

A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:

<http://wrap.warwick.ac.uk/167306>

Copyright and reuse:

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it.

Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

Evolutionary Debunking Arguments and Explanatory
Constraints On Belief

By

Christopher Noonan

Submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy in Philosophy

University of Warwick, Department of Philosophy

December 2021

Table of Contents:

Acknowledgements: 5

Declaration: 6

Abstract: 7

Introduction: 8

Chapter One: Evolutionary Debunking Arguments: 12

- 1.1: Introduction: 12
- 1.2: Ruse: 14
- 1.3: Joyce: 24
- 1.4: Street: 33
- 1.5: Similarities and Differences: 40
- 1.6: A Common EDA: 44
- 1.7: Ambiguities: 44
- 1.8: Commitments: 46
- 1.9: Conclusion: 56

Chapter Two: Minimalist Responses to Debunking Arguments: 57

- 2.1: Introduction: 57
- 2.2: The Modal Security Response: 58
- 2.3: Third-Factor Replies: 65
- 2.4: Reliance on Moral Beliefs: 72
- 2.5: What Defeats Our Beliefs: 75
- 2.6: A Dialectical Point: 78
- 2.7: Conclusion: 84

Chapter 3: Explanationism vs Modal Security: 85

- 3.1: Introduction: 85
- 3.2: Explanation: 86
- 3.3: Standards of Justification: 88
- 3.4: Counterexamples to Modal Security: 93
- 3.5: These Examples Fail to Support Explanationism: 100
- 3.6: Modal Security *via* Explanationism: 111
- 3.7: Defeat Without Explanatory Concessions: 112
- 3.8: Conclusion: 115

Chapter 4: Objections to Explanationism and (IEC): 116

- 4.1: Introduction: 116
- 4.2: TF-connected beliefs: 117
- 4.3: An Inclusive Explanatory Constraint: 121
- 4.4: Future Beliefs: 124
- 4.5: Not Necessary For Defeat? 126
- 4.6: (IEC) and Minimalist Replies: 131
- 4.7: Alternative Explanatory Constraints: 132

- 4.8: An Objection to These Explanatory Constraints: 135
- 4.9: Back to (IEC): 141
- 4.10: Conclusion: 146

Chapter Five: An Expanded Evolutionary Debunking Argument: 147

- 5.1: Introduction: 147
- 5.2: The Structure of Normative Reality: 149
- 5.3: Non-naturalism and the Structure of Normative Reality: 151
- 5.4: Minimalist Replies and the Structure of Normative Reality: 156
- 5.5: The Basic Argument: 159
- 5.6: The First Non-Basic Argument: 163
- 5.7: The Second Non-Basic Argument: 168
- 5.8: Minimalist Replies to Debunking Arguments Fail: 171
- 5.9: Just a Small Miracle: 174
- 5.10: Going Modal: 177
- 5.11: Conclusion: 179

Chapter 6: Inferring From Our Non-Basic Beliefs: 181

- 6.1: Introduction: 181
- 6.2: Leveraging Justification From Our Non-Basic Beliefs: 182
- 6.3: (IEC) and Attenuated Explanatory Links: 185
- 6.4: Why This Response Fails: 188
- 6.5: An Independence Clause: 196
- 6.6: An Objection: 199
- 6.7: Conclusion: 202

Conclusion: 204

Abbreviations: 208:

Bibliography: 209:

List of diagrams:

- 1). Figure 1: 66
- 2). Figure 2: 67
- 3). Figure 3: 118
- 4). Figure 4: 118
- 5). Figure 5: 123
- 6). Figure 6: 125
- 7). Figure 7: 136
- 8). Figure 8: 136
- 9). Figure 9: 143
- 10). Figure 10: 143
- 11). Figure 11: 145
- 12). Figure 12: 150
- 13). Figure 13: 150
- 14). Figure 14: 163
- 15). Figure 15: 164
- 16). Figure 16: 166
- 17). Figure 17: 184
- 18). Figure 18: 188
- 19). Figure 19: 193
- 20). Figure 20: 200

Acknowledgements:

Special thanks to Fabienne Peter and Kirk Surgener. Thank you both for being so generous with your time and your feedback, and for being so patient during the development of these ideas. Your insight has been invaluable.

Thanks to Josh Turner for helping me with the design of the diagrams.

Finally, thanks to Mum, Dad and Ben. Thank you for being excited during the good times and supportive during the bad.

Declaration:

All material contained in this thesis is my own work.

The work in this thesis has not been submitted for a degree at any other University.

Abstract:

Evolutionary debunking arguments (EDAs) claim that the evolutionary origins of our moral beliefs imply that those beliefs cannot be justified under the assumption of moral realism. In chapter one I outline three prominent evolutionary debunking arguments in the literature, and in chapter two I outline two types of “minimalist” replies to debunking arguments. These replies grant that our moral beliefs are not explained by the moral facts and then rely on substantive moral claims to show that our moral beliefs might still be reliable.

One way of objecting to these minimalist responses is to claim that the explanatory disconnect implied by debunking arguments defeats our moral beliefs, and so minimalist replies are not entitled to rely on substantive moral claims once they have granted this explanatory disconnect. In chapter 3 I argue that information can undermine our beliefs even if does not imply that our beliefs are modally insecure, and that this removes one way that the minimalist might deny that the explanatory disconnect implied by debunking arguments defeats those beliefs. In the next chapter I argue for a specific explanatory constraint on belief: (IEC). I point out, however, that according to (IEC), EDAs do not, by themselves, establish the explanatory disconnect that defeats our belief.

In chapter five I supplement EDAs with my own arguments which, when combined with EDA's, establish that our moral beliefs are defeated according to (IEC). When EDAs are combined with these arguments, minimalist replies fail because they crucially rely on defeated moral beliefs. In the final chapter, I consider and reject a potential avenue of response by the minimalist, and considering this response sheds further light on the kind of attitude we can have towards our belief that defeats that belief.

Introduction:

Roughly a year before starting my PhD, I read Sharon Street's "A Darwinian Dilemma for Realist Theories of Value" (2006). It struck me as making an unusually strong case against moral realism; if our moral beliefs really can be given an evolutionary explanation, and the forces of natural selection really are indifferent to realistically construed moral facts, then this seems to imply that our moral beliefs are likely to be largely off-track. As Street puts it, our moral beliefs could only be *coincidentally* true, and it seemed to me impossible that the realist could accept this point and then continue to justifiably hold their moral beliefs. The moral realist, I felt, was in hot water (and not just the kind of hot water in which we all find ourselves, from time to time, when confronted with generic sceptical arguments against the possibility of knowledge or justified beliefs).

I then read a number of *other* arguments, grouped together under the title of "evolutionary debunking arguments", that seemed to express the same underlying concern in different ways. I felt that Street's particular formulation of the argument was a particularly vivid and convincing portrayal of a problem that had been expressed or gestured towards by other authors as well. The problem, roughly, is that we ultimately only hold our moral beliefs because it was evolutionarily beneficial to do so, and it is therefore difficult to see why we should expect those moral beliefs to correlate with realistically construed mind-independent moral facts. And if we accept that they do not correlate with such facts, then (under the assumption of moral realism) we seem forced to accept some unappealing form of moral scepticism.

In chapter one of this thesis I outline three such evolutionary debunking arguments from Michael Ruse (1986), Richard Joyce (2005), and Street (2006). I then try to distil what I take to be the common thought that underlies them all: the evolutionary origins of our moral beliefs imply that our moral beliefs are not *explained by* the moral facts. And, the thought goes, if our moral beliefs are not explained by the moral facts, then these moral beliefs are ultimately unjustified.

As much as I initially found evolutionary debunking arguments against moral realism convincing, I then started to find a particular style of objections to these arguments

unconvincing. These responses (which have since been dubbed “minimalist” responses to debunking arguments¹) made substantive moral assumptions in the process of arguing that realistically construed moral beliefs might be justified *even granting* that our moral beliefs are not explained by the moral facts². I felt that evolutionary debunking arguments posed a serious problem for our realistically construed moral beliefs, and something struck me as too easy about this proposed style of response. The primary goal of this thesis will be to assess minimalist responses to debunking arguments and ultimately argue that those responses fail.

In chapter two, I outline two different kinds of minimalist responses to debunking arguments. The first is the modal security response, which aims to show that our moral beliefs can satisfy modal conditions like sensitivity or safety even if our moral beliefs are not explained by the moral facts (Clarke Doane, 2015; Jonas Olson, 2019). The second kind of minimalist response are third-factor replies, which point out that our moral beliefs and the moral facts might both be explained by some third-factor, even if neither one explains the other (Enoch, 2010; Erik Wielenberg, 2010; Kevin Brosnan, 2011; Knut Skarsaune, 2011). Both kinds of response crucially assume at least some of our moral beliefs are accurate in the process of replying to debunking arguments.

I then note one potential diagnosis of why these responses fail. Perhaps the explanatory disconnect implied by evolutionary debunking arguments defeats the realists moral beliefs, and so once the realist accepts this disconnect, they are not entitled to rely on their moral beliefs because such beliefs have all been defeated (Matt Lutz, 2018; Daniel Korman and Locke, 2020).

I argue that there are two ways that we might resist this diagnosis, each of which corresponding to one of the minimalist responses that I have just outlined. The first is to argue that modal considerations, and not explanatory ones, are really crucial when it comes to information undermining our beliefs (Clarke-Doane, 2015; Clarke-Doane and Dan Baras, 2020). The second is to argue that the specific explanatory connection implied by debunking arguments does not defeat that belief, because there might still

¹ Dustin Locke (2014), as far as I am aware, is the first person to use this terminology to refer to these responses.

² See, for example, Justin Clarke-Doane, (2015) and David Enoch (2010).

be *alternative* explanatory connections that hold between our moral beliefs and the moral facts.

In chapters three and four I tackle each of these possibilities in turn. In chapter three, I argue that accepting that our belief in P has no explanatory connection with P can defeat that belief *regardless* of whether it implies that the belief could easily have been false in some close possible world. This implies that evolutionary debunking arguments can defeat our realistically construed moral beliefs even they do not imply that those beliefs fail some modal condition like sensitivity or safety.

In chapter four, I argue for a particular “explanatory constraint”³ on belief: (IEC). This is a principle that stipulates the kind of attitude we could have towards the explanatory history of our belief that could defeat our justification for holding that belief. According to this constraint, our belief in P is unjustified if we withhold belief that our belief is explained by P *and* we withhold belief in some third-factor that explains them both.

I argue for (IEC) by considering and rejecting alternative constraints, and by defending (IEC) against objections that have been raised against related principles. I end this chapter by arguing that evolutionary debunking arguments, as they are usually construed, do *not* imply that our moral beliefs are defeated given (IEC), because they are consistent with the possibility that there is some third-factor that explains both our moral beliefs and the moral facts. This, according to (IEC), is enough for our moral beliefs to remain justified.

I do not therefore conclude, however, that minimalist replies to evolutionary debunking arguments are successful. This is because there are independent arguments that can be attached to evolutionary debunking arguments such that they do in fact render minimalist responses void. In chapter five I present these arguments, which stem from the fact that, according to structure of normative reality to which non-naturalist realists⁴ are committed, there are good reasons for thinking that our basic moral beliefs cannot be explained by some third factor that also explains the basic moral facts. When these considerations are combined with evolutionary debunking arguments, it implies that the minimalist’s basic moral belief fails (IEC). And, once

³ The term “explanatory constraint” is borrowed from Korman and Locke (2020).

⁴ I provide an account of non-naturalist realism below.

basic moral beliefs are off the table, the non-basic beliefs are soon to follow. This implies that minimalist replies fail because they crucially rely on defeated moral beliefs.

Finally, I devote chapter six to considering a potential objection to this argument, according to which the realist's basic moral beliefs can be justified because they are inferred on the basis of beliefs in non-basic moral facts. I argue that this response fails, and the reason that it fails sheds further light on the conditions under which our beliefs can lose their justification.

This concludes my assessment of evolutionary debunking arguments and minimalist replies. Traditional evolutionary debunking arguments have been incomplete, and this possibility left open by such arguments has been exploited by minimalist replies to them. Fortunately for the debunker, there are ways of plugging this gap which ensure that minimalist replies to debunking arguments fail. I conclude by considering what alternative avenues of response are available to the realist, and by indicating where further research is needed.

Chapter One: Evolutionary Debunking Arguments.

1.1: Introduction:

In this chapter I will outline three of the major evolutionary debunking arguments in the literature. The purpose of doing so is to provide an account of some of the most influential evolutionary debunking arguments, and to get a general sense of why the evolutionary origins of our moral beliefs have been thought to have troubling implications for certain meta-ethical positions. My purpose is also to draw attention to the various ways in which various debunking arguments can and have differed from each other, which will set up my discussion of the similarities and differences between these arguments in section 1.5. With these in mind, in section 1.6 I will try to formulate an evolutionary debunking argument that retains the spirit of the debunking arguments so considered, and avoids extraneous issues that are irrelevant to the goals of this thesis. The conclusions that I will ultimately draw about evolutionary debunking arguments in later chapters will therefore apply most directly to my own specific interpretation of such arguments, but I think that interpreting debunking arguments in this way is worth-while for reasons that will be outlined towards the end of this chapter.

As the name indicates, “evolutionary” debunking arguments crucially feature a hypothesis about the development of our capacity for moral judgements through natural selection (call this the empirical premise)⁵. As such, they typically involve an account of the empirical evidence that is supposed to support this hypothesis. However, the empirical discussion in the following overview of such arguments will be extremely brief. My justification for this is both practical and philosophical. On the practical side: I am not an evolutionary biologist, and I am not qualified to give a comprehensive or persuasive account the evolutionary considerations that debunkers typically use to motivate the empirical premise of their arguments.

Philosophically, defending the empirical premise is not, strictly speaking, necessary for the primary goal of this thesis, which is to determine what follows philosophically *under the assumption* that the empirically motivated premise is true. The minimalist

⁵ I flesh out this empirical premise below.

responses to debunking arguments that I will be considering and arguing against for the bulk of the thesis all grant, for the sake of argument, the empirical premise that evolutionary debunkers defend. They then deny that this empirical premise, if true, would have any major troubling implications for the targeted meta-ethical position. In order to argue against these responses, then, I just have to establish what follows philosophically if we assume this empirical premise is true, rather than having to establish the empirical premise itself.

Still, I have included some minor account of these empirical considerations. I have done so because understanding the way in which evolution influenced our capacity for moral judgements is necessary to fully understand why debunkers have taken this influence to have meta-ethical ramifications. I have also included these brief passages in order to justify the philosophical investigation that follows; it is hard to see the point in establishing the philosophical implications of these empirical premises without any sense at all of why these premises might be plausible. In order to motivate the discussion of the following five chapters, then, I provide some minimal comments on why debunkers have found the empirical premise worth endorsing.

When doing so, I use the terminology that is used by debunkers themselves, and I do so without meaning to imply that this terminology is appropriate (let alone optimal). If there are grounds to object to either their use of empirical evidence, or the terminology that is used to express it, then this could make for a compelling and important objection to evolutionary debunking arguments. But making or assessing such an objection is not among the goals of this thesis⁶.

In focusing on just three evolutionary debunking arguments, I have necessarily had to be highly selective, and I have ignored a number of evolutionary debunking arguments presented by other commentators⁷. I have chosen to look into these three debunking arguments specifically because they are among the most influential, and they are the debunking arguments that minimalists respond to most directly. As I am concerned with assessing these minimalist responses, it makes sense to outline the specific versions of debunking arguments to which they respond.

⁶ For an assessment of the empirical details of evolutionary debunking arguments, see Fraser (2014) and Levy and Levy (2020).

⁷ For alternative evolutionary debunking arguments, see Singer (2005), Greene (2007), Kahane (2011), Morton (2016).

One final caveat: as we will see, evolutionary debunking arguments differ on the meta-ethical positions that they claim are vulnerable to their arguments. In what follows I will use the meta-ethical terminology and taxonomy that is used by the debunkers themselves, and I wait until section 1.8 to outline how I will be interpreting the target of such arguments.

1.2: Ruse:

Perhaps the first fully formed evolutionary debunking argument is in Michael Ruse's *Taking Darwin Seriously* (1986). In the final chapter, Ruse argues that our "moral capacity" is an evolutionary development. What does Ruse mean by our moral capacity?

Importantly, this moral capacity should be distinguished from our capacity to have mere preferences for and against certain courses of action. According to Ruse:

"The Darwinian's claim is that we have genetically based dispositions to disapprove of certain courses of action and to approve of other courses of action. But they are more than mere likes and dislikes. Here we start to move towards genuine morality and its evolution" (1986: 221).

So the moral capacity, for Ruse, is something more than a capacity to have preferences for or against certain courses of action. What is this "something more"? Further down, Ruse writes:

"It is absolutely fundamental to the Darwinian case that, in order to spur us into action – perhaps indeed to go against other self-directed emotions – we have rules incorporating that prescriptive force which is distinctively characteristic of morality. As in the case of sibling incest, our feelings are backed by a (likewise innate) sense that approved actions are "right" and that disapproved actions are "wrong".

It is not just that we do not want to go to bed with our siblings. We feel that we ought to not have intercourse with them" (1986: 221-2)

And later:

"we are talking of more than a mere feeling that we want to help others. It will be an innately based sense of obligation towards others [...] The Darwinian's point is that our moral sense is a biological adaptation, just like hands and feet. We think in terms of right and wrong." (1986: 222).

The idea seems to be that our moral capacity involves our ability to feel that by failing to behave in a certain way we would be transgressing some rule that applies universally and does not depend on our subjective preferences (Ruse, 1986: 68). Our moral capacity does not merely allow us to feel repulsed by incest; it enables us to feel as though there is an obligation against incest that applies indiscriminately both to those who find it repulsive and those who find it enjoyable.

These feelings involve what we *ought* or *should* do, and they are not feelings about how people do in fact behave (Ruse, 1986: 69). The feeling is that we *ought not* engage in incest, and this is distinct from the feeling that people do or not do engage in incest. This seems to be the “prescriptive force” which Ruse takes to be “distinctly characteristic of morality” (1986: 222).

