Clarifications

The Darwinian Hypothesis claims that the evolutionary approach explains morality. However, phenomena of moral psychology which do not seem to comply with the evolutionary approach could render the hypothesis problematic. If so, the evolutionary debunker would be relying on a suspect starting point and the debunking arguments would be groundless. It might help to clear away the above explanatory difficulty for the Darwinian Hypothesis by clarifying what it does and does not say and imply.

First, the Darwinian Hypothesis does assert that at least *some* moral thoughts can be straightforwardly explained by the principle of "survival of the fittest". Nevertheless, the hypothesis does not imply that *all* moral thoughts were directly produced by natural selection such that every moral thought would be connected to adaptiveness in a straightforward manner. The claim that the evolutionary approach explains morality

only means that evolutionary forces have heavily influenced our moral psychology.[6] That is not to deny that some parts of human morality cannot be explained in terms of adaptiveness. Still, for those who attempt to assert that evolutionary forces have an overall influence on moral psychology, the problem remains. If there are moral thoughts that cannot be straightforwardly explained in the form: "MT (a certain moral thought) emerged through the process of natural selection because it provided selective advantages in ancestral environment", then why should we not simply say that the moral thoughts in question are *not* under the influence of evolutionary forces? Accordingly, there is still the explanatory problem for those evolutionary psychologists and the debunkers who insist on the overall impact of natural selection. These theorists have to explain more about how evolutionary forces could comprehensively influence moral psychology including the moral thoughts which do not seem to bear adaptiveness.

Indeed, the utilitarian Peter Singer (2005) utilizes of the idea that some moral thoughts, namely the utilitarian ones according to Singer, were generated by a mechanism which is immune to or untouched by evolutionary forces, namely reasoning.[7] If there are a set of moral beliefs that are separate from those which are straightforwardly adaptive, then of course, I think, the debunkers would face a huge difficulty of explaining why the

[6] Here *influence* means the extent to which evolutionary forces have given rise to the appearance of moral psychology and have shaped its content.

[7] For his argument which is primarily directed against non-utilitarian claims, see Singer (2005)

truth of the Darwinian Hypothesis still promotes a challenge to those which seem evolutionarily irrelevant or immune. Suppose there are two distinct classes of moral thoughts: one of which is heavily influenced by evolution, the other is not. In other words, evolutionary forces, which have influenced the first class of moral thoughts, have played no role in giving rise to and in shaping the content of the second class of moral thoughts. Accordingly, the justification of the latter might well be intact even if the former is completely debunked by the truth of Darwinian Hypothesis. If so, the realist can simply reply that the evolutionary story does not provide reason for believing that the moral beliefs untouched by evolution are nonetheless unjustified or false.

Details about how the *evolutionary* and the *non-evolutionary* set of moral beliefs are bound together can vary from one debunker to another.[8] However, at least all of them would have to assert that the content of the non-evolutionary set is restricted by the content of the evolutionary set or they are generated by the same set of capacities which are products of natural selection. As a result, if evolutionary forces have had a tremendous influence on the evolutionary set of moral beliefs, they have also greatly affected the non-evolutionary set of moral beliefs. Here I shall distinguish two types

[8] According to my usage, here "the non-evolutionary set of moral beliefs" simply means "the set of moral beliefs which cannot be given straightforward adaptive explanations" or "does not seem to bear adaptiveness", it does not mean that the non-evolutionary set is completely detached from and/or unlimited by the evolutionary process. Likewise, "the evolutionary set of moral beliefs" means "the set of moral beliefs which can be given straightforward adaptive explanations"

(there could be more) of assertions which are made in order to safeguard the claim that our moral psychology is comprehensively influenced by evolutionary forces. The first type is the *capacity* claim, as I shall name it, that natural selection produced not particular moral thoughts or dispositions but a general capacity or a set of capacities to make moral judgment in various environments. The second type is the *raw material* claim that evolution produced the basic material from which the whole moral psychology is derived. According to the raw material claim, we can analogically say that if our moral psychology is a skyscraper, then natural selection is the factory which has selected and provided building materials from which the skyscraper has been built. It should be noted that the capacity claim and the raw material claim are not mutually exclusive. In other words, it could be the case that natural selection gave rise to both the general capacity to make moral judgment and basic moral thoughts as construction material.

According to the above two claims, when evolutionary forces provided foundations of moral thoughts, they also shaped the boundaries of their development. Analogically speaking, the ability to perceive and the ability to calculate can be reasonably regarded as adaptations. The ability to find out how fast a butterfly flies seems to be based upon these two abilities. Although it is hard to see how knowing the butterfly's speed can be

selectively beneficial, this ability still bear an indirect relation to evolutionary forces. Thus, given the ability claim, it seems reasonable to say that the ability and what it produces are influenced or restricted by evolutionary forces. As for the raw material claim, evolution provided initial moral thoughts like emotion, tendency and intuition as input on which psychological mechanisms operate.[9] Since the input is evolution-laden, the output is also restricted by the evolutionary forces to a certain extent. In short, no matter what the evolutionary forces shaped as the starting point of our moral psychology (ability and/or innate material), the rest of moral psychology is neither completely detached from the biological basis nor free from the barriers created by Darwinian forces.

So far, we have already got the first clarification of the Darwinian Hypothesis: although it says that the evolutionary approach explains morality, it does not imply that every single moral thought is a literal product of natural selection. Some of them are and some of them are not. Nonetheless, all of them are either directly or indirectly influenced by evolutionary forces. Therefore, the evolutionary account can explain morality not because every moral thought is a direct result of the selective process, but because evolutionary forces have played a crucial role in shaping all these thoughts. An analogy

[9] Here I use the word "intuition" as equivalent to "intuitive judgment or belief", so it does not mean "the capacity to intuit".

should make this point clearer. Imagine you are seriously sick and have shown a lot of symptoms of sickness. Your doctor tells you that an invasion of certain virus *explains* your sickness which is identical to the collection of your symptoms. Does the doctor's claim imply that every single symptom is a direct product of the viral infection? Of course not. Presumably, there could be some symptoms which are indirect products of the infection, i.e. complications. However, this does not mean that the occurrence of complication is detached from the viral invasion.

Now, given the above inference, the Darwinian Hypothesis would seem compatible with the phenomena that some moral thoughts have no adaptiveness (do not have clear connection to selective advantage). But how about those moral thoughts the existence of which seems to go *against* the evolutionary because it can be selectively disadvantageous to have these thoughts? For example, the thought that we morally ought to take care of disabled people or other socially vulnerable groups at the cost of consuming some social resources could be selectively disadvantageous. However, this thought could also be bound to the evolutionary set of moral thoughts in the sense of relying on some native tendencies which has been produced by natural selection, for instance the tendency to come up with the thought that pain is bad and other basic *evaluative* thoughts. The story of how these non-evolutionary thoughts have been

derived from evolutionary thoughts and/or capacities can be very complicated, but it should not prevent the Darwinian Hypothesis from insisting on the binding force of evolved thoughts and/or capacities. Now it should be safe to conclude that although not every moral thought has an adaptive explanation, this does not undermine the claim that evolutionary forces have a comprehensive impact on moral psychology, given the assertion that the non-evolutionary set of moral thoughts is bound to the evolutionary set.

But even if we grant that the Darwinian Hypothesis is correct that evolutionary forces have a comprehensive impact on moral psychology, the hypothesis might have another problem. Such a hypothesis might preclude other factors which are of importance in shaping our moral psychology, such as cultural influence, social factor and rational reflection. However, it is a mistake to say that the Darwinian Hypothesis suggests that morality is fully determined by biological evolution and other factors like cultural influence, social factor and rational reflection are absent in shaping human moral psychology. Just like the retirement of the traditional nature-versus-nurture debate, there is no clear-cut answer to whether moral thoughts are inherited or acquired. The Darwinian Hypothesis only asserts that our moral psychology is heavily but not exclusively influenced by evolutionary forces. Therefore, the hypothesis leaves room

for these factors to come into play in moral psychology.

1(d) Summary of the Darwinian Hypothesis

The Darwinian Hypothesis is deployed by the evolutionary debunkers to attack moral realism. As I define it, the hypothesis states that evolutionary forces have played a significant role in shaping moral psychology as a whole in the sense that the relevant psychological traits are heavily influenced by the evolutionary forces, either in a direct or indirect manner. Not necessarily all evolutionary psychologists have to hold such hypothesis because some of them might be willing to admit that the evolutionary impact on human morality is not comprehensive in the sense that Darwinian forces are totally absent (or at least negligible) in bringing about and shaping a certain group of moral thoughts. However, for psychologists who wish to say that no part of moral psychology is completely detached from natural selection and for the evolutionary anti-realist whose task is to debunk *all* moral thoughts, they must insist the *overall* impact of Darwinian forces.

In the rest of this thesis, I will not propose empirical arguments associated with empirical data collected by the scientists as my goal here is not to prove that the Darwinian Hypothesis is true. Nonetheless, I would like to assume that it is at least a

promising theory, otherwise my project and the whole debunking debate would not be worthwhile. There are two traditional criticisms of the hypothesis and evolutionary psychology in general. One is that the evolutionary account is a just-so-story which does not having any empirical grounding, whereas the other is that such account is not scientifically testable. I shall assume that both criticisms are not convincing and the evolutionary explanation should be taken seriously as a testable scientific theory which has at least some positive empirical evidence.

Thus far, I will only assume this much without entering the empirical and methodological debates on the plausibility of the Darwinian Hypothesis. The aim of this thesis is not to discuss whether the Darwinian Hypothesis is empirically plausible but to examine the metaethical implication of the hypothesis. The following chapters will focus on the debunking debate on whether such a hypothesis, if successful, would undermine moral realism. In chapter II, I shall introduce three prominent debunking arguments against moral realism. In chapter III and chapter IV, I will examine responses proposed by the non-naturalist realist and the naturalist realist.

Chapter II: Debunking Arguments Against Moral Realism

Evolutionary debunking arguments are in the form of the conditional: if the Darwinian Hypothesis of Morality is at least roughly true, then moral realism would be undermined. Nonetheless, there is more than one way a theory could be undermined. For example, Alvin Plantinga, in *Warranted Christian Belief* (2000), distinguishes two kinds of challenges to Christian beliefs. The first is categorized as *de facto* argument which concludes that Christian beliefs are simply not true. The second is categorized as *de jure* argument which concludes that Christian beliefs, whether true or not, are not justified. Likewise, the debunking arguments in metaethics can also be thus distinguished. We have the *de facto* one - the *metaphysical* argument aiming at the conclusion that, because of the success of the evolutionary explanation of morality, none of our moral beliefs can be true. We also have the *de jure* one – the *epistemological* argument aiming at the conclusion that our moral beliefs, even though they could be true, are not justified (and thus cannot be knowledge) given the truth of the Darwinian Hypothesis. I shall now introduce both versions of debunking.