By “moral capacity”, then, Ruse means our capacity to view the world as containing these sorts of universal and prescriptive rules. Specifically, Ruse characterises our moral capacity in terms of our feelings, senses or “basic moral inclinations” (1986: 235) rather than having it include fully formed reflective moral beliefs (though as will emerge later, Ruse, like other debunkers, takes there to be an intimate connection between this basic capacity for moral feelings and our more fully formed moral beliefs and theories). It is our capacity to feel as though the world contains these rules for which Ruse seeks to provide an evolutionary explanation.

To argue for this evolutionary explanation of our moral capacity, Ruse first cites two mechanisms by which evolution is likely to have encourage co-operative behaviour among humans: reciprocal altruism and kin selection (1986: 219-20). Reciprocal altruism describes the process of behaving in ways that increase other organisms’ reproductive success because doing so makes it more likely that those organisms will behave in-kind towards us, thus increasing our own chances of reproductive success. Kin-selection involves behaving in ways that increase the reproductive success of closely related organisms that share our genes, thus increasing the chances of those shared genes being passed on to future generations. Though the evolutionary benefit of reciprocal altruism relies on co-operative behaviour being returned, co-operative behaviour towards one’s kin can be evolutionarily beneficial even if our kin do not return the favour and behave co-operatively towards us.

Through these mechanisms, Ruse argues that evolution encourages “genetically based dispositions to approve of certain courses of action and disapprove of other courses of action” (1986: 221). In order to effectively steer us away from self-interest, these dispositions are “backed by a (likewise innate) sense that approved actions are right and that disapproved actions are wrong” (1986: 222). Thus, evolution encouraged us to have a specifically moral sensibility, and it did so to encourage us to behave co-operatively towards one another.

What evidence is there in favour of this hypothesis? Ruse explains that, according to Darwinism, adaptations evolve very slowly. If our moral capacity is the result of evolution, we should therefore expect it (or some earlier signs of it) to also be present in our close biological relatives (1986: 227). Our divergence from these relatives is likely to have occurred more recently than the beginning of the development of our moral capacity, implying there should be some trace of this moral capacity in these biological relatives. Thus, in support of the claim that our moral capacity is an evolutionary adaptation, Ruse cites evidence concerning co-operative behaviour among higher-level primates with patterns one would expect of a species that has been subject to reciprocal altruism and kin-selection (1986: 228-30).

Furthermore, Ruse cites anthropological evidence that helping behaviours in humans between kin and non-kin are different in just the way we would expect if our moral capacity developed because of kin-selection and reciprocal altruism (1986: 232-3). Altruistic behaviour towards kin is performed unconditionally, without expectation of reciprocation. By contrast, altruistic behaviour towards non-kin is typically conditional on the prospect of reciprocation, with such behaviour being predicated on the expectation of being paid back. This is what we would expect if our altruistic behaviour were explained in part by kin selection and, between non-kin, by reciprocal altruism. Informed by a background of knowledge the potential for evolutionary pressures to influence behaviour, (1986: 230-31), Ruse concludes that (while this evidence is far from decisive) we at least have a “strong hypothesis” for the “evolution of human moral capacity” (1986: 234).

Ruse uses this claim about the evolution of our moral capacity to argue for a particular meta-ethical position. In making this argument, Ruse divides meta-ethical theories into “objective” and “subjective”. According to objective meta-ethical theories, moral

norms exist independently of humans (and/or human emotions), and we “intuit or otherwise rationally grasp” morality (1986: 214). Moral norms are “conditions laid upon us, independent of our contingent nature” (1986: 214). The general idea behind Ruse’s characterisation of objectivism seems to be that there are moral truths that hold independently of, and are not constituted by, the human moral capacity, and in order to have correct moral beliefs we humans need to accurately represent these independent moral truths. These human-independent moral norms might be grounded in God’s will, or “might exist in a kind of disembodied way, as it were in a supersensible world” (1986: 215).

By contrast, according to subjectivist meta-ethical theories, “without humans there is no right and wrong” because morality depends on human feelings and inclinations (1986: 216)⁸. There are no independent moral norms laid upon us from outside; the nature of morality is just a function of human nature. Ruse cites emotivism and prescriptivism as two major examples of subjectivist theories (1986: 216). Again, the main thought for Ruse here seems to be one of dependence; the nature of morality depends on the moral capacity of humans, such that if our moral feelings and inclinations changed then so would the content of morality, and without the human moral capacity there would be no morality to speak of.

Ruse argues that the evolutionary origins of our moral capacity imply that we should accept subjectivism rather than objectivism. How does Ruse move from a point about the evolutionary origins of our moral capacity to the further claim that there are no moral facts existing independently of our moral capacity? Or, as Ruse puts it, “Could you not accept virtually everything the Darwinian claims about morality being a human phenomenon [...] and yet argue that our human faculties are perceiving or intuition objective truth?” (1986: 254). Why does providing an evolutionary explanation of our moral capacity rule out the possibility that objective morality exists, or that our moral capacity is accurately representing these phenomena?

For the sake of the proceeding analysis of Ruse’s argument (and the proceeding thesis as a whole) it pays to quote Ruse’s response to this question in full:

⁸ Presumably the traces of a moral capacity found in non-human animals is too primitive to actually count as morality, hence the focus on human nature.

“You would believe what you do about right and wrong, irrespective of whether or not a “true” right and wrong existed! The Darwinian claims that his/her theory gives an entire analysis of our moral sentiments. Nothing more is needed. Given two worlds, identical except that one has an objective morality and the other does not, the humans therein would think and act in exactly the same ways.

Hence the objective foundation of morality is redundant. But, surely, the last thing the objectivist can accept is that his/her principles are redundant. ‘God wants us to be good, and that is the ultimate defining source of good, but it wouldn’t matter whether He did or not!’ If anything is ruled out by objectivism, it is this. Thus, we must conclude that not only is Darwinian ethics a subjectivist ethics, but it is one which positively excludes the objectivist approach.” (1986: 254)

These two paragraphs contain the major steps in Ruse’s argument from the empirical claim that our moral capacity is an evolutionary development to the meta-ethical claim that objectivism is false. In the rest of this section, I will try to unpack this passage in order to get clear on the structure of Ruse’s argument.

The first thing to note about the first paragraph is that the focus is just as much on moral *beliefs* as it is on more basic inclinations and moral feelings. Ruse claims that our moral beliefs would be unchanged if objectively construed moral facts did not exist. Ruse therefore thinks that the point about the evolutionary origins of our more basic moral capacity also has implication for our more sophisticated moral beliefs as well. Presumably this is because, according to Ruse, our more fully-formed moral beliefs are an outgrowth of these more basic moral dispositions, and so if we give an evolutionary account of the origins of our more basic moral capacity we also thereby give an account of the origins of our more sophisticated moral beliefs (1986: 223).

A second, more subtle about this first quoted paragraph is that Ruse makes two distinct claims about the relationship between our moral capacity and/or moral beliefs and objectively construed moral facts, and it is not clear how each claim relates to each other, or which one is supposed to entail the redundancy that sinks objectivism.

One of these claims concerns the explanation of our moral capacity and/or beliefs. It says that our moral capacity and/or beliefs can be entirely explained without reference to objective moral facts. We can call this the explanatory claim. The Darwinian does not need to invoke these objectively construed moral facts in order to provide a plausible account of why we have moral thoughts and feelings. Nothing in Ruse’s

account of why we have a moral capacity makes reference to, or presupposes, mind-independent moral facts. Hence, unlike an evolutionary explanation of our other capacities for belief (like mathematics), the evolutionary explanation of our moral capacity implies that moral facts nowhere explain that capacity (1986: 254)⁹.

Now, Ruse himself acknowledges the influence of cultural forces in shaping our moral beliefs, and so it is unlikely that he takes the evolutionary story that he argues for to provide a *complete* explanation of those beliefs (1986: 223). I therefore think we should read Ruse (and, as we shall see, subsequent evolutionary debunkers) as claiming that, if the evolutionary explanation of our moral beliefs is true, then it implies that the complete explanation of our moral beliefs will not include moral facts. This is because plausible ways of adding to and completing this explanation do not involve positing objectively construed moral facts. It is hard to see how the cultural portion of the explanation of our moral beliefs could make reference to such moral facts, and the same goes for any other way in which we might try to fill out this genealogy of our moral beliefs (Street, 2006: 155). Once the evolutionary explanation of our moral beliefs has been established, then, we have reason to think the complete explanation of our moral beliefs will not involve moral facts. This is how Ruse can move from the admittedly incomplete evolutionary explanation of our moral capacity/beliefs to the claim that moral facts do not explain our moral capacity/beliefs. This explanatory claim is one of the two points that Ruse makes in the first paragraph of the quoted passage.

The second is a modal one: that our moral capacity and beliefs would be the same in the counter-factual scenario in which there were no objective moral facts. If mind-independent objective facts did not exist, then nothing about our moral sensibility or beliefs would change. In the first and last sentence of the first quoted paragraph Ruse is making this point, which I will call the modal claim.

In the quoted passage, then, Ruse argues from his evolutionary explanation of our moral capacity to both the explanatory claim and the modal claim, and then argues that these claims show that mind-independent moral facts are redundant in a sense that is unacceptable for the objectivist. Though closely related, the modal claim and the

⁹ This difference between an evolutionary genealogy of our mathematical and our moral beliefs will become clearer when outlining Joyce's evolutionary debunking argument below.

explanatory claim are distinct. It is worth briefly considering the relationship between the two because, as Korman and Locke (2020) point out (and as I will argue later) the viability of certain minimalist responses to debunking arguments actually depends on whether we should interpret debunking arguments as essentially relying on the explanatory or the modal claim. Outlining the relationship between the two will therefore be useful for the purposes of this thesis.

Very often, if the modal claim is true for two facts (i.e, if A would be unaffected if B did not exist) then so is the explanatory claim (B does not explain A). Similarly, if the explanatory claim is true then, usually, so is the modal one. The explanatory claim can be used as evidence for the modal claim: if our moral capacity is entirely explained by facts other than mind-independent moral facts, then in the counter-factual scenario in which there remains the capacity-explaining facts but no mind-independent moral facts, our moral capacity will remain unchanged. This is because, in this counter-factual scenario, everything that is currently responsible for our moral capacity will still obtain, and so presumably our moral capacity will get explained in exactly the same way in this counter-factual scenario as it does in the actual one.

Similarly, the modal claim can be used as evidence for, or even as an illustration of, the explanatory claim (Faraci, 2019). If A would exist exactly as it currently does even if B did not exist, then it is hard to see how B could explain A. If B did explain A, then the absence of B would presumably imply some sort of alteration in A, which would be left without one of the things that explains its existence in the actual world. And so if A would remain unchanged were B not to exist, this is usually good evidence for the claim that B does not explain A.

So in the above passage, Ruse makes both a modal claim and an explanatory claim. In the following paragraph, Ruse argues that it follows from these claims that objectively construed moral facts are redundant in a way that cannot be countenanced by the objectivist. But it is not clear whether the redundancy is supposed to follow more immediately from the explanatory claim or the modal claim (or perhaps a combination of both). Because of their close relationship, it may be that the explanatory claim is employed as evidence for the modal claim, and the modal claim is what ensures the relevant kind of redundancy. Alternatively, the modal claim may be employed as

evidence for the explanatory claim, and it is from the explanatory claim that the redundancy follows.

Accordingly, it is unclear whether the modal claim or the explanatory claim is really essential to Ruse's argument. If the explanatory claim is being used just as evidence for the modal claim, then (so long as the modal claim is still true) then the explanatory claim might be false and Ruse's argument would remain unaffected. Similarly, if the modal claim is employed primarily to support the explanatory claim, from which follows the redundancy of objectively construed moral facts, then the modal claim might be false, but (as long as the explanatory claim were still true) Ruse's argument would be unaffected.

Various commentators have taken Ruse to be arguing for the modal claim by way of the explanatory claim, with the former doing the more fundamental work in his argument (Wielenberg, 2010; Tomas Bogardus, 2016). What is supposed to defeat objectivism is that, counter-factually, we would have the very same moral capacity and/or moral beliefs even if there were no mind-independent moral facts, and the explanatory claim is made only in support of this modal claim. This interpretation of Ruse is supported Ruse's comment, slightly further down, where he writes: "Had evolution taken us down another path, we might well think moral that which we now find horrific, and conversely. And this is not a conclusion acceptable to the traditional objectivist." (1986: 254).

But, because of the close relationship between the explanatory and modal claims, this is not the only possible interpretation of Ruse's argument. It is also possible that the modal claim is being used as evidence for the explanatory claim, and it is the explanatory claim Ruse really takes to be anathema for the objectivist. I think that Ruse's later comments support this interpretation of his debunking argument. On the question of whether his argument transgresses the is/ought barrier, Ruse writes:

"[The Evolutionary Debunker] is trying to derive morality from a factual theory, in the sense of *explaining* our moral awareness, by means of the theory. More particularly, he/she is explaining away the apparent objective referent of the theory, that which gives it its binding prescriptive nature, and which takes it in meaning beyond the mere factual" (1986: 256).

And later:

“You cannot justify killing is wrong in the sense of deducing it from factual premises. What you can do is explain why we hold this belief. This is all that can or need be offered.” (1986:256)

I therefore think that Ruse takes his explanatory claim to imply the falsity of the objectivist’s position, and uses the modal claim in support of, or by way of illustrating, this explanatory claim. If this interpretation is correct, then Ruse’s evolutionary debunking argument can be characterised as follows: Given the evolutionary origins of our moral capacity, our moral capacity and/or beliefs are not explained by moral facts, as is illustrated by the fact that nothing would change if those moral facts did not exist. This implies that mind-independent moral facts are redundant, and accepting that moral facts are redundant in this sense is inconsistent with objectivism. Hence, we should reject objectivism and accept some kind of subjectivism, according to which morality is constituted by the feelings and dispositions of humans, rather than existing independently of those dispositions.

We can now get onto the questions that surround the concept of “redundancy”. Grant that our moral capacity/beliefs are not explained by objectively construed moral facts. In what sense does this imply that the objectivist’s mind-independent moral facts are redundant? Recall, according to Ruse, that the objectivist who accepts this kind of redundancy would be in a position to say: ““God wants us to be good, and that is the ultimate defining source of good, but it wouldn’t matter whether He did or not!”” (1986: 254). There certainly is something amiss about the objectivist who locates the objectivity of morality in God’s will and simultaneously affirms that God’s will does not matter at all.

It seems as though the objectivist is committed to normative facts being *normatively* significant in the sense that they do in fact determine what we ought to do. If the objectivist posits mind-independent moral facts, then they cannot simultaneously affirm that it would not matter whether or not those moral facts existed when it comes to things like what we would have reason to do, or what obligations we would have, or what would be of value. How could moral facts exist if they had no such implications? If Ruse has successfully shown that the objectivist is committed to moral facts being redundant in *this* sense, he will indeed have shown that moral facts are redundant in a way that cannot be countenanced by the objectivist.

But it is hard to see why the objectivist should accept this kind of redundancy, even if they accept the explanatory and/or the modal claim. The objectivist can accept that neither our moral capacity nor beliefs are explained by mind independent moral facts. But it seems as though, if these mind-independent moral facts exist, they would still have all the required normative implications about what we have reason to do, what obligations we have and what is of value. Their being non-redundant in *this* sense is in no way dependent on their explaining (or modally co-varying with) our moral capacity. So neither the modal nor the explanatory claim imply that objectively construed moral facts are redundant in this sense.

There is another sense of redundancy that *is* entailed by the explanatory claim. Given the explanatory claim, objectively construed moral facts are *explanatorily* redundant, in the sense that they are not needed to give a full explanation of our having a moral capacity and/or moral beliefs. Recall that Ruse states:

“[The Evolutionary Debunker] is trying to derive morality from a factual theory, in the sense of *explaining* our moral awareness, by means of the theory. More particularly, he/she is explaining away the apparent objective referent of the theory, that which gives it its binding prescriptive nature, and which takes it in meaning beyond the mere factual” (1986: 256).

This implies that, *when it comes to explaining why we feel and believe that objectively construed moral facts exist*, objectively construed moral facts are redundant. If this is the kind of redundancy to which Ruse is referring, then it follows from the explanatory claim because it essentially just *is* the explanatory claim. Our moral capacity and/or beliefs are not explained by the moral facts, and moral facts are therefore redundant in the sense that they do not explain our moral capacity and/or beliefs. According to Ruse, the objectivist cannot accept that moral facts are redundant in this sense, and so we should reject objectivism and accept subjectivism.

Summarising, if my interpretation of Ruse’s evolutionary debunking argument is correct, it goes as follows. Our moral capacity can be given an evolutionary explanation according to which our moral capacity and beliefs are nowhere explained by the moral facts (as evidenced by the fact that they would remain unchanged if the moral facts were different). Because moral facts are redundant when it comes to

explaining our beliefs and feelings about them, we should give up our belief in such facts and accept a kind of subjectivism according to which such facts do not exist.

So interpreted, Ruse's argument relies on some major assumptions (assumptions which, as we will see, various responders are quick to reject). Most importantly for the purposes of this thesis, it relies on the assumption that, if we accept that our feelings and beliefs about a particular kind of fact are not explained by those facts, then our belief in these facts is undermined and we should believe they do not exist. If these assumptions are true then, by giving an evolutionary explanation of our moral capacity according to which our moral capacity and beliefs are not explained by the moral facts, Ruse has shown that we should believe that human-independent moral facts of the sort that characterise objectivism should not exist. One natural response to rejecting the existence of such facts is to accept subjectivism, according to which objectively construed moral facts do not exist.

The epistemological ramifications of accepting that our belief is explanatorily disconnected from the fact that makes it true is something that I will devote considerable attention to later on and is a central issue in my discussion of minimalist replies to debunking arguments. But, for now, I forestall this issue and move on to a more recent evolutionary debunking argument.

1.3: Joyce:

Another major evolutionary debunking argument can be found in the work of Richard Joyce. The fullest version of Joyce's debunking argument is presented in his *Evolution of Morality* (2005), though an earlier version of this argument can be found in *The Myth of Morality* (2001), and Joyce makes further clarifications to his debunking argument in later work as well (2016). In this chapter I will mainly focus on Joyce's argument as it is presented in (2005), and will address Joyce's later clarifications as needed.

Like Ruse, Joyce begins his debunking argument with the claim that our capacity for moral thought is an evolutionary development. And, like Ruse, Joyce is keen to distinguish between a capacity for moral thought from a capacity to merely prefer certain courses of action over others. If we make a moral judgement that someone should or should not do something, we do not take this to be contingent upon whether

or not they would like to do it, and they cannot shrug off an apparent obligation towards some course of action by citing subjective preferences against that action. In this sense, we take moral prescriptions to be *inescapable* (2005: 62).

According to Joyce, we also take moral prescriptions be *authoritative* (2005: 62). Not only do they apply to everyone, regardless of their interests or preferences, but they are also taken to provide a genuine reason to behave in the required way. Though the norms of etiquette might apply to everyone regardless of preferences, we do not necessarily think that someone is making a mistake by ignoring them, or that they genuinely provide all agents with reason to behave in according with those norms. By contrast, moral norms do provide such a reason, and a person who ignores them is making a mistake by doing so.

Joyce calls refers to the combination of the inescapability and the authority of morality as its “practical clout” (2005: 62). Joyce’s claim is that our ability to think using moral concepts and to make moral judgements can be given an evolutionary explanation.

Joyce’s hypothesis is that evolution imbued us with a capacity for moral thought in order to help facilitate certain co-operative or helpful behaviour towards other organisms¹⁰. Like Ruse, Joyce hypothesises that evolution encouraged us to perform this behaviour by imbuing us with a “moral sense” that involves experiencing certain scenarios as *demanding* action. Desires can be fickle, and in order to curb rationalisation evolution encouraged us to believe that some courses of action were not only preferable but objectively required of us in the sense characteristic of morality (2005: 111).

Joyce cites a large range of empirical evidence in favour of this claim, and I will therefore have to be selective in my discussion of it (2005: 133-140). Joyce notes that, if morality were invention rather than something to which we were innately predisposed, we might expect its invention to be attributable to a specific civilisation, and its subsequent influence to be traceable through later civilisations. But such a history is not available; a capacity for moral thought seems to have always existed, implying that a disposition to make moral judgements is innate rather than learned

¹⁰ Like Ruse, Joyce cites reciprocal altruism and kin selection as processes through which such behaviour might have been encouraged, but Joyce also cites two more: mutualism and group selection (2005: 13-45). For reasons of space, I do not include this in my account of Joyce’s debunking argument.

(2005: 135). This supports the contention that our capacity for moral judgements is an evolutionary adaptation rather than a cultural invention.

Joyce then bolsters this point with evidence from developmental psychology about the development of the moral sense. Young children can distinguish between moral norms and prudential norms without explicit instruction. This implies that the human brain is “prepared” to make moral judgements in the same way it might be prepared to acquire languages (2005: 135). The specific course of moral development in children “exhibits an extremely reliable sequence, it gets underway remarkably early, its developmental pathway is distinct from the emergence of other skills, and its unfolding includes abrupt maturations” (2005: 135). Joyce cites Haidt (2001: 826-7), who notes a sudden shift, around age 4, when children become sensitive to issues of fairness and start over-applying them at every possible opportunity. The reliability of the specific pattern of moral development across cultures strongly suggests that the capacity to form moral judgements is innate, rather than having been acquired in response to some cross-culturally consistent external factor (2005: 137).

Joyce then cites the “large body of body of research in developmental psychology” that suggests this development tracks genuinely *moral* judgements (2005: 135). Again cross-culturally, children as young as 3 demonstrate a sensitivity to the difference between conventional and moral transgressions, with the latter being more serious, more generalisable, and independent of authority. Joyce explains:

“Concerning a conventional transgression, such as a boy wearing a dress to school, when asked “but what if the teacher were to say its okay?”, children will allow that the rule is no longer binding. But concerning a moral transgression, such as punching another student, children will tend to maintain that it is wrong regardless of what the teacher says on the matter.” (2005: 136).

As this kind of objective bindingness is characteristic of moral judgement, this speaks to the claim that children have a natural tendency to make such judgements.

Joyce therefore concludes that there is considerable evidence in favour of the hypothesis that our capacity to make moral judgements is innate rather than learned. Given that there are plausible evolutionary processes through which this distinctively moral capacity could have been encouraged, it is plausible to conclude that this capacity is the result of natural selection.

Next, Joyce argues from this hypothesis about the origins of our capacity for moral thought to the position that our moral beliefs are unjustified. An important part of this move is argument by analogy. The analogy is as follows: suppose that there are belief pills that “dispose you to form beliefs involving a particular concept” (2005: 181)¹¹. Specifically, say these beliefs pills cause you to form beliefs about Napoleon (though the exact content of these beliefs is not specified) such that, without these pills, one would have no beliefs about Napoleon at all.

Now imagine that we find out, beyond any reasonable doubt, that a couple of years ago we were secretly given one of these pills. We only believe that Napoleon was short, that he was a military leader, or that he even *existed* because of this belief pill. Surely this discovery would undermine all our beliefs about Napoleon. This discovery does not count as evidence in favour of our Napoleon beliefs being *false*, but it does undermine whatever justification we initially had for these beliefs. And, until we can find some independent evidence for or against our beliefs about Napoleon, we should become agnostic about all things Napoleon.

The forces of natural selection are, according to Joyce, just like the influence of the belief pill (2005: 181). We find out that we only have beliefs involving *moral* concepts because of the process of natural selection. Just as the belief pill genealogy undermined our Napoleon beliefs, so should the evolutionary genealogy undermine our moral beliefs. Until we find some kind of evidence that is unsullied by this genealogy, we should withhold all beliefs involving moral concepts. We should become agnostic about our previously held first-order moral beliefs, but also more generally about “whether there exists *anything* that is morally right and wrong” (2005: 181-2).

One thing to flag is that, until now, Joyce has argued that our capacity for moral thought and judgements can be given an evolutionary explanation. But now the focus is on moral *beliefs*. We do not need to get into the relationship between judgement and belief¹². But it is worth noting that, by giving an explanation of our capacity for moral judgements, Joyce also takes himself to have provided an explanation of our moral

¹¹ Joyce actually provides multiple analogies, but this is the one that is supposed to most accurately reflect our position with respect to moral thought.

¹² See Campbell (2007) for a discussion of whether or not we should view moral judgements as moral beliefs.

beliefs. This might be because beliefs are included in the category of judgements, or because the ability for moral beliefs is dependent on the ability of moral judgements. Either way, we can now just note the shift in focus from the explanation of our ability for moral judgements to the explanation of our moral beliefs.

Let us grant that, by giving an evolutionary explanation of our capacity for moral judgements, Joyce has thereby given an explanation of our moral beliefs. We can now get on to assessing Joyce's argument. This argument only succeeds if the analogy with belief-pills is a fair one. And what reason do we have for thinking that the analogy is fair? In both cases we have given a genealogy of our capacity for a certain kind of beliefs, but that cannot be the reason that we feel, intuitively, that our Napoleon beliefs are unjustified in the belief pill example. Clearly, not *all* genealogies of our beliefs undermine those beliefs. We could accept a totally pedestrian genealogy of any one of our beliefs, but surely accepting such a genealogy does not imply that we lose justification for all such beliefs. So there is clearly something particular about the belief-pill genealogy that gives it its undermining force, and until we have been given reason to think that this particular property is also present in the case of natural selection, we are under no obligation to view the analogy as a fair. We can view the evolutionary genealogy of our moral beliefs to be more properly analogous to one of the many possible non-threatening genealogies, thus retaining the possibility of justification for those beliefs.

In defending the fairness of the analogy, Joyce notes that "in the case of the belief pills the story has been carefully stipulated such that forming a belief as the result of taking a pill is entirely independent of whether or not the state of affairs necessary to render the belief true obtains in the world" (2005: 182). This suggests that in order for the analogy to be fair, and in order for our moral beliefs to be similarly undermined as the hypothetical Napoleon beliefs, the genealogical influences responsible for our moral beliefs (i.e. natural selection) have to be totally "independent of" moral truths. If this is true, then the analogy succeeds and our moral beliefs have been undermined. But what reason do we have for thinking that this is true?

Joyce argues that the existence of moral facts seems to be totally independent of the evolutionary benefit of being disposed to form beliefs about such facts (2005: 182). At no point in his explanation of why evolution would have encouraged moral beliefs

did Joyce have to invoke or make reference to moral facts themselves. As such, the evolutionary explanation that Joyce provides for our moral beliefs does not include moral facts.

Because, like Ruse, Joyce recognises the influence of non-evolutionary environmental factors in shaping the content of our moral beliefs (2005: 140), it is unlikely that he takes himself to have provided a complete explanation of those beliefs. As we did with Ruse, then, we should read Joyce as claiming that the any plausible way of completing this evolutionary explanation of our moral beliefs will not include the moral facts. From Joyce's evolutionary explanation of our moral beliefs, Joyce therefore infers that moral facts do not explain our moral beliefs.

As Joyce notes, evolutionary explanations of our *other* kinds of beliefs do make reference to the truth of those beliefs. It may well be that our mathematical beliefs have been selected for by natural selection, but we can only make sense of it being evolutionarily beneficial for us to hold certain mathematical beliefs under the assumption that such beliefs are *true*¹³. Joyce writes:

“So does the fact that we have such a genealogical explanation of our simple mathematical beliefs serve to demonstrate that we are unjustified in holding these beliefs? Surely not, for we have no grasp of how this belief might have been selected for, how it might have enhanced reproductive success, independent of its truth. False mathematical beliefs just aren't going to be very useful. Suppose you are being chased by three lions, you observe two quit the chase, and you conclude that it is now safe to slow down. The truth of “ $1 + 1 = 2$ ” is a background assumption to any reasonable hypothesis of how this belief might have come to be innate.” (2005: 182)

The idea here is that creatures with systematically incorrect mathematical beliefs would have a far tougher time successfully navigating their environment than those with correct mathematical beliefs, and so we can expect there to have been significant evolutionary pressure towards having mathematical beliefs that aligned with the mathematical facts. As a result, in providing an evolutionary genealogy for such beliefs we have not argued that the forces that are responsible for these beliefs are independent of their truth, and this genealogy therefore does not have any undermining

¹³ Clarke-Doane argues that, actually, mathematical facts are no more involved in the evolutionary explanation of our mathematic beliefs than moral facts are involved in the evolutionary explanation of our moral beliefs (2012).

implications for our moral beliefs. Joyce points out that similar comments apply to our other non-moral faculties for belief that can be given an evolutionary explanation (2005: 183).

So Joyce has given an evolutionary genealogy of our moral beliefs that does not make reference to moral truths. If our moral beliefs are not explained by moral facts, then the belief-pill analogy is a fair one and our moral beliefs become unjustified just as our Napoleon beliefs became unjustified. But this does not, as of yet, imply that our moral beliefs actually have been undermined. Joyce writes:

“That the evolutionary genealogy of morals contrasts with other cases (such as arithmetical and scientific beliefs)- in that the former does not presuppose the truth of the beliefs- is an important observation. But it doesn’t suffice for establishing that in the former case we have a debunking genealogy, for the possibility remains that an identity or supervenience relation may hold between the items denoted in the genealogy and the moral properties represented in the beliefs content, in such a way that the genealogy renders the belief true after all.” (2005: 184)

This passage makes it clear that, thus far, Joyce has only established that justification for our moral beliefs is undermined *under the assumption* that moral facts are not just reducible to those natural facts that are invoked in his evolutionary genealogy. If they are so reducible, then one can accept Joyce’s epistemological claim about the undermining implications of a plausible, non-truth-invoking genealogy for our beliefs, but deny that such a genealogy has been given for our moral beliefs. The relevant moral facts *have* been invoked, via the invocation of the natural facts with which they are identical.

Joyce defines moral naturalism as the view that moral facts and properties exist, and that they are the sort of thing that they can be “comfortably integrated” into the naturalistic view of the world which science can investigate (2005: 145). Non-naturalism, as defined by Joyce, is the view that moral properties exist but are not “identical to, reducible to or supervenient upon any natural properties”¹⁴. One category of non-naturalism is moral supernaturalism, according to which moral properties depend on some supernatural phenomenon like God.

¹⁴ I will challenge this definition, particularly the supervenience cause, below.

Thus, in order to conclude that our belief in moral facts have been undermined, Joyce first has to argue against naturalism. Joyce therefore dedicates a significant portion of the chapter containing his debunking argument to showing that naturalism is untenable. Some commentators have therefore taken Joyce to be presenting an unconditional debunking argument, according to which our moral beliefs are undermined regardless of our meta-ethical commitments (Wielenberg, 2016: 505)¹⁵.

But importantly, the considerations that Joyce invokes to argue against naturalism are totally independent of any evolutionary considerations. They have nothing to do with the proposed genealogy of our moral beliefs that Joyce spends the majority of his book defending, and instead involve more familiar arguments about the alleged incapability of naturalism to account for the practical authority associated with moral facts (2005: 190-209).

Because these arguments are totally independent of any evolutionary considerations, I am hesitant to include them as being part of Joyce's evolutionary debunking argument. Though presented as such, there is nothing evolutionary about them. If there is a distinctive meta-ethical implication of the evolutionary origins of our moral beliefs, then, it seems only to apply under the assumption that meta-ethical naturalism is false.

Whether or not meta-ethical naturalism can accommodate the practical clout of moral facts (and whether or not this is required of a meta-ethical position) is surely one of the most significant issues in meta-ethics. But this thesis is concerned with the distinctive issues that arise from the potential evolutionary origins of our moral beliefs, and so I am not here concerned to appraise Joyce's treatment of these issues.

What do we get when we consider Joyce's debunking argument stripped of these non-evolutionary, independent arguments? We get the conditional claim that *if* we deny the possibility of reducing moral facts to natural facts, then our moral beliefs are unjustified because moral facts are not included in the empirically supported evolutionary explanation of our moral beliefs. Or, what seems to be equivalent for Joyce (2005: 219): assuming moral non-naturalism, our moral beliefs are undermined

¹⁵ The "unconditional" terminology is taken from Korman (2019a).

according to his proposed genealogy¹⁶. It therefore seems that the conclusion of Joyce's evolutionary debunking argument, as I think we should understand it, is conditional: if we accept moral non-naturalism, then our moral beliefs are not justified. Joyce then employs independent philosophical arguments to show that meta-ethical positions *other* than non-naturalism should be rejected, thus transforming the conditional to the unconditional. But this transformation is independent of the evolutionary considerations, which furnish only the conditional conclusion.

Towards the end of his argument, Joyce considers some major theories of justification and argues that, according to each of them, the evolutionary origins of our moral beliefs undermines the justification that we have for those beliefs (2005: 221-219). But I need not outline Joyce's discussion of each of these theories, because Joyce usefully summarises his position at the end of this discussion:

“I acknowledge that there are complexities and epistemological theories that have not been discussed. However, I contend that on no epistemological theory worth its salt should the justificatory status of a belief remain unaffected by the discovery of an empirically supported theory that provides a complete explanation of why we have that belief while nowhere presupposing its truth. We should reject or modify any theory that would render us epistemic slaves to the baby bearing capacity of our ancestors.” (2005: 219)

We therefore need not consider the various ways in which specific theories of justification might deal with the relevant explanatory claim, because Joyce's position is clear. The explanatory claim undermines justification for our moral beliefs. And so, regardless of what particular theory of justification one accepts, if it is plausible then it will have to be consistent with our moral beliefs being undermined by and empirically supported genealogy that does not involve their truth.

Another similarity with Ruse's argument is that some commentators take Joyce to be primarily making a modal claim (Shaffer-Landau, 2012: 17, Bogardus 2016: 646, Klenk 2019a: 87, Clarke-Doane and Baras: 2020:169)). That is, they read Joyce as claiming that if the moral facts had been different, our moral beliefs would have been the same, and this is what undermines our moral beliefs under the assumption of non-

¹⁶ Earlier on, Joyce rules out non-cognitivism on grounds independent of evolutionary debunking arguments (2005: 51-7).

naturalism. This interpretation is supported by comments where Joyce emphasises the modal instability of our moral beliefs given his evolutionary genealogy of our moral capacity. For example, when discussing the belief-pill analogy, Joyce notes “Were it not for a certain social ancestry affecting our biology, the argument goes, we wouldn’t have concepts like *obligation*, *virtue*, *property*, *desert* and *fairness* at all” (2005: 181). But as I noted in the previous section, such a modal claim can in fact be employed as evidence for the explanatory claim. And given that the various passages I have just quoted in this section, I see no reason to interpret Joyce’s argument as primarily relying on the modal claim over and above the explanatory one.

The reading that I accept for Joyce’s evolutionary debunking argument is therefore as follows. There is good empirical evidence for the hypothesis that our capacity for moral beliefs can be given an evolutionary explanation. What’s more (unlike possible evolutionary explanations of our *other* capacities for belief) this evolutionary explanation implies that our moral beliefs are not explained by the moral facts, under the assumption of non-naturalism. What’s more, according to Joyce, accepting that our beliefs are not explained by the facts that make them true undermines our beliefs, and we should therefore remain agnostic about our moral beliefs under the assumption of non-naturalism.

In the next section, I provide an account of the final pre-existing evolutionary debunking argument that I will be concerned with in this chapter.

1.4: Street:

The target of Sharon Street’s influential Darwinian Dilemma are realist theories of value. Street characterises such positions as follows:

“The defining claim of realism about value, as I will be understanding it, is that there are at least some evaluative facts or truths that hold independently of all our evaluative attitudes. Evaluative facts or truths I understand as facts or truths of the form that X is a normative reason to Y, that one should or ought to X, that X is good, valuable, or worthwhile, that X is morally right or wrong, and so on. (2006: 110).

The basic idea behind these realist theories of values is that our attitudes (conscious or unconscious, pre-theoretical or theoretical) towards whether something is good or bad, valuable or not valuable, or counts for or against some action, is independent of

whether those attitudes are correct. Facts about value are not determined or constituted by our evaluative attitudes; some things can be morally good, for example, independently of what we judge to be morally good.

The first premise of Street's argument against realist theories of value is the claim that evolutionary forces have had a major impact on the content of human evaluative judgements. Street does not claim that the content of our judgements has been *entirely* determined by the forces of natural selection, and accepts the role that non-evolutionary factors have played in influencing those beliefs (2006: 113-4). Street's claim is just that the forces of natural selection have had a significant impact as well, "such that our system of evaluative judgements is thoroughly saturated with evolutionary influence" (2006: 114).