2(a) Metaphysical debunking – Michael Ruse

The representative of metaphysical debunking is Michael Ruse. Ruse (1986) thinks that the evolutionary story, which does not assert the existence of objective moral truth, is

all that is needed to explain moral psychology. He argues that

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

And he adds,

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

The conclusion he makes seems very nihilist, declaring that "morality is a collective illusion foisted upon us by our genes." (1986, 253) To claim that something is an *illusion* is not merely saying that our beliefs about that thing are easily mistaken or unjustified, but also saying that the thing in which we believe does not even exist; that is, it is not real at all. For me, Ruse's debunking argument can be interpreted as an extension of Gilbert Harman's (1977) argument against the existence of objective moral facts.

An analogy Ruse uses to demonstrate the metaphysical argument is the following. In order to explain why a grieving mother believes that her son spoke to her at a séance, all we need to appeal to are psychological facts about her wishful thinking and vulnerable state of mind. There is no need to assume that she was really in contact with a supernatural being, namely her son's soul. In Ruse's metaphysical debunking, the existence of objective moral truth, just like the son's soul, is metaphysically redundant in explaining observable phenomena. According to the principle of parsimony of Ockham's razor, it would be reasonable to conclude that there is no such truth. I think that his argument can be summarized as follows:

R1 The Darwinian theory without assertions about objective moral truths successfully explains moral psychological phenomena. [The Darwinian Hypothesis]

R2 There is no need to appeal to these moral truths to explain any other phenomena.

R3 ∴ Objective moral truths are not needed in any explanation of anything. [R1&R2]

R4 If something's existence is metaphysically redundant in explaining things, then it does not exist or at least we have good reason to believe it does not.

R5 ∴ Objective moral truths do not exist. [R3&R4]

If Ruse's metaphysical argument succeeds, there is a direct defeater to moral realism which insists the existence of objective moral truths. In other words, if the above

argument is sound, then moral realism will be literally rendered false. However, there is another type of debunking arguments – the epistemological debunking - which is weaker than Ruse's metaphysical debunking in the sense that the epistemological debunkers do not plan to strictly render the falsehood of moral realism, but rather plan to bind moral realism to unfavorable skepticism.

2(b) Epistemological debunking – Richard Joyce

Two prominent versions of epistemological debunking are proposed by Richard Joyce (2006) and Sharon Street (2006) respectively. Both versions have targeted on the justificatory status of moral beliefs which are influenced by evolutionary forces. Joyce (2013) claims that his argument is best approached via the following thought experiment mentioned in his previous book, *The Evolution of Morality* (2006). Suppose you find out that your belief that Napoleon lost the battle of Waterloo simply because you unknowingly took a Napoleon-lost-Waterloo belief pill which makes you believe that the King of France did lost the battle no matter whether it is true or not. Joyce suggests that the reasonable thing to do to cast doubt on the justificatory status of your belief and look for some truth-tracking connection between your belief and the historical truth about the King losing the battle. If no such story is discovered, one should remain doubtful and conclude that the belief is not knowledge. Applying the

thought experiment to morality, Joyce argues that the belief pill is analogous to the Darwinian forces because both can effectively yield beliefs without any connection to relevant truths. Just like the unjustified belief about Waterloo, moral beliefs are also unjustified.

Here is Joyce's own summary of his justification-debunking argument.

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

I shall reconstruct Joyce's argument as follows:

J1 Moral beliefs are explained by the genealogy which does not presuppose or imply any kind (naturalist, non-naturalist or supernaturalist) of moral truths. [J1.1&J1.2&J1.3&IBE]

J1.1 Explanations which assert non-natural or supernatural moral truths are explanatorily superfluous, given the presence of the genealogical explanation.

J1.2 Explanations which assert naturalist moral truths cannot account for the

intrinsically-reason-giving essence of morality – the practical clout – and thus are explanatorily inadequate.

J1.3 The genealogical explanation which does not presuppose or imply moral truths adequately explains morality (including moral beliefs). [The Darwinian Hypothesis]

J2 If moral beliefs are explained by genealogy which does not presuppose or imply moral truths, then moral beliefs are explained without appealing to realist moral truths.

J3 ∴ Moral beliefs are explained without appealing to realist moral truths. [J1&J2]

J4 If moral beliefs are explained without appealing to moral truths, then they products of a process which is independent of or insensitive to their truth.

J5 If moral beliefs are products of a process which is independent of or insensitive to their truth, then they are unjustified.

J6 ∴ Moral beliefs are unjustified. [J3&J4&J5]

J7 If moral beliefs are unjustified, then we do not have moral knowledge.

J8 ∴ We do not have moral knowledge. [J6&J7]

Joyce points out that we have three available explanations of moral beliefs on the table. The first is the purely causal genealogy asserted by the Darwinian Hypothesis which does not presuppose or imply independent moral truths. Second, moral naturalists can propose the naturalist explanation which allows moral facts to be "implicitly buried in a scientific genealogical hypothesis" (Joyce 2006, 209) Third, realists who are not naturalists can propose explanations which assert non-naturalists and supernaturalist

moral truths. Then Joyce starts to close down the latter two options by comparing them with the genealogical account. He argues (see premise J1.1) that non-naturalist or supernaturalist options are eliminated on grounds of parsimony because "non-naturalism and supernaturalism do posit extra ontology in the world, but the presence of the non-moral genealogy shows this ontology to be explanatorily superfluous." (2006, 210) As for naturalism, Joyce admits that the naturalist explanation, which asserts a genuine kind of reduction to the natural facts "invoked in the genealogical explanation", "cannot be eliminated on grounds of parsimony, any more than cats should be eliminated from our ontology because we can explain them in terms of physics." (2006, 189) However, he argues that the naturalist explanation is inadequate because it fails to account for an essential feature of morality – practical clout (see premise J1.2).[10]

It should be noted that premise J1.2 is supported by an independent philosophical argument against moral naturalism which claims that the practical clout is an essential feature of morality and naturalism fails to account for such feature. First, the argument is not driven by the Darwinian Hypothesis and thus is not an evolutionary argument. Second, it would take too long, presumably the length of another thesis, to settle the issue whether practical clout really is an essential feature of morality and the issue

[10] For more about practical clout, see Joyce 2006, 190-209

whether naturalism really fails to account for it. For these two reasons, the discussion of premise J1.2 and Joyce's independent philosophical argument is beyond the scope of my thesis which focuses on evolutionary debunking, so I shall put the discussion aside.

From the comparison between the three explanations and the criteria of inference to the best explanation (IBE) including parsimony and explanatory power, Joyce infers that the non-moral genealogical explanation is the best one among the three in question. The rest of argument is spelled out by the following quote, which is Joyce's own summary on his argument:

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

As an epistemological debunker, Joyce concludes that assuming the genealogical story of moral beliefs is true, our moral beliefs are not justified and thus are not knowledge.

Here we can see the highlighting difference between metaphysical debunking and epistemological debunking. Ruse's conclusion is that there are no moral truths, implying that all realist moral beliefs should be false, whereas Joyce explicitly, in his debunking argument, retains the possibility that our moral beliefs might be true.

2(c) Epistemological debunking – Sharon Street

Let us now turn to a similar argument, another epistemological debunking, proposed by Sharon Street (2006) which is named the Darwinian Dilemma. I see the argument running as follows:

S1 If moral realism is true and moral beliefs are justified, then the connection between moral beliefs and realist moral truths must be more than sheer coincidence.

S2 Any account that the realist uses to establish a connection more than sheer coincidence, namely tracking, would be less parsimonious, less clear and less explanatorily powerful than the Darwinian hypothesis.

S3 ∴ According to inference to the best explanation, tracking is false. [S2, IBE]

S4 If tracking is false, then the connection between moral beliefs and realist moral truths cannot be more than sheer coincidence.

S5 ∴ The connection between moral beliefs and realist moral truths cannot be more than sheer coincidence. [S3,S4]

S6 ∴ Either moral realism is false or our moral beliefs are not justified.[S1,S5]

Premise 1 is a common or intuitive epistemological assumption. Suppose there is an objective truth about the number of people in the classroom which is, let's say, 10. The lecturer looked at her phone and the first number appeared randomly on the screen was 10. She nonetheless decided to believe that the number of people in the classroom is 10 because of this random experience. Now even if her belief happens to be true, it is not regarded as knowledge because the connection between the belief and the truth is sheer coincidence. Her belief can be justified only if there is a tracking relation between her belief and the truth; for example, she could have asked her students to report the number or counted the heads herself.

For premise 2, Street writes

> "We now have rough sketches of two competing evolutionary accounts of why we tend to make some evaluative judgments rather than others. For reasons that may already have begun to suggest themselves, I believe that the adaptive link account wins this competition hands down, as judged by all the usual criteria of scientific adequacy. In particular, there are at least three respects in which the adaptive link account is superior to the tracking account: it is more parsimonious; it is much clearer; and it sheds much more light on the explanandum in question, namely why human beings tend to make some

evaluative judgments rather than others." (2006, 129)

As for parsimony, Street argues that "the tracking account obviously posits something extra that the adaptive link account does not, namely independent evaluative truths". (2006, 129) As for clearness, she claims that it is unclear how grasping the independent evaluative truths can promote an organism's reproductive success. As for explanatory power, she states three sorts of things that cannot be well explained by the tracking account. The first is "the remarkable coincidence that so many of the truths it posits turn out to be exactly the same judgments". (2006, 132) The second is "our observed predispositions to make other evaluative judgments which (we may decide on reflection) are not true" (2006, 133), for example the tendency to treat in-group and out-group people distinctively. The third is "those normative judgments that human beings could make but don't" (2006, 133) such as the judgmen4t that infanticide is laudable. She concludes that "the tracking account has nothing comparably informative to say. It can just stand by and insist that such judgments are false—reaffirming our convictions but adding nothing to our understanding of why we have them." (2006, 133)

The conclusion of Street's argument, as it also belongs to the epistemological type of debunking, is that our moral beliefs are not justified under the framework of moral

realism. This conclusion, according to my understanding, is shared by Joyce's debunking argument. Apart from the conclusion, there are more things that they have in common. First, both arguments deploy criteria of inference to the explanation, namely parsimony (simplicity) and explanatory power, to eliminate possible realist explanations. We can see the application of these criteria in the inferences from premises J1.1, J1.2 and J1.3 to premise J1 in Joyce's argument and from premise S2 to S3 in Street's argument. Moreover, both debunkers hold that the best explanation of our moral beliefs, which has adequate explanatory power and enjoys simplicity, is that they are products of an evolutionary mechanism insensitive to independent moral truths.