In support of this premise, Street notes the potential evolutionary impact of different evaluative judgements (2006: 114). A creature disposed to judge that the fact that something endangers survival is a reason to do it will surely be weeded out by the forces of natural selection, whereas one disposed to judge that they ought to do what will promote their survival will more successfully pass on their genes. We can therefore expect there to have been significant evolutionary pressure towards making those evaluative judgements that are conducive to reproductive success.

Street invokes actual patterns of evaluative attitudes in support of this contention: there is widespread acceptance of evaluative attitudes that would be evolutionarily beneficial even across cultural divides. The widespread, cross-cultural acceptance that "We have greater obligations to help our own children than we do to help complete strangers" or "The fact that someone has treated one well is a reason to treat that person well in return" is exactly what we would expect if evolution had influenced the content of our evaluative judgements (2006: 115).

Like Ruse, Street also supports this claim by citing the similarities between these evaluative tendencies and those apparently held by our closest non-human biological relatives (2006: 117). Chimpanzees, for example, appear to view the fact that they have been helped by another chimpanzee as counting in favour of helping that chimpanzee in future. This reflects our own evaluative tendency towards reciprocity, and further supports the hypothesis that our own evaluative tendencies have been influenced by the forces of natural selection.

More specifically, Street claims that evolution directly encouraged certain “proto” evaluative judgements: motivational tendencies “to experience something as “called for” or “demanded” in itself, or to experience one thing as “calling for” or “counting in favor of” something else” (2006: 119). These proto judgements in turn had a large effect on our “fully-fledged” reflective, linguistic evaluative judgements (2006: 120), such that had our proto judgements been very different, so would our fully-fledged evaluative judgements. Via the influence of our proto-judgements, the forces of natural selection have therefore had a large impact on the content of our evaluative judgements.

Thus, Street argues that the forces of natural selection have had a significant influence on the content of our evaluative judgements. With this in mind, Street presents a dilemma for the evaluative realist. The realist can either deny a relation between the forces of natural selection and mind-independent evaluative truths, or they can affirm a relation. Street argues that neither option is tenable for the realist, and we should therefore abandon evaluative realism.

According to the first option, the realist denies any relationship between the forces of natural selection and mind-independent evaluative truth. About this option, Street writes:

“The key point to see about this option is that if one takes it, then the forces of natural selection must be viewed as a purely distorting influence on our evaluative judgements, having pushed us in evaluative directions that have nothing whatsoever to do with the evaluative truth. On this view, allowing our evaluative judgements to be shaped by evolutionary influences is analogous to setting out for Bermuda and letting the course of your boat be determined by the wind and tides: just as the push of the wind and tides on your boat has nothing to do with where you want to go, so the historical push of natural selection on the content of our evaluative judgements has nothing to do with evaluative truths.” (121).

Like Joyce, Street argues by analogy. Just as we should not expect the Bermuda-independent winds to dock us in Bermuda, so we should not expect the morality-independent forces of evolution to provide us with evaluative judgements that correspond with evaluative truth. Given their independence, if our evaluative judgements are correct it would have to be by chance. And, given the number of logically possible moral systems, this chance alignment could only occur as a result

of a highly unlikely coincidence¹⁷. As we have no good reason for thinking that coincidence did occur, we should assume that it probably did not. This implies that our evaluative judgements are largely off-track, in the sense that they do not represent the genuine, attitude-independent evaluative truths and are therefore false.

Street does not make explicit the effect that this realisation is supposed to have on our moral beliefs under the assumption of moral realism. She does not say whether this realisation is supposed to stop such beliefs from being justified, or qualifying as knowledge, or both. But, given that Street takes this to be a *sceptical* result, it is reasonable to conclude that the realisation that our evaluative judgements are likely mostly off-track defeats whatever justification we had for those judgements, and also removes the possibility that such judgements might constitute knowledge. This is the implausible sceptical result that awaits the realist who embraces this first horn of the dilemma.

Street is also not explicit about *why* this sceptical hypothesis is supposed to be unacceptable for the evaluative realist. Given her characterisation of realism as a metaphysical hypothesis about the stance-independence of evaluative facts, it is unlikely that she takes this kind of scepticism to *entail* the falsity of evaluative realism. It is more likely, given her frequent reference to it as implausible, that she takes this form of scepticism to be unacceptable just because of its implausibility. *Of course* we know (and are justified in believing) that murder is wrong, and if any meta-ethical position implies otherwise then we should consider it a *reductio* of that position.

We can therefore summarise Street's argument thus far as follows: Our evaluative judgements have been shaped by the forces of natural selection. *If* there is no relationship between the forces of natural selection and evaluative truth, then this implies that our evaluative judgements could only be true as the result of an unlikely coincidence, and they are probably off-track. This realisation defeats our justification for those evaluative judgements, and/or prohibits their candidacy as knowledge. Due to its implausibility the realist cannot accept this sceptical result.

This is Street's argument that the realist cannot embrace the first horn of the dilemma by denying a relationship between the forces of natural selection and evaluative truth.

¹⁷ Some commentators have objected to the claim that there are as many logically possible moral systems as Street claims. For example, see Foot (1958) and Shaffer Landau (2014).

As with previous debunking arguments, it is worth noting that in the process of making this argument Street makes both an explanatory claim (about the absence of any relationship between our the forces of natural selection and realistically construed evaluative truths) and a different claim: that our moral beliefs could only align with the moral facts as the result of an unlikely coincidence. Depending on our understanding of coincidence, this further claim can be taken to mean that there are a high number of close possible worlds in which our evaluative attitudes and the evaluative facts do not align, and that there is no reason to think that any of these are less likely than the actual world in which, the realist supposes, they do in fact align. As such, this further claim about evaluative judgements only being true as a result of an unlikely coincidence can be seen as a modal claim about how things could have been, and how lucky we must be if we find ourselves in one of the few possible worlds in which there is enough alignment between our attitudes and the facts to allow for evaluative knowledge¹⁸.

The structure of Street's argument implies that the explanatory claim is used to argue for the modal claim, and it is from this modal claim that the sceptical implications are supposed to follow directly. Street argues that if our evaluative attitudes are disconnected from the evaluative facts, then they might only be true as the result of an unlikely coincidence, and that we ought to conclude that they are probably off-track. This implies the explanatory claim supports the modal claim, and the modal claim ensures evaluative scepticism. However, Street (just like debunker's previously considered) is not completely explicit about this, and so it may still be possible that Street is primarily making the explanatory claim and using all subsequent modal claims to justify this explanatory claim. Either way, Street concludes that the absence of a connection between the forces of natural selection and realistically construed moral facts somehow leads to moral scepticism under the assumption of realism, and that this is unacceptable for the realist.

We therefore come to the second horn of the dilemma. The realist cannot deny a relationship between the forces of natural selection and evaluative truth, so it appears they are forced to affirm such a relationship. One way of affirming this relation is to suggest that natural selection tracked mind-independent evaluative truths and

¹⁸ This is how Clarke Doane seems to interpret Street (2015). See David Faraci (2019) for an alternative understanding of coincidence.

encouraged creatures to have evaluative judgements that reflected those truths. Street dubs this position the “tracking account” (2006: 126). As Street notes, the tracking account is a scientific hypothesis about why evolution encouraged proto-judgements with the specific content that it did. It is therefore subject, just like any other scientific hypothesis, to evaluation according to its parsimony, illumination and clarity.

The problem is that there is an alternative explanation that wins out in all of these categories: namely, that certain proto-judgements were encouraged because they formed a link between a particular circumstance and behaviour that was adaptive in that circumstance (i.e, the “adaptive link account) (2006: 127-135). As well as being more parsimonious and more illuminating, the adaptive link account lacks the obscurity of the hypothesis that natural selection tracked evaluative truths. This latter hypothesis is obscure because it is unclear *why* the fact that a certain proto-judgement corresponds to a mind-independent evaluative fact would make it more evolutionarily beneficial to have that evaluative attitude than an evaluative attitude that lacks any corresponding mind-independent fact (2006: 131). It seems as though, given the debunker’s account of why evolution encouraged certain evaluative attitudes over others, there is no reason to think that a certain evaluative attitude’s being true makes it more likely to be so encouraged. It would have been evolutionarily beneficial to experience caring for our offspring as “called-for” even if there was no corresponding mind-independent moral truth such that we ought to care for our offspring. We should therefore reject the tracking account in favour of the adaptive link account, according to which there is no relationship between the forces of natural selection and mind-independent evaluative truths.

Street concludes that neither horn is tenable for the evaluative realist. They can neither deny nor affirm a relationship between the forces of natural selection and evaluative truth. The only way out, according to Street, is to accept a meta-ethical position according to which evaluative truths are *not* independent of our evaluative attitudes. This position furnishes an answer to the Darwinian dilemma: there *is* a relationship between the forces of natural selection and evaluative truths. Namely, the forces of natural selection shape our evaluative attitudes, and these evaluative attitudes determine evaluative truths. So the forces of natural selection determine, and are thereby related to, evaluative truths. Because of this relationship, the forces of natural selection are not distorting, and our evaluative judgements are no longer likely to be

mostly off-track (2006: 152-5). In order to avoid evaluative scepticism, then, we should therefore endorse some form of anti-realism over realism.

Street extends this argument against evaluative realism to naturalist theories of value (or, more specifically, to those naturalistic theories of value that endorse the stance-independence of evaluative facts that Street takes to be characteristic of realism (2006: 137)). The value naturalist holds that all evaluative facts are identical to natural facts. There must also be facts about *which* natural facts are identical to which normative facts. Street terms these “natural-normative identities” (2006: 137). Given the stance-independence of these kinds of naturalism, these facts about natural normative identities are independent of our evaluative attitudes; the value of pleasure, for example, is independent of our attitudes towards it.

How do we establish the nature of these natural-normative identities? That is, how do we establish which natural properties are identical with which normative properties? The only option appears to be to rely on our substantive evaluative judgements and to use these to justify our beliefs in natural-normative identities (2006: 140). But then Street asks: “what is the relation between evolutionary influences on our evaluative judgements, on the one hand, and the independent truths about natural-normative identities posited by the realist, on the other?” (2006: 140). If we deny any such relation, then scepticism follows because the evolutionary influence on our evaluative judgements is bound to be a distorting influence to the extent that we use those judgements to justify our beliefs in natural-normative identities. And if the realist affirms that there is such a relation, then they are forced to claim that the forces of natural selection tracked facts about natural-normative identities. But, according to Street, the notion that the forces of natural selection encouraged us to track natural-normative identities is (at least) as implausible as the view that the forces of natural selection encouraged us to track non-naturalistically construed moral facts (2006: 141). This tracking hypothesis should therefore be rejected in favour of an analogous adaptive link-account, according to which there is no relationship between the forces of natural selection and natural-normative identities.

Street therefore concludes that, as long the realist endorses stance-independence, one must opt for one of two unacceptable positions. One must take there to be a relationship between the evaluative facts (in the naturalist’s case, the natural-

normative identities) and the forces of natural selection, or one must deny such a relationship. Either option is untenable, and we should therefore endorse anti-realism as the most plausible way out of the dilemma (2006: 141).

1.5: Similarities and Differences:

In order to more easily see the relationship between these different debunking arguments, it is useful to briefly outline the major steps in each of them.

If my interpretation of Ruse is correct, his debunking argument can be summarised as follows: There is good empirical evidence that our moral capacity is an evolutionary development. If this account of our moral capacity is true then mind-independent, objective moral facts do not explain our moral capacity or our moral beliefs. Mind-independent moral facts are therefore redundant in the sense that they do not explain why we believe in such facts or feel that they exist, and when something is redundant in this sense we ought to believe it does not exist. Hence, objectivism is false, and we should accept some kind of subjectivism, according to which there is no more to morality than the feelings and dispositions of humans.

Joyce's debunking argument also begins with the claim that there is good empirical evidence for the hypothesis that our capacity for moral thought can be given an evolutionary explanation. What's more (unlike possible evolutionary explanations of our capacities for *other* kinds of thought) this evolutionary explanation nowhere involves moral facts under the assumption of non-naturalism. This implies that our moral beliefs are not explained by the facts that make them true. Joyce then argues that accepting an explanation for our belief that does not involve the truth of that belief undermines the justification for that belief. Justification for our moral beliefs has therefore been undermined and we should remain agnostic about any judgements that involve moral concepts under the assumption of non-naturalist realism.

Street argues that the content of our evaluative judgements has been significantly influenced by the forces of natural selection. If there is no relationship between evaluative truth and the forces of natural selection, then our evaluative judgements might only be true as the result of an unlikely coincidence. As we have no good reason for thinking a coincidence occurred we should assume that it probably did not, and

that (if there is no such relationship) then our evaluative attitudes are mostly off-track. This leads to evaluative scepticism, according to which we can have no moral knowledge and/or justification for our evaluative judgements, and (due to its implausibility) this position cannot be accepted by the moral realist. But the realist also cannot *affirm* a relationship between the forces of natural selection and evaluative truths, because this hypothesis is less plausible than the alternative adaptive-link account. The realist can therefore neither affirm or deny a relationship between the forces of natural selection and evaluative facts, and the only way out of the dilemma is to adopt some form of anti-realism that can accommodate such a relationship.

There are some things, then, that these debunking arguments have in common. The most obvious is an empirically supported hypothesis that our capacity to make moral judgements has been in some way influenced by the forces of natural selection. Debunkers also typically begin with an evolutionary explanation of our more basic moral attitudes, and then take this to imply that our more sophisticated attitudes can be similarly explained because these more sophisticated attitudes are outgrowth of (or at least heavily influenced by) the more basic ones. Another common feature is the emphasis that, given the most plausible evolutionary explanation for our judgements and/or beliefs, moral facts play no role in explaining why we have moral judgements and/or beliefs. This absence of a connection between our beliefs and the facts that make them true that is supposed to commit the realist to some sort of sceptical result.

A final feature that all these debunking arguments have in common is that they are not completely explicit about when exactly our moral beliefs get undermined (under the assumption of moral realism). Specifically, they make both an explanatory and a modal claim, and it is not made explicit which entails the sceptical result. Though I have argued that Ruse is best read as primarily arguing for an explanatory claim, I have also noted that at various points he implies the modal claim is crucial to his argument, and that this has led various commentators to (mistakenly, in my view) interpret him as such. Street moves from the absence of a relationship between the evaluative attitudes and the facts to the claim that they could only be true as a result of an unlikely coincidence, to the further claim that our evaluative attitudes are probably false. The relationship between these premises is not made explicit, and it could be that these latter comments are supplied in support of the more crucial claim

about the absence of a relationship between our evaluative attitudes and the evaluative facts (though I do not think this is how Street should be read).

Joyce is perhaps the clearest, stating that any plausible epistemic theory will have to imply that our beliefs are undermined when there is an empirically supported hypothesis about the origins of those belief that does not involve the facts that make them true. But it is still possible to interpret these comments as stating that this explanatory claim removes justification *because* it invariably leads to some modal claim, and this is exactly how multiple commentators have interpreted Joyce.

One potential reason for this ambiguity is that the explanatory and modal claim are so closely related that they typically run together, implying no need to specify which one does the defeating. I also think that these debunkers took the combined intuitive pull of their explanatory and modal comments to be sufficiently compelling to forgo the need to specify exactly at what point in these comments the realist (or non-naturalist realist) becomes defeated. By the end of their argument, it should be clear that realistically construed moral beliefs have been undermined, and there is no need to specify the exact point as which the undermining occurred, especially if there is no reason to think that anything of significance depends on this issue.

So much for their similarities. The major differences between arguments, as far as I can tell, are as follows. Firstly, the meta-ethical targets of these debunking arguments are different. Ruse and Street take evolutionary debunking arguments to show that moral realism of any kind is untenable (though they do so in different ways). Joyce, by contrast, takes his argument to be restricted to non-naturalist realism, and uses independent arguments in his attempt to rule out naturalism.

Secondly, Ruse and Joyce claim that natural selection is responsible for our capacity to make moral judgements, whereas Street's argument focuses on the influence of natural selection on the *content* our evaluative judgements. Furthermore, Ruse and Joyce imply that our belief in the existence of moral facts is among the beliefs subject to this evolutionary explanation, whereas Street uses evolutionary considerations to explain our first-order moral beliefs.

I think that these differences are related. The claim that our very capacity to form moral judgements can be given an evolutionary explanation supports the claim that our belief in moral facts can be given a non-truth involving explanation. Presumably, without capacity to form moral beliefs we would never have believed in moral facts to begin with. By contrast, if the empirical data just supports the hypothesis that the influence of natural selection has been largely on the content of our moral judgements (but that we would have made moral judgements of some kind independently of the forces of evolution) then this better supports the claim that our particular substantive moral judgements can be given an evolutionary explanation while our belief in moral facts cannot.

A further difference is on what this disconnect between facts and attitudes implies for the relevant mental states. Ruse's debunking argument is best viewed as claiming that, because of this disconnect, and we should believe that objectively construed moral facts do not exist. According to Joyce, this disconnect implies that we should become agnostic about all-things moral. For Street, the implication seems to be that evaluative judgements lose justification and/or cannot qualify as knowledge.

These debunking arguments also differ in what they imply about how we should cope with the supposed blow to their intended meta-ethical target. Ruse and Street both imply that we should accept a meta-ethical theory according to which moral truth is a function of our attitudes, because such a theory is apparently immune to the kind of evolutionary debunking that sinks realism. Joyce ultimately rejects these anti-realist positions as well, and ends his book by endorsing moral scepticism according to which should be agnostic about all moral claims (though his reasoning for doing so is strictly speaking independent of any evolutionary considerations).

A final, more subtle (but highly significant) way in which these debunking arguments differ is in how exactly they take the realist's beliefs to have become defeated (though as I have just mentioned, none of the three debunkers just considered are explicit about this). Ruse and Joyce appear to take the explanatory claim to undermine the moral beliefs "directly", whereas Street appears to think that the explanatory claim defeats

“indirectly”, by virtue of implying a further claim about our evaluative attitudes being, at best, only coincidentally reliable¹⁹.

These differences make it difficult get one common debunking argument on the table. Still, in the following section, I try to articulate an argument that captures the significant common features of such arguments. In places it is deliberately ambiguous, so as to be compatible with various differences between these arguments that I want to, at least for now, avoid assessing. In other places I have had to opt for the formulation that best sets up the discussion of the rest of the thesis. I outline my reasoning below.

1.6: A Common EDA:

EDA:

P1). We should accept an evolutionary explanation of our moral beliefs.

P2). If this evolutionary explanation of our moral beliefs is true then our moral beliefs are not explained by the moral facts under the assumption of non-naturalist realism²⁰.

P3). If we accept that moral facts do not explain our moral beliefs, then this undermines our moral beliefs under the assumption of non-naturalist realism.

C). Our moral beliefs are undermined under the assumption of non-naturalist realism.

1.7: Ambiguities:

This argument is ambiguous in various ways. First a foremost, premise 3 is ambiguous, and is consistent with both explanatory and modal readings of debunking arguments. Explanatory debunking arguments stipulate that (P3) is true because the explanatory concession undermine *directly*, whereas modal debunking arguments stipulate that (P3) is true *indirectly*, because of some further modal implication that follows from this explanatory claim. In either case, the accepted absence of an explanatory connection between our moral beliefs and the moral facts undermines those beliefs, as

¹⁹ This direct/indirect terminology is taken from Korman and Locke (2020).

²⁰ I will give an account of what I mean by non-naturalist realism below.

(P3) states. My interpretation of the debunking argument is therefore deliberately consistent with both explanatory and modal debunking arguments. This is because my own opinion on this issue can only emerge as the result of the substantial arguments that I will provide in later chapters.

Secondly, the “moral beliefs” to which each of the premises refer can be taken strictly as our first-order moral beliefs, or they can also be taken include meta-ethical beliefs about the very existence of moral facts. As I have noted, Ruse and Joyce argue that our very capacity for making moral judgements can be given an evolutionary explanation, and therefore claim that our belief in the existence of moral facts can be fully explained independently of the existence of those facts. By contrast, Street’s empirical considerations focus on the influence on the content of our moral beliefs, and therefore imply we should interpret (1) and (2) as being restricted to first-order moral beliefs; the distorting influence of evolution targets the content of our moral beliefs, rather than the fact that we have moral beliefs at all.

The reason I am ambiguous between these readings is that the difference is largely irrelevant to the purpose of this thesis. I am not concerned with assessing the empirical evidence to see whether it better supports the hypothesis that our first-order moral beliefs can be given an evolutionary explanation or whether our very belief in the existence of moral facts can be given such an explanation as well. Instead, I am concerned with if and how accepting that our beliefs are not explained by the facts that make them true can undermine those beliefs. This is what I take to be at issue between myself and the minimalist replies to evolutionary debunking arguments to which I am hoping to respond. And I take my comments on this issue to apply equally regardless of the content of the belief and, hence, regardless of whether we take the beliefs in question to be purely first-order moral beliefs or to include meta-ethical ones about the existence of moral facts²¹. I can therefore bracket the issue going forward, and invite the reader to interpret “moral beliefs” in whichever way they think is best supported by the empirical evidence. This will have no effect on the conclusion that

²¹ I defend my claim that moral beliefs have no special status, such that they might be immune to defeat via in virtue of explanatory considerations, at the beginning of the chapter 3.

our moral beliefs, so interpreted, are or are not undermined given non-naturalist realism.

Now, the difference between these conclusions is surely important; it is important whether our belief in moral facts is undermined, as Joyce claims, or whether our first-order moral beliefs in the rightness and wrongness of specific actions is undermined, as Street argues. But from the perspective of whether or not non-naturalism is tenable, it is unlikely that this distinction is going to make much of a difference. Either conclusion is surely going to be vigorously denied by the non-naturalist realist, who will have to show the debunker's argument fails according to either interpretation in order to argue that non-naturalism is a genuinely tenable position.

In what follows, I will generally focus on our substantive moral beliefs when considering what follows from a non-truth-involving explanation of those beliefs. The reason is that, perhaps due to the influence of Street's Darwinian Dilemma, this is what minimalist responses to debunking arguments typically focus on²². As I am concerned with engaging with minimalist responses, I take my lead from them (though, as I have just emphasised, I do not take anything of significance, for the purposes of this thesis, to follow from my doing so).

1.8: Commitments:

EDA makes reference to our beliefs being "undermined". By this, I mean that our previously justified moral beliefs have become unjustified²³. In what follows, I will also describe our moral beliefs as being "defeated", which I take to mean the same thing²⁴. Defeaters have traditionally been split into two kinds: rebutting and undercutting (John Pollock and Joseph Cruz, 1999: 196-7). Roughly put, rebutting defeaters remove justification for our belief in P by giving us reason to believe not-P.

²² David Enoch (2010) and Knut Skarsuane (230) both offer minimalist responses that are framed specifically in response to Street's Darwinian Dilemma.

²³ As I will outline in greater detail in chapter two, debunking arguments grant that our moral beliefs can be justified independently of the evolutionary considerations they take to debunk those beliefs.

²⁴ In this context, the term "defeat" was popularised by John Pollock (see, for example, 1974 & 1984), and has often been used in reference to the support provided to our belief by a particular reason, rather than the justification of a belief as a whole. I am using it in the latter respect, in line with Bergmann (2005).

We have a rebutting defeater for our belief that murder is always wrong if we encounter a scenario in which murder is not wrong, because this constitutes a reason for thinking that murder is not always wrong.

By contrast, undercutting defeaters remove justification for our belief in P by giving us reason to believe that our reasons or evidence in favour of P do not actually support P. If we come to learn that we only ever believed murder is always wrong because we were hypnotised to do so, then this defeats our belief by giving us reason to doubt that our apparent evidence for that belief actually supports that belief. This is true even though the fact that we were hypnotised to believe that murder is always wrong could not be employed as evidence in favour of the claim that murder is not always wrong (in a sense, what we were hypnotised to believe has nothing to do with the wrongness of murder)²⁵.

Evolutionary debunking arguments are best construed as attempting to undermine our beliefs by providing an undercutting defeater for those beliefs (under the assumption of moral realism) (Wielenberg, 2016; Matt Lutz, 2018). The claim that our moral beliefs can be given an evolutionary explanation cannot be used as evidence in favour of the claim that murder is not always wrong, but it does potentially give us reason to think that whatever evidence we had in favour of that belief does not actually support that belief. According to EDA, then, accepting that our moral beliefs are not explained by the facts that make them true constitutes an undercutting defeater for those beliefs (under the assumption of moral realism).

Some commentators distinguish between metaphysical and epistemological debunking arguments (Wielenberg, 2011; Lutz, 2018:1106). Metaphysical debunking

²⁵ More recently, some commentators have suggested that there is a third category of defeaters: “higher-order” defeaters (David Christensen, 2010). It is not always clear how best to differentiate higher-order defeaters from undercutting defeaters (for more on this distinction see Joshua Dipaolo (2018)) but the basic idea seems to be that higher-order evidence gives us reason to think that we have some cognitive impairment such that our belief was formed unreliably *without* giving us reason to doubt that our evidence supports that belief (see, for example, Christensen (2010: 194-5) and Ru Ye (2020b: 5438-9). Because the evolutionary debunking arguments defeat, if they succeed, by giving us reason to doubt that our usual evidence for moral belief actually qualifies as good evidence for those moral beliefs, I think they are best construed as attempting to provide an undercutting defeater for our beliefs. But nothing of any major significance depends on whether we view debunking arguments as attempting to provide undercutting or a higher-order defeaters for our realistically construed moral beliefs. What matters, for the purposes of this thesis, is whether evolutionary debunking arguments succeed in defeating at all.

arguments are those that conclude that moral facts do not exist, and epistemological arguments conclude that our moral beliefs are not justified or cannot qualify as knowledge. The conclusion of EDA is about the status of our moral beliefs, and I have therefore opted to interpret evolutionary debunking arguments epistemologically.

My reason for this is that I take the following line of argument to be the most plausible way of supporting a metaphysical reading of evolutionary debunking arguments: if we take the scope of “moral beliefs” to include our belief in moral facts, and we understand “undermining” to mean we should actually believe that such beliefs are *false*, then the conclusion of EDA would imply that we should believe that our belief in realistically construed moral facts is false. In other words, we should believe that realistically construed moral facts do not exist, and we therefore get a metaphysical conclusion of EDA.

But, as I have noted, I am understanding “undermining” in line with our beliefs having an undercutting, rather than a rebutting, defeater. Undercutting defeaters do not provide us with good evidence for believing our belief is false, and just remove the justification we initially had for our belief. As such, the evolutionary genealogy of our moral beliefs is not good evidence for the claim that our belief in realistically construed moral facts is false, and so we should not conclude, as a result of this genealogy, that realistically construed moral facts do not exist²⁶. My understanding of the way in which the evolutionary genealogy of our beliefs might undermine our beliefs therefore does not support this metaphysical interpretation of evolutionary debunking arguments.

Next, we can note that the target of my debunking argument is non-naturalist realism, rather than all forms of realism, as is the case with Ruse and Street’s debunking argument. What do I mean by non-naturalist (or: “robust”) realism?

This is the position that (i) moral judgements express propositions that can be true or false (ii) some moral propositions are true, (iii) the truth of these propositions is independent of our attitude towards them, (iv) moral truths are not identical with or reducible to natural truths, and (v): (i) – (iv) do not rely on a deflationary theory of

²⁶ For a similar thought in a different context, see Bergmann (2005: 424).

truth²⁷. Non-naturalist realism is an increasingly popular position, with some of its major advocates including Huemer (2005), Parfit (2011), and Enoch (2011)²⁸.

I think that non-naturalist realism is the meta-ethical position that is most obviously susceptible to evolutionary debunking arguments. In order to see why, and in order to further explain my understanding of non-naturalism, consider how theorists who deny each of the above conditions might respond to the evolutionary origins of our moral capacity²⁹.

The fundamental problem posed by an evolutionary genealogy for our moral judgements is a disconnect between moral judgements and moral facts, which leads to questions about how we can reasonably expect those beliefs to be justified. But a non-cognitivist who denied (i) need not worry about whether such judgements align with the facts that make them true, because, according to them, no such facts exist. Moral judgements are not the kind of thing that *can* align with facts, and their genealogy cannot imply they are “off-track” in any way because questions about them being off-track are misplaced to begin with³⁰.

A non-cognitivist would presumably take issue with EDA on the grounds that it presumes that we have moral *beliefs* at all. Under the assumption that we have no moral beliefs, my formulation of EDA would therefore fail in virtue of the fact that it harbours a commitment to the existence of moral beliefs as a background assumption. As such, a non-cognitivist should be unphased about the evolutionary origins of our moral attitudes in general and about the specific formulation of EDA.

An error theorist who accepted (i) but denied (ii) would be similarly be unphased by the evolutionary origins of our moral beliefs, and would readily accept the conclusion that all such beliefs had a full explanation independent of moral facts; this seems to be

²⁷ Note that this does not necessarily imply that non-naturalists cannot accept a deflationary theory of truth. They may do, they just cannot take to truth of (i)-(iv) to *depend* on that deflationary theory of truth such that, if the deflationary theory of truth were false, one of more of these clauses would also be false.

²⁸ See also Fitzpatrick (2008) and Wielenberg (2014).

²⁹ I refer to moral capacity rather than beliefs because some theorists who deny (i) might deny that we have moral beliefs.

³⁰ Philip Kitcher actually argues that the evolutionary origins of our moral capacity supports non-cognitivism, because they imply that the function of moral thought is to motivate us to action (1998; 2005). I will not consider this claim here, though I think Joyce (2006: 175-77) correctly diagnoses why Kitcher’s argument fails.

Joyce's own position. If we swap out all mentions of non-naturalism with "error-theory" in EDA, the argument would still succeed (if, that is, it succeeds against the non-naturalist) but the conclusion would be readily accepted by the error theorist. They could accept an evolutionary explanation of our moral beliefs, and that such an explanation does not include moral facts, and that when beliefs are so disconnected from the facts that make them true they are undermined. This would imply that our moral beliefs are undermined under the assumption of error theory, but the error theorist is already committed to the claim that our moral beliefs are all incorrect, and so we can expect them to be perfectly okay with this conclusion.

A theorist who accepts (i) and (ii) but denies (iii) takes moral truth to exist but to be in some way that is dependent on human moral attitudes. Constructivist and ideal-observer theories spring to mind. Such positions have a ready answer to the question of how our moral beliefs might be related to moral facts, given the evolutionary origins of our beliefs. That is, our moral beliefs are influenced by our moral attitudes. It therefore does not matter *how* we came to have the moral attitudes that we have. So long as we have access to them, there is no reason to suppose that we cannot allow them to influence our beliefs, thus connecting our moral beliefs with the moral facts. Swap out non-naturalism with constructivism in EDA, and the inference from (P1) to (P2) fails. That we are forced to accept such a position given the evolutionary origins of our moral beliefs is exactly the point of Ruse and Street's debunking arguments.

Clause (iv) is intended to distinguish non-naturalism from naturalism. One reason I have chosen this specific branch of realism, rather than having the argument appeal to naturalist realism as well, is that I think that Joyce is correct in his insistence that it is possible for the naturalist to claim that moral facts are identical with those natural facts cited in the evolutionary explanation of our moral beliefs. And, by doing so, the naturalist can reject the claim that the debunker has defended a non-truth involving explanation of those moral beliefs. If so, the naturalist can reject (P2).

As I have noted above, Street extends her argument to apply to naturalist realism as well as non-naturalism. Street's additional argument is interesting, and may well show that naturalism is just as subject to debunking arguments as non-naturalism. But if it does succeed, it does so in part by extending the same methods that are used in the debunking of non-naturalism. That is, it presumes that if our beliefs in NN-identities

are disconnected from the moral facts, this disconnect undermines those beliefs. Street then goes on to argue that the realist cannot affirm that our NN-beliefs are connected to the moral facts, that their NN-beliefs are thereby undermined, and that the undermining of such beliefs sinks naturalism.

As such, Street's extension to her argument only succeeds if a disconnect between our beliefs that the facts that make them true does in fact undermine those beliefs, and her initial argument against non-naturalism succeeds as well. Once this has been established, one can attempt to make the further argument there exists such a disconnect with our NN-identity beliefs as well, and that the undermining of such beliefs cannot be countenanced by the naturalist. But if there is something wrong with this argumentative strategy as applied to non-naturalism then the extended application of it to naturalism will fail as well. Because of this, the assessment of the debunking argument applied to non-naturalism is prior to an assessment of Street's extended argument against naturalism because, if we are yet to establish that this initial argument succeeds, we are in no position to establish that the extension succeeds either. So even if one takes debunking arguments to successfully rule out naturalism in the way that Street argues, one will first have to establish that the argument framed against non-naturalism works. This is why I take it to be worth-while considering the argument framed against non-naturalism, and leaving it up to the reader to establish whether it can then be extended to debunk naturalism also.

The final clause (v) is intended to distinguish non-naturalism from quasi-realism. Quasi-realists are typically characterized as denying a realist metaphysics (moral facts do not actually exist in the metaphysically weighty sense that realists have in mind) whilst justifying talk of moral truth in moral discourse by utilising a deflationary notion of truth. According to a deflationary theory of truth, when we say that a statement "is true" we do nothing more than express the statement itself; the predicate of truth does not actually denote any further property that was not denoted by the initial statement. We can also accept deflationism about facts, according to which saying "P is a fact" is nothing more than simply saying "P".

Wielding a deflationary theory of truth, the quasi-realist can claim that "it is true that it is morally wrong to murder" just in case it is morally wrong to murder. Nothing of any metaphysical weight is added by calling the statement "true". As they already have

an account of what it means to make first-order moral judgements, talk of moral truth is granted automatically. Same goes for the statements about moral facts. Thus, the quasi-realist may be able to deny realist metaphysics whilst accepting clauses (i)-(iv) above. By instating clause (v), according to which the truth of (i)-(iv) need not depend on a deflationary theory of truth, I attempt to establish that the targeted meta-ethical position takes there to exist stance-independent, non-natural moral facts in the metaphysically weighty sense usually thought to be characteristic of non-naturalist realism in opposition with quasi-realism.

Now, with the resources of a deflationary notion of truth at hand, it is notoriously difficult to successfully distinguish realism from quasi-realism. The quasi-realist could presumably accept that mind-independent moral facts exist in a “metaphysically weighty” sense, so long they extend their deflationism to talk about metaphysical weight. This is the threat of “creeping minimalism” to which James Drier (2004) draws our attention, and is a threat to the very possibility of individuating between realism and quasi-realism. It may in fact be possible for the quasi-realist to accept clauses (i) to (iv) *and* (v), so long as (v) can somehow be interpreted in a deflationary way.

I hope I have said enough for the reader to get the gist of the difference between non-naturalism and quasi realism, even if everything that I have said about realism being “metaphysically weighty” can actually be accommodated by the persistent quasi-realist as well. To the reader who objects that, because of the threat of creeping minimalism just described, clause (v) fails to successfully distinguish between non-naturalism and quasi realism, I ask: is there any difference between realism and quasi-realism? If not, then my failure to distinguish between them is to be expected, and the arguments in the rest of the thesis, which concern the evolutionary origins of our moral beliefs under the assumption of non-naturalism, will apply straightforwardly quasi-realism, which is just another term for non-naturalism.

If there *is* a difference between realism and quasi-realism, then I invite the reader to substitute (v) with their’s own anti-quasi-realism clause that successfully individuates non-naturalism from quasi-realism. I hope they will forgive me shifting this burden onto them, especially if it spares them my own attempt to solve the problem of creeping minimalism. If the reader is sure that there is a difference between quasi realism but has no solid opinion on how exactly to capture this difference, they can by

all means substitute (v) with some sort of non-specific “whatever separates non-naturalism from quasi-realism” clause. This should be enough to establish that, in what follows, I am concerned with non-naturalism and not quasi-realism.

Why have I decided to have EDA target non-naturalism but not quasi-realism (assuming the two are genuinely distinct)? Because of the problem just discussed (of how exactly to distinguish quasi-realism from realism) it is very difficult to know whether the commitments of quasi-realism leave it susceptible to evolutionary debunking. A denial of a realist metaphysics seems to put realism into the camp of meta-ethical positions that need not worry about the potential relationship between moral beliefs and mind-independent moral facts. On the other hand, if quasi-realism countenances discourse involving mind-independent moral facts and moral beliefs, should it not also countenance discussion about the relationship between the two? (Street, 2011).