The main difference between Joyce and Street, I think, is how they deal with naturalism. As mentioned above, Joyce proposes an independent philosophical argument, which has nothing to do with the Darwinian Hypothesis, to reject possible naturalist replies in the first place. Street's strategy is to firstly distinguish two possible readings of the naturalist reply: 1. morality/reasons internalist and 2. morality/reasons externalist. Street then argues that the naturalist reply, on the internalist reading, is question-begging, whereas the externalist version is not actually targeted by her Darwinian Dilemma which is primarily directed against realism about *normativity* but not *morality*. In chapter IV, I shall discuss her reply to the naturalist reply in more detail.

Another important difference between two debunkers is beyond the content of their debunking arguments reconstructed above. Recall that they share the conclusion that moral beliefs are unjustified under the framework of moral realism. As an error-theorist, Joyce endorses the additional connecting premise that ordinary moral beliefs and moral semantics are essentially realist in the sense that they are pointing to realist moral facts. Accordingly, if the justificatory status of realist moral beliefs and judgments is debunked by evolutionary arguments, then no justification of moral beliefs and judgments will be left. With this connecting premise, Joyce's debunking argument, if succeeds, will lead to what he (2015) categorizes as evolutionary debunking of *moral justification*. In contrast, Street does not accept the connecting premise above as she does not think that moral beliefs and semantics are essentially pointing to independent moral truths. As a constructivist, she holds that morality are about evaluative-attitude-dependent moral truths. Accordingly, even if the justificatory status of realist moral beliefs is debunked, moral beliefs could still be justifiable under the framework of her response-dependent (anti-realist) theory, namely constructivism. So, Joyce (2016) categorizes her debunking argument as evolutionary debunking of *ethical theory* (moral realism) since her debunking is directed against justification of realist moral beliefs but tolerate justification of constructivist moral beliefs. As for Ruse, he seems to take the

connecting premise as he claims that "ethics is subjective, but its meaning is objective". (2006, 22) Therefore, Joyce categorizes his argument, extended via the connecting premise, as truth-debunking which implies that there is no (any kind of) moral truth at all.

So far I have introduced the Darwinian Hypothesis deployed by the debunkers and have synthesized their arguments against moral realism. In the following chapters, I shall critically examine various realist responses to the epistemological type of debunking arguments. The main reason that I will firstly put aside Ruse's metaphysical argument is that its conclusion is too strong and is drawn hastily. The debunking account might be an adequate proof that we do not have reason to believe in the existence of independent moral truths. However, whether we have good reason to believe that these independent truths do not exist is another story. This is why agnosticism is different from atheism. And I think that Ruse, given his argument, can at best prove the moral agnostic conclusion that we have no reason to believe that those moral truths exist. Indeed, his premise R4: "if something's existence is metaphysically redundant in explaining things, then it does not exist or at least we have good reason to believe it does not" is highly controversial. Therefore, it seems that he is jumping to the conclusion that there are no independent moral truths and all our moral beliefs are

illusions. A minor reason for neglecting Ruse's argument is that most if not all prominent realist replies are directed against the epistemological debunking. For these two reasons, I shall focus on critics of Joyce and Street. I aim at showing that neither the non-naturalist replies nor the naturalist replies discussed below are convincing.

Chapter III: Non-naturalist Responses

3(a) William FitzPatrick

A key methodological point employed by the debunkers is the principle of parsimony. As I mentioned above, the principle of parsimony, or Ockham's Razor, is used in both epistemological debunking arguments. We can see the application of parsimony in premise R4 of Ruse's argument, in premise J1.1 of Joyce's debunking and in premise S2 of Street's argument. If the principle of parsimony is refuted, then the debunkers will have no grounds (or at least lose one important basis) for arguing that the non-truth-tracking genealogy is the best explanation of our moral beliefs. Among the non-naturalist vindicators, William J. FitzPatrick (2015) is one who refutes the use of parsimony in evolutionary debunking. He firstly distinguishes two forms of epistemological debunking arguments: the *capacity etiology argument* and the *content etiology argument*. He proposes two responses to these two types of arguments respectively. One response is that the content etiology argument fails because it unfairly relies on a methodological premise – parsimony – which begs the question against moral realism from the start. In the following, I shall argue that this response fails because the principle of parsimony does not beg the question against moral realism from the start. Before giving my criticism of FitzPatrick's response to the content etiology argument, I will firstly criticize the one he proposes to the capacity etiology

argument, according to which the capacity argument fails to raise genuine problem for realism. In my opinion, both responses are not convincing.

First, let me explain the distinction that FitzPartick draws between the capacity etiology argument and the content etiology argument. Recall the two different claims I have classified in section 1(c) which should be made in the Darwinian Hypothesis:

> **The capacity claim:** natural selection produced a general capacity or a set of capacities employed in making moral judgment.
>
> **The raw material claim:** natural selection directly shaped the content of basic moral thoughts as starting points/raw materials of moral reflection.

The capacity etiology argument, classified by FitzPatrick, is distinctively based upon the capacity claim that our basic mental capacities employed in making moral judgment are products of natural selection. Likewise, on his classification, the content etiology argument is distinctively based upon the raw material claim that natural selection has greatly shaped the very content of intuitive moral beliefs as the starting points of reflection. In section 1(c), I stated that these two claims are not mutually exclusive because natural selection could produce both the capacities and the raw materials. But notice that FitzPatrick treats these two claims and the two versions of debunking arguments more distinctively than I do. For the sake of argument, I set aside whether

his way of making such a distinction is legitimate and accept his classification.

The capacity etiology argument, as FitzPatrick classifies, says that basic cognitive capacities involved in moral-judgment-making, given their genealogy, are not reliable in tracking independent moral truths because natural selection would not have specifically designed them to be thus truth-tracking. FitzPatrick replies that although natural selection has not tailored the basic capacities employed in making moral judgments to the need of tracking independent moral truths, it does not mean that they cannot reliably track the truths at all. One example he uses to illustrate this point is the reliable beliefs about prime numbers. Clearly, using FitzPatrick's words, natural selection did not hand us a *ready-made* faculty for recognizing prime numbers correctly. However, the genealogical story, according to which we come to believe that 2, 11, 37, 163, 457, 769 are prime numbers because natural selection has shaped our mathematical, reasoning and concept sophisticating capacities which enable us to recognize them, would not be worrisome for realist of prime numbers. Facts about prime numbers do not seem to play any role in the adaptive explanation of the mathematical beliefs in question because it does not matter in terms of biological fitness whether our ancestor correctly recognized which are prime numbers and which are not. But as long as natural selection has given us reliable capacities to yield correct judgments about *numbers* and

to reason or to conceptually sophisticate, the sum of these basic capacities should also generate correct beliefs about prime numbers.

From the example of prime numbers, FitzPatrick (2015, 888) argues that moral reasoning can likewise viewed as "an extension of forms of reasoning our capacities are designed to do accurately". Apart from the same logical and analytic abilities which are employed in other forms of reasoning like the one for prime numbers, FitzPatrick (2015, 888) adds that moral reasoning "is continuous with broader evaluative and normative thinking that our cognitive capacities were plausibly designed to do accurately". I think that this is a crucial but also a fragile point in his response to the capacity etiology debunking argument. While taking the point that moral reasoning is continuous with broader evaluative and normative thinking such as "survival is *good*", "having strong hunters is *good*" and "poisonous food is *bad*", I reject the idea that our cognitive capacities were designed by natural selection to track the *normative* or *evaluative* truths in the realist sense. In the following, I will continue to compare moral beliefs with beliefs about prime numbers to show that there is a significant disanalogy between them even if we guarantee that both are extensions from basic cognitive capacities designed by natural selection.

The first problem which should be addressed here is what grounds we can find in the genealogy of prime number belief for saying that the belief in question and the basic capacities deployed in yielding the belief are reliably truth-tracking. The production of beliefs about prime numbers, as FitzPatrick claims, relies on the basic mathematical capacities and other reasoning abilities evolved in human's biological history. The evolutionary origin of these abilities provides a good reason for believing in their reliability of capturing independent truths because the reliability is necessarily asserted in the evolutionary explanation of the beliefs in question. For instance, when an ancestral tribe needed to decide whether to fight against enemies or run away, they had to know whether they were having more fighters than their enemies, meaning that they had to do simple mathematical reasoning. Presumably only those reliable mathematical capacities would become products of natural selection because the capacities would not have promoted selective advantages if they did not help our ancestors to reach independent truths about the actual numbers of enemies and in-group fighters and about which one is greater. In fact, it would be fatal to the tribe if they were relying on unreliable mathematical abilities in the sense of failing to track independent mathematical truths such as the fact that the tribe were having less fighters than their enemies. Regarding mathematical capacities (as well as other basic reasoning capacities), the connection between truth-trackingness and adaptiveness is obvious: if

the abilities are not reliable in telling independent truths about the world, they would not have become adaptations.

Then, how about moral beliefs? As FitzPatrick claims, moral thinking is derived from more general normative thinking about what is good and what is bad. To establish the analogy between moral belief and prime number belief, he insists that basic normative capacities are as reliable as basic mathematical capacities in the sense of yielding independent truths. I think a reasonable objection from the debunking camp is that given the success of genealogical explanations of moral beliefs and also *normative* beliefs, the reliability of the capacities producing these beliefs are thus undermined because grounds for their reliability in telling independent truths cannot be found in the evolutionary explanation. In contrast to the basic mathematical capacities, the independent-truth-trackingness of general normative capacities has no obvious connection with the adaptiveness of these capacities. Even if the capacities in question did not tell our ancestors independent normative truths such as "survival is good", they would still promote selective advantage as long as they disposed the ancestors to act in the way that eventually led them to have more babies than their competitors. In other words, the adaptive function of normative capacities does not seem to have anything to do with whether or not the capacities track independent normative truths posited by

normative realist. More importantly, going along with Street's approach, to serve the adaptive function, it seems sufficient for the normative capacity to merely track *response-dependent* normative truths which are anti-realist.

FitzPatrick does notice the problem and tries to build a bridge between the adaptive function and independent-truth-tracking. He writes

> Indeed, debunkers such as Street, who despite rejecting realism wish to save moral knowledge by moving to a subjectivist constructivist account of moral truth, presumably agree that such extensions are possible and afford us moral knowledge. So their thought must just be that this is problematic if the moral facts are taken to be objective or independent of our contingent desires or attitudes, as realists maintain. But nothing about evolution considered so far supports any special worry along those lines. The independence or objectivity of the goodness or badness of potential partners as hunters, or of certain mushrooms as food, was no hindrance to accurate evaluative judgment about them. Similarly, there is no reason so far why the independence or objectivity of the badness of racist voting laws should be thought to pose a special obstacle to our discovering that badness, at least as far as evolutionary considerations are concerned. (2015, 889)

The reason why I think the response above is wanting is that in establishing the connection between truth-trackingness and adaptiveness, it is of no help to merely

emphasize that the independence of normative truths itself does not pose special obstacle to the reliability of the capacities to discover them. Indeed, the debunking issue faced by the realist is that the best explanation of our normative beliefs does not contain any assertion about the independence of normative truth. Accordingly, the evolutionary mechanism has generated and shaped normative beliefs without appealing to independent normative truths. Here the realist still owes us explanations of why we should add independent normative truths into the allegedly completed genealogical picture and how independent normative truths serve the adaptive function.

So far, I have discussed FitzPatrick's response to the capacity version of debunking argument. This response attempts to invoke an analogy between moral beliefs and other reliable beliefs by appealing to the claim that both kinds of belief are based upon some basic evolved capacities. I argue that such attempt fails because the connection between truth-trackingness and adaptiveness is much more obvious in the genealogies of other beliefs, for instance mathematical beliefs, than in the genealogy of moral beliefs. In other words, the Darwinian explanation has given us good reason to think that beliefs about prime numbers are reliable in tracking realist facts, but has not given us good reason to think the same regarding moral beliefs and normative beliefs.

Another response proposed by FitzPatrick is directed against the content version of debunking argument. Alleging the capacity argument is refuted, FitzPatrick considers the variant that "the real problem stems from what natural selection *has* done in the way of shaping the very content of our moral concepts and beliefs." (2015, 890) So, now the evolutionary premise shifts from $P_{capacity}$: the basic capacities we employ in moral judgment are products of natural selection to $P_{content}$: natural selection has greatly shaped the content of our moral thoughts. FitzPatrick thinks that the shift allows the debunker to create a genuine problem for the realist. If the best explanation of the content of moral beliefs is the truth-irrelevant genealogy, then the beliefs in question really seem unreliable. Why the truth-irrelevant genealogy is the best explanation of moral beliefs? One important reason for thinking so lies in parsimony: the truth-irrelevant genealogy is simpler than what the realist can offer. If independent moral truth plays no role in the best explanation of moral beliefs – the genealogical explanation – then there would be no grounds for asserting the truth-trackingness of these evolved moral beliefs. And if there is no truth-tracking, then it would be a massive coincidence that our moral beliefs match the independent moral truths.

The principle of parsimony is a crucial point shared by all three prominent debunking arguments. Ruse uses it to debunk the existence of moral truths. Joyce uses it to

preclude the non-naturalist and the supernaturalist account of justification. Street uses it to debunk realist accounts which assert truth-trackingness of evolved moral beliefs. Here comes FitzPatrick's methodological criticism of the use of the principle of parsimony in debunking:

> Greater parsimony is a theoretical virtue only where the world is obligingly austere, and that is exactly what is at issue in this debate. The realist's position (at least for the realist I am defending) is precisely that the world does contain objective values that figure crucially into our having at least some of the moral beliefs we hold, in which case greater parsimony in explaining our beliefs exclusively in other terms is not a virtue at all: it is just a misrepresentation. (2015, 893)

Here FitzPatrick does not seem to propose a general complaint about validity of parsimony in the way that some skeptics of this IBE criterion do. Instead, his point is that the deployment of parsimony in the debunking context is illegitimate. If we accept the principle in the first place, then the realist's central claim that there are independent moral truths or objective values would be automatically undermined. If moral realism is true, then parsimony is not a theoretical virtue. Accordingly, if we deploy parsimony as a virtue, then we just beg the question against realism from the start, which is unfair.

However, FitzPatrick's concept of parsimony is relatively rough. I shall argue that the

principle of parsimony does not beg the question against realism from the start by clarifying what the principle really means and how it should be used. A simple and intuitive description of parsimony is: the simpler or the less complicated a theory or a candidate explanation is the better. But this description is misleading because parsimony is not just about the quantity of theoretical entities asserted by theories. Consider the naïve science taught in primary schools or even kindergartens. These naïve versions of scientific theories are clearly much simpler than the real ones conducted by expert scientists. But we would not simply apply parsimony and say that the naïve ones are better theories than the expert ones. A better way to describe parsimony should be: do not assert more than what is actually needed in explaining things. Of course, one can say that this is just because of another criterion of inference to the best explanation (IBE), namely explanatory power. But I shall argue that the principle of parsimony is never conceived of independently. Considerations of explanatory power are always premised when considering whether to use Ockham's Razor to shave off certain accounts. Otherwise, it would be meaningless or highly controversial to deploy such principle.

Given the assertion that parsimony also takes care of how much is *needed* to explain phenomena, there is no beforehand conflict between moral realism and parsimony

unless realism is the thesis that there are objective moral values which are *not needed* in explaining things (presumably moral phenomena). The central ontological claim of realism does not entail that the existence of those independent values does not possess any significant explanatory power and thus is explanatorily superfluous. If the concept of parsimony is like the one I have described above: do not posit more than what is needed in the adequate explanation, then moral realist explanations would be knocked out by the use of parsimony only when the success of the truth-irrelevant genealogy implies that independent moral truths are not needed in the adequate explanation. If the use of parsimony does not beg the question against realism from the start, then FitzPatrick has offered us no reason to reject it. Notice that here I am not trying to deal with deep methodological skeptics about parsimony. I maintain that the validity of parsimony is supported by intuitive methodological thinking, which is debatable at deep levels of philosophical methodology. Perhaps FitzPatrick is trying to argue, like the deep skeptics about parsimony, that parsimony is not a theoretical virtue in any case. Then, my final word to say here is that if parsimony is forsaken in this way, then we would be left with little to preclude redundant theories which are rationally unsatisfying. This would lead to an awful result that any theorist would be allowed to defend any explanatorily superfluous account by simply complaining that the use of parsimony begs the question against her account from the start.

3(b) Erik J. Wielenberg

In last section, we have seen how FitzPartick tries to save truth-trackingness of evolved moral beliefs on the basis of the capacity reading of evolutionary debunking and his attempt to refute the use of parsimony in debunking arguments. Now we may turn to another realist reply, which is proposed by Erik J. Wielenberg (2010). Recall that Joyce's debunking argument has the following premise:

J4 If moral beliefs are explained without appealing to moral truths, then they are products of a process which is dependent of or insensitive to their truth.

Wielenberg argues that this premise is false. His project has both a positive line and a negative line. The positive line is to sketch a reliabilist account of moral knowledge in the evolutionary context, whereas the negative line is to argue that the analogy employed by Joyce to support premise J4 fails. I think that both ways of replying are mistaken. First, Wielenberg's reliabilist approach commits the same problem as FitzPatrick's response to the capacity etiology argument. Second, his accusation that Joyce has not provided substantial support for premise J4 is based on a misunderstanding of the debunking strategy. As the latter response is less complex than the former, I shall firstly introduce and respond to Wielenberg's criticism of Joyce's analogy and then explain his reliabilist approach.

Wielenberg points out that what Joyce has offered to support premise J4 is an analogy. The analogy is that moral beliefs are like a kind of imaginary belief about the plant's magic power which encourages the belief-holder to eat the plant in particular circumstances, which is selectively beneficial for medicinal reasons. Now both moral beliefs and the plant belief are products of evolved psychological mechanism which does not require their truth. Presumably the plant belief is product of an unreliable belief-forming mechanism. Analogically, the conclusion is that moral beliefs are also product of unreliable mechanisms and thus are unjustified.[11] Wielenberg argues that the analogy fails because "the weakness of this reasoning lies in the fact that the plant-belief has a relevant feature that we cannot simply assume human moral beliefs possess, namely, falsehood". (2010, 462) Here I agree that the analogy might be misleading or, more precisely, sneakily intuition-steering. However, I think what Joyce tries to demonstrate with the analogy is a counterfactual condition that the evolutionary process, as a reliable process producing true belief, must meet. If an evolutionary mechanism is reliable, then

A. If the belief were true, it will contribute to selective fitness.

B. If the belief were not true, it will not contribute to selective fitness.

[11] For the analogy, see Joyce (2006, 215)

According to my understanding, Joyce wants to point out that some evolutionary mechanisms just do not fit in the second conditional: they produce beliefs that are nonetheless selectively beneficial even though they are false. And we have good reasons to think that evolutionary mechanisms which produce moral beliefs cannot meet the counterfactual condition because it is unobvious that the falsehood of a moral belief will undermine its contribution to fitness.

Taking a step back, even if it should be admitted that the plant-belief analogy is unfair, Joyce has other analogies that simply do not presuppose the falsehood of beliefs, but these analogies are still capable of supporting premise J4. Recall Joyce's Waterloo analogy: a belief-pill makes people believe that Napoleon won the Waterloo regardless of whether it is the fact. This analogy does not presuppose that Napoleon did not win the Waterloo (in fact, he did win the Waterloo). But the scenario is that if we are the ones who believe that Napoleon won the Waterloo after taking the truth-irrelevant pill, what should we do after we know that our belief is the product of this pill? The natural answer is that we should case doubts upon the justificatory status of our Napoleon belief and upon the reliability of this belief forming process. Therefore, I shall conclude that Wielenberg's negative task of undermining Joyce's support for premise J4 fails for two reasons. First, the alleged falsehood of the plant belief does not create genuine disparity

as long as the realist has not provided any good reason that evolved moral belief can meet the counterfactual condition: if the belief were not true, it will not contribute to selective fitness. Second, even if the magic plant analogy fails for unfairly generalizing the falsehood of the plant belief to moral beliefs, the debunker still has other utilizable analogies which does not presuppose the falsehood of beliefs.