In order to avoid the difficulties surrounding this issue, I have focused my argument specifically on non-naturalism. If it is legitimate to ask the quasi-realist questions about the relationship between our moral beliefs and mind-independent moral facts, and if these questions will have to be answered in some way that reflects the realist’s answers to such questions, then my discussion of EDA as it applies to non-naturalism will be easily extendable to quasi-realism as well. This completes my justification for why I have chosen to target non-naturalism, as defined by clauses (i)-(v), in EDA.

Another notable feature of EDA is that it is framed in terms of what we accept or ought to accept, rather than what is objectively true. For example, (P1) says that we *ought to accept* an evolutionary explanation of our moral beliefs, rather than saying that an evolutionary explanation of our moral beliefs is *true*. (P3) says that *if we accept* the relevant explanatory disconnect then our moral beliefs have been undermined, and does not say that our moral beliefs have been undermined because there does in fact exist the relevant explanatory disconnect between our moral beliefs and the moral facts.

I have framed EDA in this way because, firstly, minimalist responses to evolutionary debunking arguments grant the evolutionary origins of our moral beliefs, and grant that our moral beliefs are not explained by the moral facts. They do not resist this

explanatory disconnect by taking issue with the debunker's evolutionary genealogy of our moral beliefs, or by arguing that such a genealogy does not imply the explanatory disconnect that the debunker thinks it does. Instead they argue that, *even if one accepts* the explanatory disconnect for which debunker's argue, our realistically construed moral beliefs are not thereby defeated. As such, they object to EDA as I have formulated it, according to which our moral beliefs are defeated if we accept that they are not explained by the facts that make them true.

Furthermore, I think that debunking arguments are particularly compelling when they are framed in this way. It seems difficult to argue that an explanatory disconnect between our moral beliefs and the moral facts might undermine those beliefs even if we (mistakenly) do not actually believe that such a disconnect exists, or even have no reason to believe in such a disconnect³¹. Someone might be epistemically blameless for holding some belief in P that is not in fact explained by P, but once they *accept* that their belief is not explained by P, the case that their belief is unjustified surely becomes much stronger. It is because minimalist responses object to this point that I want to discuss those responses, and so by framing EDA in this way I can focus on the issues that will be of most interest to me in the forthcoming chapters. I can avoid having to address issues that would come with a defence of the corresponding debunking argument that is not framed in terms of what we accept or have reason to accept.

Another decision I have made regarding EDA is that its conclusion is just that our moral beliefs are not justified given non-naturalist realism. I do not argue that we

³¹ It seems difficult make such an argument without presupposing externalism about epistemic justification, according to which the justification of our beliefs can be affected by factors to which we have no access. And the same issues that have been raised against externalism about justification generally can be raised against its application in evolutionary debunking arguments (Cohen and Stewart: 1983, Cohen 1984). For example, imagine two agents that are exactly alike mentally but are in different possible worlds. One of these agents is in a world where their moral beliefs are the result of evolution, and this implies their moral beliefs are explanatorily disconnected from the facts that make them true. The other is in a world their moral beliefs cannot be given an evolutionary explanation, and in which God has endeavoured to provide them with moral beliefs are explained by realistically construed moral facts. As such, there is no valid evolutionary debunking argument that could be made against this second agent. Given that these individuals have exactly the same mental states, it is implausible that their moral beliefs could have different degrees of justification. This implies that the mere existence of a valid debunking argument cannot defeat our realistically construed moral beliefs *unless* we have some access to the debunking argument or some reason for thinking that it might exist.

should thereby give up realism. I also do not profess to have any idea what meta-ethical position we should accept if we were to do so. The reason, once again, is because I take these extremely important and difficult questions to lie outside the scope of this thesis, and I content myself with the less ambitious question of whether or not non-naturalist realism is consistent with our having justified moral beliefs. And, of course, the way in which debunkers argue for these further conclusions is by showing that our moral beliefs cannot be justified given the debunker's genealogy. My establishing whether this claim is valid is therefore of primary importance to those commentators who take the proper conclusion of debunking arguments to be that we should reject realism in favour of some alternative meta-ethical position.

A final point about EDA that requires justification: I have framed my argument exclusively in terms of a disconnect between our moral *beliefs* and realistically construed moral facts. I have therefore not formulated this premise as proceeding from a claim about the origins of our more basic moral capacity, as with Ruse and Joyce. I also have departed from Street's evolutionary debunking argument in this respect, in which any mention of moral beliefs is absent, with the argument being framed entirely in terms of a disconnect between our moral *judgements* and the moral facts.

One reason for framing the argument exclusively in terms of moral beliefs is that it avoids the issue of whether, by explaining our more basic capacity for moral feelings we thereby explain our moral beliefs. If we could arrive at moral beliefs without relying on our evolutionarily influenced moral capacity, one could object to this part of the debunkers argument. This is an interesting strategy that some commentators have taken, but it is not the one that I will be concerned with in this thesis³². By framing the argument in terms of beliefs I effectively grant the debunker this move to moral beliefs order to see what might follow from it. This is in line with minimalist responses to debunking arguments, which also grant, for the sake of argument, that our moral *beliefs* can be explained without reference to the moral facts, and then investigate whether our moral beliefs are therefore undermined.

³² For the claim that we can somehow circumnavigate the evolutionarily influence moral capacity and arrive at moral beliefs in some other way, see Copp (2008), Shafer Landau (2012:19-20), and Cuneo & Shafer Landau (2014).

Furthermore, as I have just clarified, the target of EDA is moral non-naturalism, according to which (among other things) there are moral facts and moral propositions can be true or false. This position naturally lends itself to the view that, when we make a moral judgement, we thereby form a moral belief. The major line of resistance to viewing moral judgements as beliefs comes from non-cognitivism, according to which moral judgements are (or express) some kind of non-cognitive mental state. As EDA is concerned exclusively with the implication of the evolutionary origin of our moral attitudes under the assumption of non-naturalism, we should be able to use beliefs and judgements interchangeably. By framing the argument in terms of our moral beliefs, then, I make this presumption explicit.

1.9: Conclusion:

This chapter began with an account of three of the most influential debunking arguments in the literature. I then noted several similarities between these debunking arguments, namely: that our moral capacity and beliefs can be given an evolutionary explanation, that such an explanation implies an explanatory disconnect between our moral judgements and the moral facts, and that such a disconnect has sceptical implications for certain meta-ethical positions.

My own EDA attempts to accommodate these similarities. It is also formulated in such a way to abstract away from various issues that do not concern me, and to focus on the issues that I do wish to discuss going forward. These are issues to do with the supposed explanatory disconnect between our moral beliefs and realistically construed moral facts. Say there does exist such a disconnect, does it genuinely imply that our moral beliefs can be justified? If so, how? Does it do so directly, or is it supposed to have some further modal implication that does the undermining?

As we will see, these questions are of central importance to minimalist responses to debunking arguments. Minimalist responses are happy to grant that our moral beliefs are not explained by moral facts but argue that accepting such a disconnect does not undermine those beliefs. As such, they argue, the non-naturalist can accept the debunker's genealogy of their moral beliefs without being saddled with any sort of moral scepticism. It is to these responses that I now turn.

Chapter 2: Minimalist Replies to Debunking Arguments:

2.1: Introduction:

In chapter one I formulated an evolutionary debunking argument (EDA) that was designed to incorporate those features of debunking arguments that are most pertinent to minimalist responses to such arguments. One of these features is that our moral beliefs are not explained by the moral facts, and another is that accepting this undermines the justification of our beliefs under the assumption of non-naturalist realism (henceforth just “realism”). Minimalist responses to debunking arguments grant that our moral beliefs are not explained by the moral facts under the assumption of moral realism but then argue that this explanatory disconnect between our moral beliefs and the moral facts does not undermine our moral beliefs, and they rely on substantive moral claims when doing so.

The two kinds of minimalist response to debunking arguments that I will be outlining in this chapter are the “Modal Security” response and “third-factor” replies, and I begin this chapter by explaining each of these responses in turn. I then consider a natural response to these replies, according to which they fail because they beg the question against the debunker. I reject this diagnosis in favour of another, more promising suggestion: they fail because the substantive moral claims on which they rely have been defeated by evolutionary debunking arguments. I note that the success of this response depends on whether the explanatory disconnect that minimalists grant between moral beliefs and the moral facts can undermine those beliefs.

I end this chapter by neutralising the concern that, even if explanatory considerations *can* undermine directly, this possibility cannot vindicate evolutionary debunking arguments against moral realism. The worry is that, because realistically construed moral facts are causally impotent, they are incapable of explaining our moral beliefs for reasons that have nothing to do with the evolutionary origins of those beliefs. As such, if explanatory considerations can undermine directly, then our moral beliefs have always been unjustified given moral realism, and evolutionary considerations are irrelevant in establishing this fact. I argue that the causal impotency of moral facts

does not entail that our beliefs are not explained by those facts, and so this worry is misplaced.

This chapter sets the stage for the following two chapters, in which I argue that accepting an explanatory disconnect between our beliefs and the relevant facts can undermine those beliefs independently of what they imply about the modal security of our beliefs³³. I will go on to argue, however, that the specific explanatory disconnect implied by debunking arguments does not in itself defeat our realistically construed moral beliefs, and further arguments are required in order to establish an explanatory disconnect which, if accepted, would in fact defeat those beliefs. I go on to provide and defend these arguments in chapters 5 and 6.

2.2: The Modal Security Response:

Minimalist responses to debunking arguments have elsewhere been treated as one group (Korman and Locke, 2020), but I think that they can be usefully broken down into two sub-categories. I will label the first minimalist response to evolutionary debunking arguments the “Modal Security” response.

The modal security response involves granting the explanatory premise of debunking arguments but then arguing that neither the moral facts, nor our moral beliefs, could easily have been different. Under the assumption that our moral beliefs are in fact true, this implies that our moral beliefs could not have easily failed to align with the moral facts even if we can give a complete explanation of our moral beliefs that does not involve those facts. We can therefore still explain how our moral beliefs might satisfy any modal condition on justified belief even if we accept that they are not explained by the moral facts (as the debunker claims). This implies that our realistically construed moral beliefs can remain justified, because those beliefs can only be defeated by their failure of some modal condition on justified belief, and accepting the relevant explanatory disconnect does not imply that they fail this condition. The primary advocate for this kind of response is Clarke-Doane (2015), and so I will first focus on the argument as he presents it, though towards the end of this section I will also indicate where other commentators have utilised this argument.

³³ I explain the notion of modal security below.

In his influential “Justification in Mathematics and Morality” (2015), Clarke-Doane interprets debunking arguments as moving from the explanatory claim that moral facts are not cited in the best explanation of our moral beliefs to the claim that there is no possible explanation of the reliability of our moral beliefs. If there is no possible explanation of the reliability of our beliefs, then those beliefs could only be accidentally or coincidentally correct, and this undermines our beliefs (2015: 84). Thus, Clarke-Doane takes the debunker to be arguing that, because of the explanatory premise, there is no possible explanation of the reliability of our moral beliefs under the assumption of moral realism and, under the assumption of realism, our moral beliefs are thereby undermined.

According to Clarke-Doane, our moral beliefs are “sensitive” if, had the moral facts been different, our moral beliefs would have been correspondingly different (2015: 87-8). Our beliefs are “safe” if we could not easily have had false beliefs (2015: 93). Showing that our moral beliefs are insensitive or unsafe are two ways in which one might argue that there is no possible explanation of the reliability of those beliefs. Clarke-Doane considers whether the explanatory claim implies either that our moral beliefs are insensitive or unsafe in turn.

First, Clarke-Doane argues that the explanatory claim does not imply that our moral beliefs are insensitive. In order to make this point, Clarke-Doane notes that “explanatorily basic” moral beliefs are, if true, necessarily true. Explanatorily basic beliefs are “our conditional beliefs which purport to state the conditions under which a moral property is instantiated” (2015: 88). The belief that an action is wrong if it needlessly promotes suffering might be an example of such a belief. Clarke-Doane then points out that, when beliefs in metaphysically necessary facts are true, those beliefs are trivially sensitive because there are no possible worlds in which the relevant facts are different, and hence no possible worlds in which those facts are different and our beliefs about them remain unchanged. As a result, then, our explanatorily basic moral beliefs are trivially sensitive.

Importantly, this argument does not make any reference to an explanatory connection between our moral beliefs and the moral facts. Those moral beliefs could have *any* genealogy, and as long as they are about necessary truths, they are trivially sensitive.

The debunker's evolutionary genealogy of our moral beliefs therefore fails to establish that our explanatorily basic moral beliefs are insensitive.

Furthermore, our non-basic moral beliefs seem more obviously sensitive than our basic beliefs. To use Clarke-Doane's example:

“Had Bush's invasion of Iraq not been wrong, it would have been different in non-moral respects, and our moral beliefs would have varied correspondingly (since, even if the explanatorily basic moral truths would not be metaphysically necessary, the closest worlds in which the antecedent is true are presumably worlds in which those truths are the same).” (2015: 88).

In the closest possible world in which Bush's invasion of Iraq is not wrong, this is not because, in that world, the basic moral principles have changed. Rather, it is because the natural facts have changed such that those same moral principles no longer imply that the war in Iraq is wrong. For example, it is still wrong to promote unnecessary suffering in that world, but the war in Iraq no longer causes unnecessary suffering, and so the war in Iraq is no longer wrong.

But, as our belief that the war was wrong is contingent on these very natural facts (such as the war's promotion of unnecessary suffering), if the war had not caused such suffering then we would not believe it to be wrong. This implies that if our non-basic moral beliefs were not true then, because they would no longer have the natural properties that are actually responsible for our belief that they are wrong, we would not believe that they were wrong. Thus, even given the debunker's genealogy of our moral beliefs, we have no reason to think that either our explanatorily basic or non-basic beliefs are insensitive.

Next, Clarke-Doane argues that the debunker's genealogy of our moral beliefs does not imply that those beliefs are unsafe, either (2015: 93). To make this point, Clarke-Doane uses the debunker's own genealogy of our moral beliefs against them. Street implies that the contents of our moral beliefs were shaped by natural selection, and that natural selection would not have allowed for organisms with massively different evaluative judgements. Given the further assumption that our explanatorily basic moral beliefs are in fact true, we get the claim that we could not easily have had moral beliefs that failed to align with the moral facts. In other words, we could not easily have had false moral beliefs, and so our moral beliefs are safe. Once again, this point

does not depend on whether or not our moral beliefs are explained by the moral facts, and so the explanatory disconnect implied by debunking arguments does not imply that our moral beliefs are unsafe.

Thus far, then, Clarke-Doane has argued that the explanatory claim gives us no reason to believe that our moral beliefs are either insensitive or unsafe. He ends his critique of debunking arguments by arguing that information cannot undermine our beliefs *unless* it gives us reason to think that our moral beliefs are insensitive or unsafe. According to his “Modal security”:

“Modal security: “Information, E, cannot undermine our D-beliefs without giving us some reason to believe that our D-beliefs are not both safe and sensitive.” (2015: 97)

Our D-beliefs are our beliefs about a particular area³⁴. So, according to Modal Security, information cannot defeat our beliefs about a particular area without giving us some reason to believe that our beliefs about that area are not both safe and sensitive. If, for example, information about our beliefs in a given area gives us reason to think there is an explanatory disconnect between our beliefs and the facts that make them true, but does not give us reason to think those beliefs are either unsafe or insensitive, then that information cannot undermine those beliefs.

According to Clarke-Doane, if information does not threaten the sensitivity or safety of our beliefs, then it does not give us reason to think that our beliefs were not “(all but) bound to be true” (2015: 97). And Clarke-Doane finds it difficult to see why we should be rationally required to give up beliefs that, for all we know, were *bound* to align with the facts. Thus, if information does not give us reason to doubt the sensitivity or safety of our beliefs, it cannot undermine those beliefs. Because, as Clarke-Doane has just argued, the debunker has not given us reason to doubt the sensitivity or safety of our moral beliefs, the debunker has therefore failed to undermine our moral beliefs.

First, then, Clarke-Doane takes the debunker to be arguing that, because our moral beliefs can be explained without reference to moral facts, we cannot explain the reliability of those beliefs and those beliefs are therefore undermined. Clarke-Doane

³⁴ I assume the “D” stands for “domain”, but Clarke Doane uses the “area” terminology when explicating the notion of D-beliefs (2015: 84). Nothing depends on this going forward.

objects by arguing that the debunker has failed to show either that our beliefs are insensitive or unsafe, which are two ways in which the debunker might have shown that the realist cannot explain the reliability of our moral beliefs. Clarke-Doane then concludes by arguing that information can *only* undermine our beliefs by giving us reason to believe they are either insensitive or unsafe, which implies that the debunker's genealogy of our moral beliefs cannot undermine those beliefs.

Olson (2019) also argues in a manner strikingly similar to Clarke-Doane. First Olson notes that fundamental moral truths are generally taken to be necessary. Next Olson points out that our beliefs in whatever plausible candidates there are for fundamental moral truths are widely believed and that, according to the debunker, such beliefs are evolutionarily advantageous. The upshot of these points, according to Olson, is as follows:

“This suggests that non-naturalists have a ready explanation of why at least some of our true moral beliefs are non-coincidentally, and hence reliably, true. First, it is no coincidence that the fundamental moral truths obtain, since they are necessary truths. Neither is it a coincidence that some non-fundamental moral truths obtain, since they follow from conjunctions of fundamental moral truths and non-moral truths about worlds like ours. Second, it is no coincidence that belief in some of the (fundamental and non-fundamental) moral truths is widespread, since believing these truths (or something close enough to them) tends to enhance evolutionary fitness. Admittedly, natural selection could have favoured other kinds of moral belief, but it is not a coincidence that it favoured the kinds of moral belief it did favour. It is thus not a coincidence that some of our moral beliefs are true; that is, it is not a coincidence that we have beliefs whose content are true moral propositions.” (2019: 292-3).