Before illustrating my critique of Wielenberg's positive task of providing a reliabilist account to show that premise J4 is false, I should acknowledge that there are several general points made by him that I shall not doubt. One is that evolutionary processes are generally reliable belief-forming processes. Here Wielenberg quotes two reasons for thinking so from Peter Carruthers' argument (1992): (i) reliable belief-forming is selectively beneficial and (ii) many proximate belief-forming mechanisms to which natural selection has given rise are reliable. Apparent examples which can lay out these two points include perception, mathematics and memory. In addition, Wielenberg makes the reliabilist assumption that beliefs produced by reliable processes are sufficiently justified to be knowledge. I accept all these points not just for the sake of argument but also because I am truly inclined to think that they are correct. Another point asserted by Wielenberg that I would like to take for granted is that some basic moral thoughts are produced by certain evolved cognitive faculties. The description of

the cognitive faculties given by Wielenberg is relatively rough: "the cognitive faculties in question are either the very ones required to form beliefs about rights or are closely linked to such faculties." (2010, 449) But I take what he has in mind to be similar to the basic mental capacities highlighted by FitzPatrick (see last section).

The reliabilist approach proposed by Wielenberg is very similar to FitzPatrick's response to the capacity etiology argument. Nevertheless, the problems from which their responses suffer are thus very similar. The first step taken by both of them is to ensure the general reliability of cognitive faculties produced by natural selection. The second step is to argue that this set of reliable cognitive capacities gives rise to a bunch of basic moral or normative thoughts which grounds the whole morality. For FitzPatrick, the basic thoughts generated by this mechanism are general normative beliefs about what is good and bad. For Wielenberg, the basic thoughts generated are moral beliefs about what he calls "moral barriers". But, after all, what are these barriers? Wielenberg maintains that moral barriers are

> Conceptualized differently in different cultures. In the West, the barrier is presently conceptualized in terms of rights. Under Islam it is conceptualized in terms of duties. In Neo-Confucianism, it is conceptualized in terms of legitimate desires. Despite various cultural differences, human beings normally believe that there are certain things that others simply

> ought not do to them, for example, rape them, enslave them, steal from them, or kill them for entertainment.

Now it might be seen that there would be no surprise that being disposed to such moral beliefs is fitness-enhancing. More importantly, he argues that as being produced by reliable mechanism, these moral beliefs are at least roughly correct (in the realist sense).

After all, how can this reliabilist model prove that J4 is false? Wielenberg asks us to consider

> a Westerner's belief that she possesses certain moral rights. The evolutionary explanation for this belief (Wielenberg sketched above) does not appeal to the truth of this belief. Yet the model says that moral rights supervene on the cognitive faculties that generate this belief, and this thus implies that, even though moral rights in no way help to explain the corresponding belief about moral rights, the belief in question is nevertheless true.

Again, I shall argue against reliability of the Westerner's belief in question, just like I have argued against reliability of basic beliefs posited by FitzPatrick. A crucial point I will deploy here is that reliability of a certain should not be over-generalized. When we consider reliability of a certain mechanism or process, we never its reliability in generating *all kinds of* beliefs. A question to be always asked is "reliable with respect to what?" When considering the reliability of the mechanism producing mathematical

beliefs, it would be inappropriate to ask whether the mechanism in question is reliable in yielding true, let's say, perceptual beliefs.

Now back to Wielenberg's model, for those cognitive faculties like logical reasoning, perception and mathematical operating, it would be bizarre if we extend the reliability of these faculties to the reliability of their making moral beliefs. Of course, natural selection has given us reliable capacities to logically reason, perceive and calculate so that these capacities always lead us to independent truths of logic, perception and mathematics, which is crucial in fitness enhancement. But it does not mean that they are also reliable in leading us to independent moral or normative truths. Presumably the set of cognitive capacities denoted by Wielenberg also including a certain amount of evaluative ones. Parallel to FitzPatrick's normative foundations, the genealogical picture has provided no ground for their independent-truth-trackingness. For it is still hard to see why promotion of fitness requires independence of truths indicated by evaluative beliefs.

3(c) David Enoch

Here I shall introduce another similar kind of response proposed by David Enoch. In this section I will not repeat the same criticisms directed towards FitzPatrick and

Wielenberg. Here I aim at illustrating what these three realists have in common and what systematic mistakes have been made by their common form of replying.

First of all, the epistemological debunking arguments (both Joyce and Street) suggest that the correspondence between realist moral truths and evolved moral beliefs would be fairly surprising given the fact that the best explanation of our moral beliefs is non-truth-tracking genealogy. Accordingly, the task in front of the realist is to explain the correspondence in the evolutionary context so that it would not be so surprising that the content of our moral beliefs matches up with independent moral truths. In order to fulfill the task, Enoch (2010) proposes his *third-factor explanation* to account for the correspondence in question. So what are the first, the second and the third factors respectively? The first is the factor A which lies in the side of independent moral truths, the second is the factor B which lies in the side of genealogical moral beliefs. Correspondingly, the independent moral facts are A-facts, and genealogical facts about moral beliefs are B-facts. To bridge the two sides, Enoch employs a third factor C, which lies in where A and B overlap, to explain why A-facts and B-facts can match up unsurprisingly. The strategy here is to propose that factor C is correlated with independent moral truths (factor A) and evolved moral beliefs (factor B).

The third-factor strategy is not the monopoly of Enoch. As we have seen above, FitzPatrick (see his response to the capacity etiology version of debunking on p.35) and Wielenberg also deploy the same strategy to defend moral realism against debunking. The main difference between these three theorists lies in which third factor they identify. Wielenberg's and FitzPatrick's choice is a set of cognitive capacities produced by natural selection. Enoch seems to focus on the goodness of survival or reproductive success. So how do these third factors explain the correspondence between evolved moral beliefs and independent moral truths? Use Enoch as an example. On one hand, it would be no surprise that natural selection has given rise to the belief that survival is good. On the other hand, it seems plausible to say that the claim that survival is true. The goodness of survival becomes the third factor which is responsible for both the evolved belief and the moral truth. As for Wielenberg and FitzPatrick, a certain set of cognitive faculties are both responsible for the content of fundamental moral truths and the evolved beliefs in these truths. Nonetheless, this way of reply does not really turn surprising coincidence between independent moral truths and evolved moral beliefs into reasonable correspondence.

The systematic mistake buried in the third-factor strategy is that it does not do anything significant. The third-factor explanation merely reasserts the correspondence between

A-facts (independent moral truths) and B-facts (evolved moral belief). It might be able to reasonably assert the connection between C-factor (alleged goodness of survival or cognitive faculties) and A-factor (moral beliefs). But it would still be question-begging to assert the connection between C-factor (alleged goodness of survival or cognitive faculties) and B-factor (independent moral truths). For Enoch, he has not given non-question-begging explanation of why the goodness of survival is or relates to independent truths. For Wielenberg and FitzPatrick, they also have not provided good reason for thinking that the cognitive capacities in question are connected to independent moral truths under the pressure of natural selection, given that the best explanation is the non-truth-tracking genealogy. To conclude, the third-factor response is a positive type of response in the sense that the proponents wish to establish something to account for the unsurprising correspondence between evolved moral beliefs and independent moral truths. I think that their projects are in general question-begging in the sense of merely reasserting the correspondence between evolved moral beliefs and independent truths. In the following, I shall examine two negative responses proposed by realists. One argues that the debunkers are replying on unsupportive interpretations of massive coincidence. The other suggests that epistemological debunking is self-underming.

3(d) Russ Shafer-Landau

Massive coincidence is an important concept explicit in Street's Darwinian Dilemma. Recall that Street's argument, in brief, is that if the best explanation of our moral thoughts or moral faculties is that they are generated by the off-track (not being hooked up with the existence and content of realist moral truths) evolutionary mechanism, then it would be a massive coincidence that our moral thoughts or moral faculties capture realist moral truths. And if it is a sheer coincidence that our moral thoughts or moral faculties capture realist moral truths, then there is a good reason to believe that they are not reliable and thus cannot give us moral knowledge. But what does "massive coincidence" literally mean? This is the question raised and answered by Shafer-Landau (2012) in "Evolutionary debunking, moral realism and moral knowledge". In the article, he offers five possible interpretations of massive coincidence and examine them one by one. He concludes that none of the five interpretations warrants the premise of Street's debunking argument.

On Shafer-Landau's reconstruction, the key premise which claims massive coincidence is the following:

D1 If moral realism is true, and if evolutionary forces have thoroughly shaped our moral faculties in doxastically discriminating ways, then it would be a massive coincidence were our moral faculties reliable.

I take premise D1 to be identical with premise S4 in my reconstruction of Street's argument:

S4 If tracking is false, then the connection between moral beliefs and realist moral truths cannot be more than sheer coincidence.

The five interpretations he offers are respectively (i) The odds, (ii) Insensitivity, (iii) The confirmed doxastic effects of selective pressures, (iv) The empiricist argument and (v) The causal argument. Before examining his critique of premise S4 (or D1) based on the first three interpretations, I shall firstly point out that these five interpretations should not be treated separately, as done by Shafer-Landau. All five meanings of "massive coincidence" matters in supporting premise S4 in a collective manner. The most powerful version of S4 should be interpreted as conjointly supported by all five interpretations. In the following, I shall argue that at least three interpretations of massive coincidence – (i) The odds, (ii) Insensitivity and (iii) The confirmed doxastic effects of selective pressures – support premise S4.

First, the odds interpretation:

> **The odds:** There is an indefinitely large set of possible moral commitments, and evolutionary pressures have led us to adopt just a small number of these. If we suspend judgment about which outlook is actually correct – as we must do, according to debunkers – then the chance that selective pressures have led us to the realistically correct set of moral commitments is extremely small. Thus it would be an extraordinary coincidence were our core moral beliefs on target, given their evolutionary origins. (Shafer-Landau 2012, 10)

Here is an analogy to illustrate the odds interpretation. There is a bag of cartoons stickers the number of which is indefinitely large. Some among these cartoon stickers are your son's favorite, others are not. Now someone, who has no connection to the fact about your son's favorite, has drawn out or picked out a small number of stickers from the bag. Leaving which are your son's favorite unknown, how likely are the ones drawn or picked our actually your son's favorite? The natural answer would be "quite unlikely".

Shafer-Landau proposes three solutions to the challenge thus interpreted. The first solution appeals to certain conceptual constraints on what can qualify as a moral view. The second solution argues that the challenge, if successful, will prove too much, that is, proving that for instance our perceptual faculties are also not reliable and thus debunked. The third solution, which deals with constructivism particularly, argues that

the constructivist does not earn special advantages over the realist on this issue of odds. I shall now examine these solutions one by one and conclude that the first is question-begging, the second relies on an inadequate analogy between the debunking effect on moral beliefs or faculties and perceptual beliefs or faculties, and the third is derived from misunderstanding of the constructivist's advantage over the realist.

Why is the odds so low for the realist? Shafer-Landau says that

> If every possible moral proposition is on the table as a candidate for stance-independent truth, and we, as a result of selective pressures, believe just a small fraction from this array of possibilities, then the odds really are stacked against us, and … the chances of our having been influenced in reliable ways are quite slim. (2012, 11)

Accordingly, his aim here is to narrow down the group of possible candidates for independent moral truths. How? Shafer-Landau suggests that the realist can argue for "a moral semantics according to which certain moral propositions are reference fixing." (2012, 11) The realist could reject the idea of Street (2008, 208) that the independent truths in question can be any set of moral propositions like survival is bad, children's lives are worthy, genocide is not wrong, rape is morally fine and so on. By appealing to semantics concepts of moral language, the realist can claim that the speaker of these "moral" propositions is not talking about morality at all. In other words, the realist can

appeal to some basic conceptual truths that grounds the talk of morality. As a result, independent moral truths cannot be everything. They must either be some basic, paradigmatic moral propositions or moral propositions which are grounded by these preconditional conceptual truths.

I think the reply in question is question-begging in two ways. First, even if we agree that there are some basic moral claims that are reference-fixing and the moral talk must be either literally about them or complying with them, it is unclear why the affirmation of such moral truths can increase the odds that our evolved moral beliefs or faculties are reliable in capturing these truths given that the evolutionary forces operated without relying on these truths. Back to the analogy of drawing stickers, we can suppose, for the sake of argument, certain propositions about your son's favorite stickers are conceptual truths (let's pretend that there are such conceptual truths). Now presumably the number of the candidates for your son's favorite will be narrowed down. For instance, the stickers must either be of Iron Man or of Bat Man otherwise we are talking in a correct semantic manner in which we use the concept "the son", "favorite", "sticker" and so on. But given that the person has drawn out the stickers from the big bag without relying on the fact whether you son likes them or not, the intuition would still be that it is unlikely what have been drawn from the big bag are actually your son's favorite.

Second, what grounds do the realist have for saying that certain moral claims are reference-fixing in the evolutionary context? Of course, they can simply respond that these are conceptual truths. But the debunker can reply that the reason why we have these concept is exactly that non-truth-tracking evolutionary forces shaped our concept in this way. So, while relying on these conceptual truths, the realist, or at least Shafer-Landau, does not seem to well notice that they are also evolutionarily "polluted". In short, the mistake in the approach seems to lie in the difference between a truth's being conceptual and its being independent. We can well accept that basic moral claims are conceptual truths without accepting that the best genealogical explanation of these claims implies or presuppose their being realistically independent.

Taking a step back, Shafer-Landau argues that even if we grant that the odds argument is successful, using the same strategy would prove too much. When we replace "moral faculties and beliefs" with "perceptual faculties and beliefs", we might have a parallel argument against the reliability of perception. But the conclusion that we do not have perceptual knowledge is clearly (expect for some extreme skeptics) unconvincing. I think the response fails because there is a significant disanalogy between the two entities in question. For it is pretty clear how the genealogical explanation, alleged

successful, provides support for the reliability of perception in tracking *independent truths*, truths about the real situation of the world. In contrast, as said above, the non-naturalist has not provided us positive account of how the independent moral truths are linked up with the adaptiveness of moral faculties and beliefs.

I have defended that the odds interpretation can support premise S4, if my defense is successful, then we have at least one interpretation of massive coincidence which can ground Street's epistemological debunking. The second interpretation I shall discuss now is the insensitivity interpretation:

> **Insensitivity**: We would hold the mora beliefs we have because of evolutionary pressure, even if they were false. (2012, 15)

This interpretation of massive coincidence appears more explicitly in Joyce's debunking argument. Indeed, while discussing the insensitivity interpretation, Shafer-Landau focuses on Joyce's critique. However, I think his understanding of the Waterloo critique (the one illustrated by the belief pill thought experiment) is deeply problematic. He claims that Joyce's critique is sound only if two conditions are true:

> (1) To the extent that our moral beliefs are the product of evolutionary forces, such beliefs are formed in ways that are insensitive to whatever moral truth there may be.

(2) Our inability to exclude the possibility of insensitive doxastic origins for a set of beliefs S mandates suspension of judgment regarding all beliefs within that set.

(Shafer-Landau 2012, 17-18)

Then Shafer-Landau set aside condition (1) and criticize (2). His first criticism is that if (2) is true than we would face wholesale skepticism because we are also unable to exclude the possibility of insensitive doxastic origins for many beliefs. Every belief, we can say, has a chance of being produced by evil demon or brain-in-the-vat. So, the success of condition (2) and Joyce's critique leads to a global skepticism which seems deeply problematic. Nonetheless, this reply to Joyce and also other debunking arguments commits the straw man fallacy. Joyce, as well as other debunkers, does not rely on the merely theoretical possibility that moral beliefs *might* be produced by an evolutionary mechanism which is doxastically insensitive. All of them distinguishes the evolutionary debunking argument from global skeptical challenge like the Cartesian Evil Demon and Brain in The Vat. All the debunkers start from the claim that the natural causal genealogy of moral beliefs or faculties is a successful theory in the sense of being empirically confirmed.[12] In the analogy of Waterloo, it is also empirically confirmed

[12] Regarding this point, Joyce states that "the view under discussion here does not come so cheap. It is not just that in this case we can make up a consistent hypothesis according to which a bunch of our ordinary beliefs are false; rather, it is that we might have empirical evidence supporting the hypothesis that explains how these beliefs came about but does not require that they be true. The argument does not depend on invoking extreme standards for epistemic justification; the skeptic is not requiring people to consider outlandish brain-in-vat-type possibilities that they would ordinarily scoff at. If the everyday standards for being morally justified take account of empirical data concerning human evolution, then if these data ultimately show moral beliefs to be unjustified it will be by ordinary epistemic standards." (2006, 188)

that the historical belief is generated by the truth-irrelevant belief pill. Therefore, condition (2) is not an accurate interpretation of what supports the debunking arguments.

If we erase condition (2), does condition (1) alone seem adequate for spelling out the debunking conclusion? It might intuitively work in this way but nonetheless seem not satisfying enough. So, what else do we have to replace condition (2)? Shafer-Landau claims that there are two which can replace condition (2): full-scale effect and distortion. Both of these two points are crucial details of the Darwinian hypothesis described in section 1(b) and 1(c). Condition (1) does not imply that *all* moral beliefs are the products of insensitive evolutionary mechanisms. So, for the debunker to debunk the whole set of moral beliefs without leaving some parts intact, they have to assert the full-scale effect of evolutionary forces. However, as mentioned in section 1(b) and 1(c), not all moral beliefs are straightforward products of evolutionary history. So, Shafer-Landau claims that "evolutionary influences are at best only indirectly efficacious in causing us to have the specific moral beliefs we do. For what those they do affect, what influence they wield is not a matter of wholesale causation." But he also notices that it would be a lame realist reply if it relies on the claim that the evolutionary influence is not a wholesale causation.[13] For if the extent to which evolutionary forces have shaped

[13] Here "wholescale causation" means that every moral belief is a direct product of natural selection in virtue of adaptiveness.

the content of our moral beliefs is high enough, it does matter even if a considerable portion of the influences is indirectly efficacious. Now for bring out the result, the debunker claims that the evolutionary influences, direct or indirect, are distorting. Shafer-Landau's reply attempts to find a gap for the realist. Here are two claims made by him that I think is refutable:

> **The possibility claim**: Yet so long as such forces are not the sole cause of our presumptively warranted moral beliefs, then even if evolutionary pressures are distorting ones, there is the possibility that other doxastic operations can compensate for the distortions and lead us to the moral truth in ways that are perfectly familiar from other instances in which we have corrected for distorting doxastic influences.
>
> **The example claim**: we do recognize cases in which agents have been able to correct for very substantial distorting pressures on their moral faculties.
>
> (Shafer-Landau 2012, 19)

I think that the possibility claim is of no help to solve the problem, whereas the example claim is utterly question-begging. First, of course it is *possible* that there would be a compensation for the distortions. But given that the evolutionary influences, though not absolute, are comprehensive to a certain degree, how high is the probability? I think it is more intuitive to say that the probability of compensation is lower than the case that the influences in question is irresistible. If the realist wants to reject this claim, the

burden is on them to prove that the probability of compensation is actually quite high. In short, asserting the mere possibility of overcoming distorting evolutionary effects does not help to respond to the debunking challenge. In addition, the example claim is begging the question "how do we know those cases are genuine *corrections* in the realist sense?" and the question "does viewing the cases as successful corrections utterly refrain from evolutionary effect?" So far, I think it is safe to conclude Shafer-Landau's reply to the insensitivity argument fails.

3(e) Katia Vavova

In the previous sections, we have seen realists who have made huge effort to argue against the use of parsimony, to account for the correspondence between evolved moral beliefs and independent moral truths and to refute the concept of massive coincidence deployed in epistemological debunking. Katia Vavova's (2014) defense of mora realism is relatively simple. She attempts to show that evolutionary arguments, if successful, would be self-undermining. In her reconstruction of EDAs, a key starting premise of epistemological debunking argument is that

P2 The true evaluative beliefs and the adaptive evaluative beliefs come apart.

In order to support P2, she argues, we have to know both the content of our evolved evaluative beliefs and the content of independent evaluative beliefs. But EDAs, if successful, would leave us no room for knowing the content of latter. Thus, the success of EDAs implies the lack of support for P2, which ultimately causes a self-undermining result. My criticism of Vavova is that her reconstruction of the EDAs is not accurate because P2 is not really a starting premise in the EDAs and debunkers do not have to assume the very content of such-and-such evaluative truths in order to support P2. I will give an analogy demonstrating this point.