First, then, Olson notes that fundamental moral truths are necessary, and these truths could not easily have been different. Some non-fundamental moral truths are guaranteed by the fundamental moral truths in combination with natural facts about our world, and so they too could not easily have been different. As a result, then, fundamental and (some) non-fundamental moral truths could not easily have been different.

Next, Olson notes that the widespread acceptance of certain fundamental and non-fundamental moral beliefs can be explained by the fact that those beliefs promote reproductive success. It is therefore no coincidence that we have the moral beliefs that

we do; natural selection encouraged those beliefs specifically because they were the beliefs that promoted reproductive success. If such beliefs are true, then this implies that natural selection could not easily have encouraged false beliefs, and our moral beliefs are non-coincidentally accurate.

Thus, Olson argues, evolution may have encouraged us to have moral beliefs that were close enough to moral truth that, what with our capacity for normative theorizing, we can reasonably think that we might arrive at moral truth. And there is no sense in which this might have been a coincidence. Olson therefore concludes that “evolutionary debunking arguments do not present an insurmountable challenge for moral non-naturalism” (2019: 295)³⁵.

It is important not to misinterpret Olson when he says that it is no coincidence that evolution selected for true moral beliefs. A possible interpretation of these comments is that evolution selected for true moral beliefs *because they are true*. Another is that it is no coincidence that evolution selected for moral beliefs with a particular content and, what do you know, beliefs with these contents are true. Only the latter interpretation is supported by his earlier comments; he has done nothing to show that, *contra* the debunker, evolution selects for beliefs with a particular content because those beliefs are true. Instead, he has just shown that certain moral truths could not easily have been different, and that evolution selects for moral beliefs that align with these truths *because these belief promote reproductive success*. This does not support the claim that evolution selects for moral beliefs because they are true, and we should avoid attributing this strategy credibility because we take it to have established this separate, unsupported claim.

David Enoch also makes comments that seem to support something like the argument presented by Clarke-Doane and Olson. Having just defended a different kind of minimalist response that I will outline below, Enoch considers in what sense it would be miraculous if the aim of evolution were in fact morally good:

“For what would have to be the case for this "miracle" not to occur? The evolutionary "aim" would have had to not be of any value. And how could that be? Fundamental

³⁵ Though Olson does think that debunking arguments can be usefully levied by the nihilist and sceptic when responding to Moorean objections (2019: 295-8).

normative truths are presumably necessary in a fairly strong sense, or at the very least so we are entitled to assume in the context of critically evaluating the epistemological challenge to Robust Realism. So the main way in which the evolutionary "aim" (which is actually of value) could have failed to be of value is if evolution had a very different "aim". But it's not clear what to make of this suggestion: For surely, it's not contingent that evolution has something to do, for instance, with survival and reproductive success rather than their opposites." (2010: 433)

Once again we see the same general strategy at work: the fundamental normative truths could not have been different and (because the aim of evolution also could not easily have been different) neither could our moral beliefs. As such, there is no sense in which our moral beliefs might easily have been false.

Though only Clarke-Doane argues explicitly that information cannot undermine our beliefs unless it gives us some reason to think that our belief is not sensitive or not safe, this seems to be a background assumption held by any advocate of the strategy under consideration. These commentators hope to establish that the debunker's non-truth-involving explanation of our moral beliefs does not imply that those moral beliefs are defeated because this explanation does not imply that either the moral facts or our moral beliefs could easily have been different. But showing that the moral facts and our moral beliefs could not have easily failed to align only ensures that they remain undefeated if information must threaten the modal stability of those beliefs in order to undermine them.

If Clarke-Doane's Modal Security principle is false, then even if these commentators have successfully shown that the debunker has given us no reason to doubt that our realistically construed moral beliefs are sensitive or safe, they have not thereby shown that the debunker has failed to undermine our moral beliefs. This is because, if Clarke-Doane's Modal Security principle is false, then the debunker may have undermined our realistically construed moral beliefs in some way *other* than by showing that they are unsafe or insensitive. They might have defeated them, for example, by showing that they are not explained by the facts that make them true. Thus, if Modal Security is false, and information need not show that our moral beliefs are unsafe or insensitive in order to defeat those moral beliefs, then showing that the debunker has failed to establish that our realistically construed moral beliefs are unsafe or insensitive does thereby succeed in vindicating our realistically construed moral beliefs.

In the previous chapter, I noted that evolutionary debunking arguments have typically been ambiguous about the point at which the realists' beliefs are supposed to be defeated. They make both an explanatory claim and a modal claim and, when these have both supposedly been established, the realist's moral beliefs are meant to be defeated. But it is often left implicit which claim is supposed to be really crucial in the defeating of our moral beliefs under the assumption of moral realism. We can now see that, whether explicitly (as with Clarke-Doane) or implicitly, the modal security response takes debunking arguments to be primarily making the modal claim because, according to their view, the explanatory claim by itself cannot undermine our beliefs. Having interpreted debunking arguments thus, they then argue that the modal claim does not follow even given the explanatory claim and debunking arguments, so interpreted, fail.

The modal security response is partially characterized by the fact that it does not attempt to establish any particular relationship between our moral beliefs and the moral facts. Rather, it takes each of them separately and establishes that neither one could easily have been different. Given the further assumption that they do in fact align, the minimalist therefore concludes that they have shown that our moral beliefs are securely, non-coincidentally true, *regardless* of whatever explanatory relationship might hold between these two facts.

In contrast, the kind of minimalist response to debunking arguments that I will consider in the next section attempts to show that, even if our moral beliefs are not explained by moral facts, there might still be some alternative relationship that holds between the two and in virtue of which our moral beliefs can escape defeat. Unlike the modal security response, this is an attempt to find an alternative relationship between moral facts and moral beliefs once the debunker has ruled out a more straightforward explanatory connection, rather than an attempt to do away entirely with the need to establish any such relationship.

2.3: Third-Factor Replies:

As with the modal security response, third-factor replies begin by granting that our moral beliefs are not explained by moral facts. They then argue, however, that there is some third-factor that explains them both. Because our moral beliefs are explained

by some fact that *also* explains the truth of those beliefs, the realist can grant that our moral beliefs are not explained by the moral facts without thereby accepting that such a genealogy has undermining implications for those beliefs.

An influential proponent of third-factor replies is Enoch, who first grants that moral facts do not explain our moral beliefs (2010: 430). One way to explain a correlation between A facts and B facts is to say A facts are responsible for B facts or *vice versa*. But for reasons related to those outlined by Street, “it is exactly these two kinds of explanation that are unavailable to the robust realist” (2010: 429). Having accepted that the moral facts do not explain our beliefs, Enoch looks to explain the correlation between the two by showing that there is a third-factor that explains them both. In the process of doing so, Enoch assumes the moral premise that “survival [...] is at least somewhat good” (2010: 430).

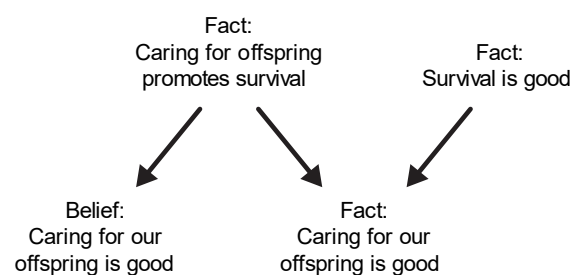
If, as the debunker claims, our moral beliefs are explained by natural selection, then we have the moral beliefs we do because those beliefs encouraged behaviour that promoted survival. Because we are inclined to behave in ways that we think are morally called for, and natural selection encouraged us to behave in ways that promote survival, natural selection will have encouraged us to believe that certain actions are morally good because those actions promote survival. For example, our belief that we ought to care for our offspring is partially explained by the fact that caring for our offspring promotes survival (2010: 431).

What’s more, under the assumption

that survival *is* morally good, the fact that caring for our offspring promotes survival will *also* explain the fact that it is morally good to care for our offspring. It is morally good to care for

our offspring in virtue of the fact that caring for our offspring promotes survival, which is a morally worthwhile end (2010: 431). We can represent this explanatory structure in figure 1³⁶. Here, the fact that caring for our offspring promotes survival grounds the

Figure 1:



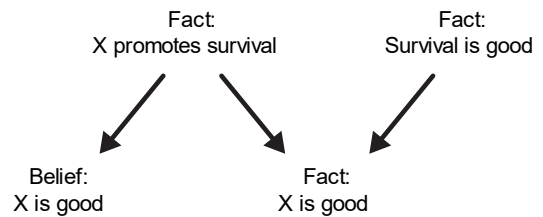
³⁶ Note that, according to this explanatory structure, the third-factor that explains *both* our belief that looking after our offspring is good and the fact that it is good is the natural fact that looking after our offspring promotes survival. The moral fact that promoting survival is good explains the moral fact that looking after our offspring is good but does not explain our belief.

fact that caring for our offspring is good. The goodness of caring for our offspring depends on, and is therefore explained by, the fact that doing so promotes survival³⁷.

This implies that we believe that caring for our offspring is morally good because it promotes survival, and that caring for our offspring *is* in fact morally good because it promotes survival. Thus, there is a third factor (that caring for our offspring promotes survival) that explains both our belief in the goodness of caring for our offspring and the moral fact that we ought to care for our offspring. It is therefore no coincidence that our belief that we ought to care for our offspring is correct; we only have the belief because of the very fact that ensures its truth. If caring for our offspring were not morally good, then it would not promote survival and we would not have been encouraged by natural selection feel motivated towards it. Given that caring for our offspring promotes survival, then, it is likely both that we would believe doing so is morally good and that doing so *is* morally good; we could not have easily been wrong about this.

The same story can be told for all those moral beliefs that are explained by natural selection; if we believe some action is good then that will be because that action promotes survival. What's more, in virtue

Figure 2:



of the fact that the action promotes survival, that action actually will be (pro tanto) morally good. This is represented in figure 2. Though our moral beliefs are not explained by the facts that make them true, they are explained by some fact which itself explains the fact that makes them true. This explains how many of our moral beliefs are non-coincidentally accurate even if the debunker is correct in claiming that there is a full evolutionary explanation of why we have them that does not involve moral facts. This is Enoch's explanation of the reliability of our moral beliefs.

In line with some commentators, I have interpreted Enoch as claiming that the third-factor that explains both our belief that *x* is good and the fact that *x* is good is the natural fact that *X* promotes survival³⁸. But some read Enoch as claiming that the

³⁷ A more explicit account of explanation in terms of metaphysical determination relations will be given in the next chapter.

³⁸ For example, see Tersmann (2017: 765).

relevant third-factor is actually the assumed moral fact that survival is good. This could be motivated by the following passage from Enoch:

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

Perhaps it is in response to comments like these that Michelle Dyke claims that, according to Enoch, the third-factor responsible for both our normative judgements and the normative facts is “that survival is in fact (by and large) good” (2020: 2117-8). Michael Klenk also attributes to Enoch the view that the third-factor that ultimately causes our moral judgements and grounds the moral facts is the moral fact that survival is at least somewhat good (2020: 430). Even Wielenberg, when offering his own third-factor reply explicitly in the style of Enoch’s, interprets Enoch as stating that the third-factor responsible for both normative judgements and our normative beliefs is “the goodness of survival or reproductive success” (2010: 450).

But there are problems with this interpretation. As I mentioned above, Enoch grants the debunker’s point that our moral beliefs are not explained by the moral facts. His strategy is to show how our moral beliefs are reliably correlated with those facts even if neither one explains the other. But if the third-factor that is supposed to explain both our moral beliefs and the moral facts is the fact that survival is morally good, this clearly involves a moral fact (that survival is morally good) explaining our moral beliefs. As Enoch is accepting that moral facts do *not* explain our moral beliefs, and is instead looking for an *alternative* relationship that might hold between them, he cannot be proposing that our moral beliefs are explained by the fact that survival is morally good³⁹.

If this *is* what Enoch is proposing, his response seems to beg the question against the evolutionary debunker, who takes herself to have provided good empirical arguments for the claim that our moral beliefs can be given an evolutionary explanation which implies they are not explained by the moral facts. Enoch’s response to evolutionary

³⁹ This point is made by Tersman (2017).

debunking arguments would involve denying this claim without offering any argument for doing so and without engaging with the empirical evidence that the debunker uses to motivate this position. Interpreting Enoch as making this claim therefore dramatically weakens his third-factor reply to debunking arguments.

Furthermore, Enoch notes that the third-factor structure is not causal in *both* directions, but allows that, on the side of the third-factor explaining our moral beliefs, the relationship may be causal (431). But how could the causally inert moral fact that survival is good *cause* any of our moral beliefs?⁴⁰ If this is indeed what Enoch is suggesting then, once again, his position becomes much less plausible and we have good reason to reject it.

But, as I have already outlined, I do not think that this is Enoch's position. He does not claim that the goodness of survival actually explains our moral beliefs, but just that the goodness of survival explains the *correlation* between our moral beliefs and the moral facts. Explaining the correlation between two facts is distinct from explaining both of those facts.

To see this, consider again the explanatory structure represented in figure 2. Clearly, there is nothing incoherent in supposing that this explanatory structure holds while *denying* that the goodness of survival explains any of our beliefs that a particular action is good. That the fact that X promotes survival explains our belief that X is good and the fact that X is good does not imply or give us reason to believe the further claim that the moral fact that survival is good *also* explains this belief and the fact that makes it true. These are clearly distinct positions, and accepting one does not, in itself, give us reason for accepting the other.

However, if the explanatory structure of figure 2 holds, then this explanatory structure as a whole explains why our belief that X is good correlates with the fact that X is good; they correlate because there is a third-factor that is responsible for them both. But this explanatory structure is only possible in virtue of the moral fact that survival is good. If survival were not good, then the fact that X promotes survival would not explain that fact that X is good, and so the third-factor structure outlined in figure 2

⁴⁰ Here I do not mean to imply that causally inert are in principle incapable of *explaining* our beliefs. As I will outline below, I think that this position is mistaken. I just mean to point out that the causal impotency of moral facts seems to render their explaining moral beliefs by *causing* them impossible.

could not hold. Because it is in virtue of this third-factor structure that there is a correlation between our moral beliefs and the moral facts, and this third-factor structure is only possible because survival is good, the correlation between our moral beliefs and the moral facts is explained by fact that survival is good. This is why Enoch comments that the *correlation* between our moral beliefs and the moral facts is explained by the fact that survival is good.

Above I noted that the explanatory structure outlined in figure 2 does not imply that the goodness of survival *also* explains both our belief that the action is good and the fact that the action is good. But I have now shown that this explanatory structure does imply that the goodness of survival explains the correlation between our belief that the action is good and the fact that the action is good. Together, these points imply that some fact can explain the correlation between two facts without thereby explaining both of those facts. For the reasons I have just outlined, someone who accepts the explanatory structure outlined in figure 2 has good reason to accept that the correlation between be belief that X is good and the fact that X is good is explained by the goodness of survival, but they do not have good reason to accept that the goodness of survival explains both their belief that X is good and the fact that X is good.

This implies that something can explain the *correlation* between A and B without explaining both A and B. I think that commentators who read Enoch as claiming that the goodness of survival explains both our moral beliefs and the facts they make them true miss this point. They assume that, because the goodness of survival is supposed to explain the correlation between the two, it must actually explain our moral beliefs and the moral facts. But, as I have just argued, this inference is invalid. I therefore think that Enoch is best read as claiming that the third-factor that explains both our belief that some act is good and the fact that the act is good is the natural fact that the act in question promotes survival, rather than the moral fact that survival is good.

Other third-factor repliers use the same general strategy as Enoch but differ in the substantive moral claim on which they base their argument. Wielenberg claims that having cognitive faculties entails having rights (2010: 444-445). It is hard to see how evolution could have encouraged the belief that we have rights without imbuing us with cognitive capacities that ensure our possession of those very rights. In other words, the fact that we have cognitive faculties explains both our belief that we have

rights and the fact that we do have rights, even if (as the debunker claims) neither one explains the other. So, once again, we only have our belief in virtue of the very fact that ensures its truth. This implies that our belief is formed by a process sufficiently reliable for knowledge, and this provides us with a base from which we can reason our way to broader forms of moral knowledge.

Third-factor replier Knut Skarsuane first assumes that pleasure is usually good and pain usually bad. (2011: 232). He then grants that evolution has had a large influence on our moral beliefs. Next, he argues that part of the way in which evolution will have encouraged us to value some things and disvalue others is by shaping us to feel that some things are pleasurable and some things are painful (233-4). But, under the assumption that pleasure is usually good and pain usually bad, the fact that a certain act is pleasurable will usually make it good and the fact that is painful will usually make it bad. Thus, the fact that a given act is pleasurable explains both our belief that the act is valuable and the fact the act is valuable. This keeps our moral beliefs reliable enough that we can get at moral truth through rational reflection, even if (as the debunker claims) the moral facts do not explain our moral beliefs.

Kevin Brosnan assumes, in the process of his reply to debunking arguments, that promoting well-being is morally good (2011: 60). He then considers why we might have a belief like co-operation is good. Part of why we have this belief (according to the debunker) is because having this belief encourages helping behaviours in our group. But helping behaviours presumably increase the well-being in our group. And, given our assumption that well-being is morally good, the fact that co-operation increases well-being will explain why co-operation is morally good. So, the fact that [believing that co-operation is good promotes helpful behaviour] explains both why we believe that co-operation is good and also the fact that co-operation is good. This implies that the probability of believing that co-operation is good is not independent of co-operation actually being good, and that our belief therefore “tracks” the truth (2011: 53). The sceptical implications that were supposed to follow from a tracking failure can therefore be avoided.

Third-factor replies share some characteristics with the modal security response. Firstly, they both accept, for the sake of argument, the debunker’s explanatory claim that moral facts nowhere explain our moral beliefs. They also crucially rely on

substantive moral claims in order to show that this explanatory claim does not defeat our moral beliefs under the assumption of moral realism. It is because of these shared features that they are both characterised as “minimalist”.

But what differentiates them from the modal security response is that the third-factor replier endeavours to use such moral claims to show how there might be an *alternative* explanatory relationship that holds between our moral beliefs and the moral facts, even if one does not explain the other, whereas the modal security response ignores any such connection and focuses exclusively on how easily either one could have been different, irrespective of the relationship between the two. As such, third-factor replies do not presuppose that accepting an explanatory disconnect can only defeat our moral beliefs to the extent that it presupposes those beliefs are modally secure. But it does presuppose that accepting our moral beliefs are not explained by the moral facts cannot, in and of itself, undermine our justification for those beliefs, so long as there remains the possibility that those beliefs might be explained by some fact that also explains the truth of those beliefs. As will become clear in chapters to come, I take this difference to be crucial in diagnosing what goes wrong with each of these responses to debunking arguments.