First of all, I think Vavova has a convincing claim which summarizes one main mistakes made by the non-naturalist reply above. She writes that

> The problem here is that our entire body of moral beliefs is suspect. It follows that we must set all of our moral beliefs aside, if we are to block such question-begging responses. We cannot, then simply assume that we have reason to avoid pain – that morality is about what is good for us, and that needlessly throwing ourselves off of cliffs just isn't that sort of thing. These assumptions aren't appropriately independent. (2014, 91-92)

I think this is a fair summary on why many of the above responses we have gone through are question-begging. However, she does not stop here and argues that "taking these assumptions off the table threatens to undermine the debunker's argument." Of

course, the debunker can always reply to the above non-naturalist responses by accuse of being question-begging when they make assumptions about certain moral truths, the basic ones which sound good enough to be picked out by evolutionary forces and supporting our moral reflection. But Vavova wants to remind us that the accusation is a double-edged sword for the debunkers.

According to Vavova's reconstruction, a key *starting* premise of the debunking arguments is that the true evaluative beliefs and the adaptive evaluative ones actually come apart. To be fair, debunkers cannot assume such-and-such moral truths, just like the realist cannot assume the content of moral truths in a question-begging manner. However, Vavova argues that avoiding doing so would lead to the lack of warrant for premise P2 in question. So, while debunking the beliefs in independent moral truths, the debunker would also eliminate the resource for supporting her own argument and thus become self-undermining. As clearly shown in the above critiques I have proposed, I agree with Vavova that in the debunking debate, anyone who does not want to be question-begging should not take any moral propositions for granted in the sense of regarding them as realist truths. But I do not agree that the debunker also suffers from this constraint because the premise categorized by Vavova is not a starting premise deployed by the debunkers.

According to my reconstructions of the epistemological debunking argument (also the metaphysical one), none of them seems to adopt the premise that independent moral truths and adaptive evaluative beliefs come apart as a starting premise. Notice that what the epistemological debunker tries to infer is that it would be a massive coincidence that independent moral truths and adaptive evaluative beliefs do not come apart (and therefore our beliefs are not knowledge). So, arguing that they do actually come apart from the start would be very problematic. Indeed, if the debunker does use the above premise P2 as a starting premise, it would make the debunking argument very vulnerable because of committing to circular reasoning or being question-begging from the start. Actually, the debunkers' texts show that they remain neutral at the begging on whether realist moral truths and the adaptive evaluative beliefs come apart, just like remain neutral about the content of independent truths. In the debunking arguments, the claim that they probably com part is in fact a premise inferred from the starting premises but does not serve as a starting one.

How do the debunkers support the premise that it would be a massive coincidence that realist moral truths and the evolved ones do not come apart? They do so not by presupposing such-and-such content of moral truths. Their strategy is to argue that the

non-truth-tracking mechanism, namely natural selection, is the best explanation of our thinking and this shows that realist moral truths and the evolved ones probably come apart or we have good reason to doubt that they will not come apart. For example, recall that Joyce's belief pill analogy does not say from the beginning that the belief in Waterloo and the historical truth come apart. Street's Darwinian Dilemma also does not make assumptions about the content of moral truths. Therefore, I shall conclude that Vavova's criticism relies on a mistaken reconstruction of the debunking arguments.

3(f) Summary on the non-naturalist approach

I shall conclude that in order to solve the epistemological debunking challenge, there are a few tasks that the realist needs to fulfill. The main task on table is to explain why moral thoughts produced by natural selection is truth-tracking in a non-question-begging manner so that it would not be a surprise that our moral beliefs are true. Given the principle of parsimony, the realist cannot propose any theory which is theoretically and ontologically more complicated than the explanatorily adequate genealogy which does not assert independent moral truths. Accordingly, the realist seems to be required to explain how the truth-trackingness of moral thoughts contributes to the selective advantage of moral thoughts within the natural, causal and genealogical map of morality. I shall conclude that the non-naturalist replies discussed in the above sections fail to

fulfill these tasks. Nonetheless, the naturalist camp of moral realism seems to be more promising in fulfilling the tasks. In the following, I shall examine the naturalist responses. I will firstly discuss the prominent naturalist reply proposed by David Copp and Street's response to it. Then I will move onto a general criticism of the naturalist reply in a general form.

Chapter IV: Naturalist Responses

4(a) David Copp

David Copp is the prominent naturalist vindicator who defends realism against the debunking challendge. He names his naturalist account "the society-centered theory" according to which morality serves "the function of enabling a society to meet its needs" (2008, 198). The corresponding truth condition of moral proposition can be spelled out by two parts of his theory: 1) the standard-based account and 2) the society-centered account of the truth-grounding status of moral standards. According to the standard-based account, a pure and basic moral proposition is true if and only if a certain morally authoritative standard grounds its truth. For example, if torture is wrong, it is prohibited by a morally authoritative standard. Furthermore, according to the society-centered account of the truth-grounding status of moral standards, a morally authoritative standard is a standard which is included in the system of standards which best serve the basic needs of a society. For example, it is true that killing people for fun is wrong because the system of standards which best serve the societal needs forbids such behavior. As for examples of the basic needs of a society, Copp roughly says that "a society needs to ensure that its population continues to exist. It needs to ensure that there is and continues to be a system of cooperation among its members. It needs to ensure internal social harmony. It needs peaceful and cooperative relationships with

neighboring societies." (2008, 200) Presumably, the moral standards which best serve basic societal needs would ground many moral propositions that we believe.

Such naturalist account, Copp argues, would allow moral realists to escape the Darwinian challenge. He argues that the evolutionary mechanism would track the moral truths posited by his social-centered theory to an epistemically sufficient degree because it would be no surprise that natural selection has picked out moral standards which best serve the society's basic needs. Accordingly, Copp's strategy is to firstly sketch a metaphysical or ontological framework, the naturalist one, in order to pair independent moral truths with evolved mora beliefs. So, Copp's goal, just like other realists', is to wipe off the charge of massive coincidence between two entities. The premise tackled by Copp's reply is Street's S2:

> **S2** Any account that realist uses to establish a connection more than sheer coincidence, namely tracking, would be less parsimonious, less clear and less explanatorily powerful than the Darwinian hypothesis.

If Copp's explanation, namely the social-centered theory, is successful, then we will have at least one account which can support a connection more than sheer coincidence – tracking.

Copp's theory and naturalism in general does have some seeming advantages over the non-naturalist approaches discussed above. In the debunking debate, non-naturalist accounts are often complained of not being as simple, clear and explanatorily powerful as the purely natural genealogy. Nonetheless, the naturalist thesis seems to help the naturalist accounts, including Copp's account, to overcome these difficulties. First, since it is a naturalist theory, it does not introduce entities and properties other than the natural ones. As Joyce says, "if the moral facts are reducible to the non-moral facts invoked in the genealogical explanation, then the former cannot be eliminated on grounds of parsimony, any more than cats should be eliminated from our ontology because we can explain them in terms of physics." (2006, 189) The talk of morality and moral knowledge could be vindicated if they can be comfortably integrated within the naturalistic picture. Second, as the naturalist claims that moral properties are identical to or constituted by natural properties, the causal efficacy of naturalistic morality is clearer than the non-naturalistic morality. If so, it would be easier for the naturalist to spell out the causal picture of how moral facts are connected with biological adaptiveness. Third, the naturalist, thus categorized, seems to have the power of explaining why it would not be a massive coincidence that Darwinian forces have selected truth-tracking beliefs. In these three regards, the naturalist approach seems to stand in a better position than the non-naturalist approach.

Street's response to Copp split into two parts. She firstly distinguishes two readings of Copp's society-centered theory: morality/reasons internalist and morality/reasons externalist. Morality/reasons internalism says that practical reason necessarily follows morality. If an agent is morally required (or forbidden) to do something, then there must be a reason for her to do so (or not to do so). Morality/reasons externalism says just the opposite: an agent may not have reason to do (or to refrain from doing something) which is morally required (or forbidden). Then Street clarifies that the target of her Darwinian Dilemma is *normative* realism according to which there are realist practical reasons which are independent of our subjective evaluative attitude. For example, if it is a normative fact that we should pursue survival, then we have a reason to do so regardless of whether we would like to survive, whether we have a desire to survive and whether we think survival is good. If Copp's account is a form of morality/reasons externalism, then it does not posit the existence of this kind of independent reason. Accordingly, even though it might well be a form of *moral* realism, it would not be a genuine form of *normative* realism which is the target of Street's debunking argument.

But what if Copp's account is a form of morality/reasons internalism? Then it will tell us that there are independent reasons and thus is a genuine form of normative realism

targeted by Darwinian Dilemma. Street argues that if Copp's theory is thus categorized, then his reply is question-begging in the sense that it "merely trivially reasserts the coincidence between the independent normative truth and what the evolutionary causes pushed us to think; it does nothing to explain that coincidence". (2008, 214) Street then uses an analogy, which is extremely similar to Joyce's Waterloo belief pill, to illustrate her point:

> Suppose you hold certain views about the planet Jupiter—about its size, surface, number of moons, and so on. Then one day you learn that you acquired these views about Jupiter by having them implanted in you by a hypnotist who picked them out of a hat. You're concerned that this might not have been a reliable method for arriving at them. In answer to that concern, it is no help to repeat your views about Jupiter and then point out that those are the very same ones that the hypnotist picked out of the hat and implanted in you. No method of arriving at your views about Jupiter—no matter how bizarre and unreliable—could fail this test, since if having led you to your actual views is the test of a method's reliability, then whatever method you actually used will come out as reliable. (2008, 215)

The internalist and the externalist interpretations are the only two possible readings of any metaethical theory, including Copp's. And both interpretation seems to be refuted by Street. But it might be noticed that there is a shift from "moral" to "normative" in

Street's reply. Is the shift purely ad hoc in the sense of defining realism in order to make it either vulnerable or not targeted by the Darwinian Dilemma? I think the shift here is not ad hoc, not because I think that these two terms can be used interchangeably, but because it is quite clear that Street, at the very beginning of proposing the Darwinian Dilemma, states that the target is *normative* realism. When she deals with a possible naturalist reply in her paper, before having the chance of replying to Copp, she states that

> In order to count as realist, a version of value naturalism must take the view that facts about *natural-normative* identities (in other words, facts about exactly which natural facts evaluative facts are identical with) are independent of our evaluative attitudes. (2006, 137)

From here we can see the shift from "moral" to "normative" is not crafty because Street has long been pointing to *normativity* and theories about value in general. In short, if Street's criticism of Copp's reply is successful, then either it does not answer the debunking challenge more than reasserting the coincidence between evolved normative (also moral) beliefs and normative (also moral) facts posited by the realist.

Nonetheless, I think that there is more to say about the naturalist reply. Street holds that the *uncompromising* normative realism targeted by her Darwinian Dilemma is an interesting theory and thus a worthwhile quarry for two reasons. The first reason is

simply that quite a few of theorists appear to hold such position. The second reason is that such combination of moral realism and morality/reasons internalism seems to be the only way of vindicating the objective bindingness of morality, namely that morality is objectively reason-giving. But how about the combination of moral realism and morality/reasons externalism? This combination might fail to be a genuine form of *normative* realism but well remain as a genuine form of *moral* realism. Is this form of moral realism uninteresting so that the debunking camp should simply ignore it? For me, it is intuitively obscure that morality does not provide reason for acting so that such metaethical theory is unattractive at first sight. But I would like to step back a bit and remain open-minded regarding this complicated issue of practicality of morality. In the following, I shall set aside all these internalist-externalist issues (just like I set aside Joyce's practical clout argument against naturalism) and argue that no matter which form of naturalism the realist takes, the naturalist approach to the debunking challenge faces some universal epistemological obstacles. The reason why I am doing this is that my primary goal here, unlike Street's, is to examine whether *moral* realism, not normative realism, is epistemologically debunked by the Darwinian Hypothesis.

4(b) From Copp to naturalist reply in a general form

This final section proposes a general argument against all forms of naturalist reply. The

argument is simply an expansion of Street's criticism of the internalist reading of Copp's reply. My attempt here is to explain the epistemological obstacles which apply to the naturalist reply, in various forms, to the debunking challenge. First, we should look closer again at Street's criticism of Copp's internalist version:

> Our substantive normative views on how we have reason to live are (on the internalist reading of the theory) merely being taken for granted as at least roughly correct. These views are then gathered together, pruned and systematized, yielding a general claim to the effect that we have reason to act in accordance with codes that would best serve society's needs. But merely gathering together, pruning and systematizing our substantive normative views does nothing to save the alleged explanation from being trivially question-begging…no matter how bizarre and unreliable one's method of arriving at one's beliefs about the independent normative truth …, there is no way this approach to vindicating one's method could fail. By definition, one's starting set of views is going to be within reach of a pruned and systematized version of those very same views. So by definition, whatever method one actually used to arrive at one's starting set of views is going to have landed one within reach of a pruned and systematized version of those views. But it obviously doesn't follow from this that one's method was a good one that is likely to have landed one on the independent normative truth. (2008, 215)

Then she concludes that Copp's account "has failed to give any satisfactory explanation

of why evolutionary forces influenced our views on how we have reason to live in such a way that they track the independent normative truth to an epistemically sufficient degree; it merely trivially assumes that they did." (2008, 216)

Here I shall summarize the criticism in this way. First, the naturalist reply assert the natural-normative identity or constitution according to which moral facts *M* is identical to or constituted by certain ordinary non-moral natural facts *N*. Second, the naturalist claims that once we confirm that there is no surprise if these natural facts *N* can be integrated into the evolutionary picture, then we can simply conclude that evolutionary forces track those natural facts which are identical to or constitute moral facts. As a result, evolutionary forces would track the moral facts to an epistemologically sufficient degree so that the skeptical worry could be evaded. But Street's point is that our views on the natural-normative identity, from the very beginning, are already heavily influenced by evolutionary forces. Taking these evolution-laden views for granted, of course the synthesis of these views would be evolutionary-explanation-friendly. But it does not explain why these views and their synthesis, yielded and promoted by evolutionary forces, represent independent truths.

I think that this criticism can be generalized and applied to naturalist replies in general.

Let's consider the variety of epistemological picture which can be proposed by the naturalist. Moral naturalism can be firstly divided into two groups: analytic naturalism and synthetic naturalism. The representative of the former is Frank Jackson's moral functionalism. Copp also claims that his society-centered theory is a form of moral functionalism (2008, 198), so I shall include him in this group. The representative of the latter is Cornell realism and certain kinds of reductionism. The criterion which distinguishes them lies in the type of epistemological approaches they respectively provide. The analytic naturalist holds that we discover moral truths and the meaning of moral terms by conducting a relational or network analysis of these terms. Once we identify the network, we can know what these terms means and find their functional realizer. Use Copp's account as an example. His account says, roughly, that morality has the function of best serving the needs of a society. Accordingly, "a behavior X is moral" simply means "X is permitted by a set of moral standards the currency of which would best serve the societal needs. Once we find the meaning of all relevant terms, we can know what "moral" means in an analytic manner and also find its corresponding natural realizer(s). The synthetic naturalist claims that we need more than the functional meaning of relevant terms in order to discover moral truths, just like how we discovered "water is H2O".

Here I shall argue that both types of naturalism face serious methodological obstacles. The traditional challenge faced by the analytic branch of naturalism would be the open-question argument according to which there is no promising candidate for identity statement in the form of "M (moral property) = N (non-moral natural properties)" such as "goodness=pleasure" or "rightness=maximization of overall utility" because the meanings on two sides cannot be analytically the same so that "goodness=goodness" is different from "goodness=pleasure". However, I think that the functionalist can roughly reply that once we have a more delicate candidate than "pleasure" and "maximization of overall utility" after adequately undergoing reflective equilibrium, the analyticity would come out more easily. This might not be a satisfying reply but I shall stop going deeper on this issue because evolutionary debunking has much to do with this open-question issue. So, I shall simply assume that moral functionalism cannot overcome this traditional challenge and move on to the debunking challenge it will suffer.

According to the Darwinian hypothesis, the whole network including the set of moral platitudes with which we are complying is heavily influenced by the evolutionary forces. That is to say, the way we use the moral terms and which natural properties we identify with these moral terms are massively shaped by evolutionary forces. No matter how delicate the synthesis of these terms is, the moral functionalist does not seem able to

make connection to independent moral truths. As Street argues against the normative realist reading of Copp, the functionalist cannot simply reassert that the moral platitudes to which natural selection have given rise indicate independent truths.

One might think that the synthesis naturalist like the Cornell realist and the reductionist will suffer less from this problem because the method of discovering moral truths is synthetic and a posteriori, just like discovering the truth that "water is H2O", so that they need not relying on evolutionarily-laden semantics. But I think that the problem is just the same. Yes, natural selection will pick out certain natural properties which are candidates nominated by the synthetic naturalist for being identical to or constituent of moral properties, but how do we know about the identity or the constitution? The evolutionary story seems to suggest that any way in which we investigate such identity or constitution would be heavily driven by evolutionary forces. So, here we have a parallel problem for the synthetic naturalist.

Conclusion

In the above chapters, I have summarized the debunking debate on whether moral realism is undermined by the success of the Darwinian Hypothesis that the evolutionary approach explains morality. Also, I have clarified crucial ideas invoked by both the debunkers and the realists. The central aim of this thesis is to critically investigate whether the prominent replies proposed by moral realists are convincing. After examining both non-naturalist and naturalist responses, I conclude that realists in question have not provided convincing solutions to the epistemological problems invoked by the evolutionary debunkers. As a result, I suggest that realists need to seek for new solutions or perhaps the only choice they is left with is to refute the empirical plausibility of the Darwinian Hypothesis.

Bibliography

Carruthers, Peter. 1992. *Human Knowledge and Human Nature*. Oxford: Oxford University Press: 109–10, 183–87.

Copp, David. 2008. "Darwinian Skepticism About Moral Realism." *Philosophical Issues* 18 (1): 186–206.

Cosmides, Leda, and Tooby, John. 1997. "Evolutionary psychology: A primer." Accessed April 10, 2017. University of California, Santa Barbara, Center for Evolutionary Psychology Web site: http://www.cep.ucsb.edu/primer.html.

Darwin, Charles R. 1866. *The Origin of Species*. 4th ed. London: John Murray.

Enoch, David. 2010. "The Epistemological Challenge to Metanormative Realism: How Best to Understand It, and How to Cope with It." Philosophical Studies 148 (3): 413–438.

FitzPatrick, William J. 2015. "Debunking Evolutionary Debunking of Ethical Realism." Philosophical Studies 172 (4): 883–904.

———. 2016. "Misidentifying the Evolutionary Debunkers' Error: Reply to Mogensen." Analysis 76 (4): 433–437.

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

Joyce, Richard. 2006. The Evolution of Morality. MIT Press.

Joyce, Richard. 2006. The evolutionary debunking of morality. *Reason and responsibility: Readings in some basic problems of philosophy*: 527-537.

Plantinga, Alvin. 2000 *Warranted christian belief*. Oxford University.

Ruse, Michael. 1986. *Taking Darwin Seriously*. Oxford : Basil Blackwell.

Shafer-Landau, Russ. 2012. “Evolutionary Debunking, Moral Realism, and Moral Knowledge.” Journal of Ethics and Social Philosophy 7 (1).

Singer, Peter. 2005 "Ethics and intuitions." *The Journal of Ethics* 9, no. 3-4: 331-352.

Spencer, Herbert. 1896. *The principles of biology*. Vol. 1. D. Appleton.

Street, Sharon. 2006. “A Darwinian Dilemma for Realist Theories of Value.” Philosophical Studies 127 (1): 109–166.

———. 2008. “Reply to Copp: Naturalism, Normativity, and the Varieties of Realism Worth Worrying About.” Philosophical Issues 18 (1): 207–228.

———. 2009. “Evolution and the Normativity of Epistemic Reasons.” Canadian Journal of Philosophy 39 (sup1): 213–248.

———. 2011. “Mind-Independence Without the Mystery: Why Quasi-Realists Can’t Have It Both Ways.” Oxford Studies in Metaethics 6: 1–32.

Tropman, Elizabeth. 2014. “Evolutionary Debunking Arguments: Moral Realism, Constructivism, and Explaining Moral Knowledge.” Philosophical Explorations 17 (2): 126–140.

Vavova, Katia. 2014. “Debunking Evolutionary Debunking.” Oxford Studies in Metaethics 9: 76–101.

Wielenberg, Erik J. 2010. “On the Evolutionary Debunking of Morality.” Ethics 120 (3): 441–464.

Wilson, Edward O. *Sociobiology: The new synthesis*. Oxford, England: Belknap Press of Harvard U Press, 1975.