2.4: Reliance on Moral Beliefs:

It is natural to wonder whether the minimalist’s reliance on substantive moral claims is legitimate. Is it fair for Clarke-Doane to assume that our beliefs about explanatorily basic moral facts actually align with those explanatory basic moral facts? Is it fair for Enoch to assume that survival is morally good? For those sympathetic to debunking arguments, it can seem illegitimate for commentators to assume that some of their moral beliefs are correct when responding to such arguments.

One potential diagnosis of why this is illegitimate is that, by relying on substantive moral claims without argument, minimalists beg the question. Minimalists attempt to show that their moral beliefs are justified, but they do so by crucially relying on moral claims. Does this not assume their moral beliefs are justified in the process of arguing that their moral beliefs are justified?

This line of argument seems to imply that, when trying to vindicate a particular class of beliefs (i.e. establish that those beliefs are justified), one can only legitimately

employ beliefs that are independent of that class. Otherwise, one can be charged with begging the question and the supposed vindication fails. Of course, in the current context, there is the further implication that the absence of such an independent vindication undermines those beliefs. If we denied this further condition, the realist could accept that the minimalist's attempted vindication of realistically construed moral beliefs fails because it begs the question, but such an admission would not have the sceptical implications that the debunker is trying to establish because no such independent verification would be required.

The problem is that this requirement for independent verification leads to total scepticism. For one thing, it is hard to imagine what such independent vindication would look like for any of our faculties for belief, including perception (Street, 2008: 216; Bedke, 2014: 107; Vavova, 2014; Clarke-Doane, 2015: 89). How might we argue that our perceptual beliefs are justified? Well, we might point out that it would have been evolutionarily beneficial to have correct perceptual beliefs, because having incorrect perceptual beliefs about hungry lions and cliff-edges would remove us quickly from the gene pool. But this argument presupposes the truth of those perceptual beliefs on the basis of which we believe in evolutionary theory. Without relying on any such perceptual beliefs, it seems that we do not have enough background knowledge to mount a defence of such beliefs (Sosa, 2002: 375). Similarly, it is hard to see how we could vindicate our logical beliefs without at any point presupposing their truth.

In fact, the independent verification requirement leads to a regress. Every faculty must be justified on the basis of a distinct faculty, which itself must be justified on the basis of a distinct faculty, and so on *ad infinitum*. Assuming we have a finite number of faculties, this entails that independent justification for any of our faculties for belief is impossible, and that none of our beliefs can ever be justified (Huemer, 2005: 108).

This can be taken as a *reductio* of the requirement that all justified beliefs are independently verified; surely, we do not accept total scepticism according to which *none* of our beliefs are justified. It isn't even clear whether this position is coherent; do we not rely on beliefs (for example, in the requirement of independent vindication for all justified beliefs) when coming to this conclusion? Is our belief in the failure of all our beliefs to live up to this requirement itself unjustified?

But even if we were prepared, somehow, to accept the total sceptical consequences of these two requirements, doing so would render evolutionary debunking arguments a specific application of a much more general sceptical problem. Our moral beliefs would be unjustified because we could provide the requisite independent justification for them. But the same goes for all of our non-moral beliefs as well. There is nothing uniquely problematic about our moral beliefs, and no special issues posed by their evolutionary origins⁴¹. It would be impossible to provide the requisite kind of justification for our moral beliefs regardless of whether those beliefs were explained evolutionarily, and so all evolutionary considerations drop out as otiose⁴². This is unacceptable to the debunker, who takes herself to be arguing that our moral beliefs face a particular problem in virtue of their evolutionary origins (Street, 2008: 216; Shaffer Landau, 2012: 18) (Vavova, 2014: 82-3).

All this implies that the debunker cannot object to minimalist responses on the grounds that they rely on moral beliefs in the process of vindicating our moral beliefs. If the minimalist is indeed begging the question by relying on their substantive moral beliefs when arguing that those beliefs can be justified, there can be nothing objectionable about their being question-begging in this sense⁴³.

Furthermore, evolutionary debunking arguments are supposed to assume that, prior to such arguments, realistically construed moral beliefs are *prima facie* justified (Shafer 2010: 475-476; Locke 2014: 230). If their moral beliefs were already unjustified prior to debunking arguments, then those arguments would become entirely superfluous. The debunker does not take themselves to be presenting an argument that the realist's moral beliefs are unjustified that only works under the assumption that those moral beliefs are unjustified. This would be of little interest to anyone. The debunker's actual ambitions are grander: they bracket all the typical objections to moral realism and assume that realistically construed moral beliefs can be justified for the sake of

⁴¹ Note that, even assuming that moral facts are causally impotent, there would be no special problem for our moral beliefs because it is impossible some faculty for belief can attain the requisite kind of independent verification *regardless* of whether the relevant facts are causally potent or not. If the absence of this independent verification is what is wrong with our moral beliefs, then this flaw applies equally to all faculties for belief, including those that concern causally potent entities.

⁴² And even if the evolutionary origins of our moral beliefs were troubling, they could not be justifiably believed in because of the total scepticism to which we now ascribe.

⁴³ For more on the issue of begging the question as it applies to minimalist responses to evolutionary debunking arguments, see David Copp (2019) and Korman and Locke (2020).

argument. They then present an argument that is supposed to remove such justification *even assuming* that those beliefs were justified prior to these considerations.

Thus, the debunker must treat the realist's moral beliefs as being *prima facie* justified prior to evolutionary debunking arguments. And there is nothing wrong, generally speaking, with utilising moral beliefs in the process of arguing that moral beliefs can be justified. So what could be wrong with the minimalist's reliance on substantive moral claims when responding to debunking arguments?

2.5: What Defeats Our Beliefs:

The suggestion that I want to pursue is that minimalists cannot rely on their moral beliefs in response to EDA's because EDA's defeat those moral beliefs (Locke, 2014; Lutz, 2018; Korman, 2019b; Korman and Locke, 2020). This explains why minimalists cannot rely on their moral beliefs even though those moral beliefs were *prima facie* justified prior to debunking arguments; this justification has been defeated by debunking arguments. This is also consistent with the claim that there is nothing wrong, in and of itself, with using the outputs of a faculty when establishing that the beliefs it produces can be justified; just so long as those outputs have not already been defeated. This suggestion therefore provides a potential justification for the claim that the minimalist's reliance on substantive moral claims is illegitimate, and one that is consistent with the points made in the previous section.

Korman and Locke have recently argued that success of this argument against minimalists depends on which claim in debunking arguments is supposed to defeat our realistically construed moral beliefs (2020). Recall that minimalists accept, for the sake of argument, the evolutionary genealogy of our moral beliefs presented by the debunker and the further claim that, according to this genealogy, our moral beliefs are not explained by the moral facts. They then rely on moral claims to argue that *even if* there exists this explanatory disconnect between our moral beliefs and the moral facts, those moral beliefs can still satisfy whatever condition is required for justified belief. Minimalists are therefore engaged in the project of granting that our moral beliefs are not explained by the moral facts and then disputing what follows from this point.

Now the suggestion is that these responses fail because their reliance on substantive moral beliefs is illegitimate, and they are illegitimate because their moral beliefs have

been defeated. But if the defeating force of debunking arguments is supposed to come from a further claim that is meant to be inferred from the relevant explanatory disconnect, then the minimalist is perfectly entitled to accept this explanatory disconnect and then rely on their moral beliefs to show why the further claim does not hold. This is because, until this further defeating claim has been established, their moral beliefs remain undefeated and can therefore be relied upon by the realist⁴⁴. Their use of substantive moral claims is therefore legitimate.

In order to see this point, imagine a supposed debunker of our perceptual beliefs. Our perceptual beliefs are debunked, they argue, because they are *insensitive*; we would have the very same perceptual beliefs even if they were false. Now imagine that we respond by appealing to the evolutionary benefit to having correct, sensitive perceptual beliefs. Then the debunker claims that this move fails because we cannot rely on our perceptual beliefs *because those beliefs have been defeated by their insensitivity*.

Clearly, there is something amiss about this response. The debunker is trying to establish that our perceptual beliefs are insensitive and thereby defeated. In the process of establishing this point, then, they cannot assume that our perceptual beliefs are insensitive and are thereby defeated. Until the insensitivity of our perceptual beliefs have been established, then, they remain undefeated and “on the table” when it comes to considering whether such beliefs are insensitive or not⁴⁵.

The same point goes for any condition that we might claim our beliefs fail and are therefore defeated. In the process of arguing that such beliefs fail this condition and are defeated, we cannot assume that such beliefs fail this condition and are thereby defeated. If our beliefs are genuinely only defeated because they fail this condition (and they are not defeated independently of failing this condition) we must therefore allow that such beliefs have not been defeated *unless* we have established that those beliefs fail the relevant condition.

This implies that, if the defeating force of debunking arguments comes from the explanatory disconnect itself, then the minimalist’s use of substantive moral claims is illegitimate. This is because minimalists grant this explanatory disconnect and then

⁴⁴ Copp (2019: 238 & 242) makes a point similar to this.

⁴⁵ And, for the reasons outlined in the previous section, we should not be squeamish about employing the beliefs produced by a given faculty in the process of arguing that such beliefs can be justified.

use substantive moral claims to argue that some additional further condition does not hold. But if the explanatory claim itself does the defeating, then minimalists cannot accept this point and then rely on substantive moral beliefs because their acceptance of this point defeats those beliefs. Having accepted this point, they are therefore not entitled to rely on such beliefs in the process of vindicating their moral faculty. And, even if they could rely on their substantive moral beliefs to show that our moral beliefs satisfy some further condition, this would not secure justification for those beliefs; this would not show that the explanatory claim is false, and it is the explanatory claim that does the undermining.

The claim that minimalist replies fail because they rely on defeated moral beliefs therefore depends on the claim that accepting the relevant explanatory disconnect between our beliefs and the facts that make them true can defeat those beliefs, regardless of whatever further implications might follow.

There are two ways in which one might deny that accepting the relevant explanatory disconnect between our moral beliefs and the moral facts might defeat our moral beliefs directly, and they correspond to each of the minimalist responses outlined above. Firstly, the explanatory disconnect might not defeat directly because it does not imply anything about the modal security of our belief, and information has to show that our beliefs are modally unstable in order to defeat that belief. This is clearly what the proponent of the modal security response would want to say.

Secondly, one might accept that an explanatory disconnect between our belief in P and P can defeat that belief regardless of what it implies for the modal security of our belief, but deny that the *specific* explanatory disconnect implied by debunking arguments defeats that belief. We might think, for example, that accepting our moral beliefs are not explained by the moral facts does not defeat those beliefs so long as we take there to be some third-factor that explains them both. This is the line of response that correlates with third-factor replies, and one can advocate for this response even if one thinks that explanatory considerations can undermine directly independently of what they imply for the modal security of that belief.

In the following two chapters, I consider both of these possibilities in turn. In chapter three, I will argue that accepting an explanatory disconnect between our moral beliefs and the moral facts can defeat those beliefs independently of what it implies for the

modal security of that belief, and this implies that the modal security response is fundamentally misguided. In chapter 4, I will side with the third-factor replier and argue that the specific explanatory disconnect implied by debunking arguments does not defeat our realistically construed moral beliefs because it leaves open the possibility that our moral beliefs might be explained by some third-factor that also explains the truth of those beliefs. However, in chapters five and six, I will argue that there are good independent reasons for thinking that our beliefs about basic moral facts cannot be explained by some fact that also explains those basic moral beliefs, and that when EDAs are combined with these independent considerations, they imply that minimalist responses ultimately fail.

But, before I get onto that, I first need to consider a dialectical point that threatens the success of this strategy before it even gets off the ground.

2.6: A Dialectical Point:

As we have seen, minimalists deny that accepting the relevant explanatory disconnect between our moral beliefs and the moral facts can defeat that belief directly. In the next chapter, I will argue that accepting the right kind of explanatory disconnect between our moral beliefs and the moral facts *can* defeat those beliefs directly, regardless of whether we have reason to doubt the modal security of those beliefs. Arguing for this point will involve an epistemological discussion that is independent of evolutionary debunking arguments. Generally speaking, I will argue, if we accept there is no explanatory relationship between our belief in P and P then that belief is defeated regardless of what it implies for the modal security of that belief. I will then apply these independent epistemological considerations to the debate between evolutionary debunking arguments and minimalist replies in an attempt to show that the modal security response to debunking arguments is fundamentally misguided.

But some commentators have argued that, even if it were true that accepting an explanatory disconnect between our belief in P and P could defeat that belief, we could not use this claim in defence of evolutionary debunking arguments against non-naturalism (Clarke Doane, 2015; Michael Klenk, 2019b). This is because non-naturalism *already* entails that our moral beliefs are not explained by the moral facts, independently of those evolutionary considerations presented by the debunker. So if

evolutionary debunking arguments do nothing more than establish that there exists such an explanatory disconnect, the evolutionary considerations that they employ become completely irrelevant.

If this line of argument is sound, then there may well be good reasons, independent of evolutionary debunking arguments and minimalist replies, for the claim that accepting an explanatory disconnect between our belief in P and P can defeat that belief directly. But we cannot apply these considerations to the debate between evolutionary debunking arguments and minimalist replies without rendering evolutionary debunking arguments superfluous. So, before I get onto the independent epistemological considerations of the next chapter, I first need to establish that they can, if legitimate, be applied to the debate between evolutionary debunking arguments and minimalist replies without rendering evolutionary debunking arguments unnecessary. I first outline this worry more fully, and then argue that, actually, the debunker need not be worried.

Clarke-Doane dubs the hypothesis that moral facts do not explain any of our observations (including our moral judgements) “Harman’s Objection” (2015: 82). He calls the position that our beliefs are undermined if the facts that make them true do not feature in the best explanation of our observations “Quinean empiricism” (2015: 82). Of the possibility that an awareness of Harman’s objection might defeat our moral beliefs in and of itself (regardless of whatever modal implications follow from this observation) Clarke-Doane writes:

“But how could Harman’s Objection undermine our moral beliefs? Again, if Quinean empiricism were true, then, rather than undermining our moral beliefs, Harman’s Objection would show that they were never justified to begin with. But even if this were what Joyce intended to show, such an argument would have no traction with Joyce’s primary targets—“non-naturalist” moral realists—who explicitly accept Harman’s Objection and so, of course, reject Quinean empiricism.” (2015: 83-84).

Clarke-Doane’s point is that (given non-naturalism) realistically construed moral facts cannot explain our moral beliefs independently of any evolutionary considerations. Given that debunkers are looking to provide an argument that shows our realistically construed moral beliefs are undermined *even if* they were justified prior to such an argument, they cannot be claiming that the absence of an explanatory connection undermines our realistically construed moral beliefs. This would imply they were

never justified to begin with, even before the debunker established their evolutionary genealogy.

What's more, Clarke-Doane claims that because of the impossibility of realistically construed moral facts explaining our moral beliefs, the non-naturalist clearly does not accept that the absence of such an explanatory connection undermines their beliefs. This would straightforwardly lead to all their moral beliefs being undermined. So a good argument against the non-naturalist cannot just assume that establishing such a genealogy defeats our moral beliefs, because the non-naturalist is likely to object to exactly this point.

Klenk echoes this point. Klenk considers the view that debunking arguments should be viewed as claiming that the absence of an explanatory connection might undermine the realist's moral beliefs, and writes:

“if an explanatory connection were required [...] then objectivist moral requirements could not be justified in the first place. Objectivists assume, and debunkers accept for the sake of argument, that there *are* justified moral beliefs whose truth is not implied by the best explanation of them.⁴⁶” (2019b: 257).

Consider the first point, about the impossibility of moral facts explaining our moral beliefs, prior to any evolutionary considerations. I deny this point. The only potential reason I can see in favour of this claim is that realistically construed moral facts, according to the non-naturalist, are causally impotent. But I do not think that the causal impotency of some realm of facts implies that those facts cannot explain our beliefs about them.

Take logical facts, for example. Assume logical realism, according to which mind-independent logical facts exist and are causally impotent. I think it is possible that, even if such facts are causally impotent, they might still explain our beliefs about them. For example, it is possible that logical facts constrain the physical world⁴⁷; it seems plausible that part of the reason why we have never encountered an object that is both completely red and completely blue is because its existence is logically impossible. In general, we only encounter objects that are not logically contradictory, and this might

⁴⁶ Klenk uses “objectivist” as I have been using “non-naturalist”.

⁴⁷ The idea that non-causal facts might explain physical phenomena by constraining them can be found in Lange (2016).

be because logical facts constrain what objects might possibly exist. The relation here, if it exists, is not causal (this is ruled out by the causal impotency of realistically construed logical facts). Instead, the logical facts have some non-causal influence on the physical world, and this non-causal influence may constitute an explanatory relation⁴⁸.

If this is true, then it is possible that creatures with correct logical intuitions were at an evolutionary advantage because they could better navigate the logically-constrained physical world. If the logical facts were different then the physical world be correspondingly different, and different logical intuitions would have been evolutionarily beneficial. This implies that, by constraining the physical world, logical facts play a role in explaining the content of our logical intuitions (and perhaps in explaining why we have a capacity for logical judgements in the first place). Assuming our logical beliefs are an outgrowth of these intuitions, it follows that our logical beliefs are explained by realistically construed, causally impotent logical facts.

These comments actually reflect a popular strategy for vindicating our mathematical beliefs, given mathematical realism, under the assumption that they can be given an evolutionary genealogy. As I indicated in the previous chapter, some commentators argue that having *correct* mathematical beliefs allows us to better navigate our environment, and this explains why evolution would have encouraged us to have those mathematical beliefs that correspond to realistically construed mathematical facts (Joyce, 2005: 182). As a result, our beliefs are in fact explained by mathematical facts, even if we assume mathematical realism according to which those facts are causally impotent. Sinnott-Armstrong writes:

“People evolved to believe that $2 + 3 = 5$, because they would not have survived if they had believed that $2 + 3 = 4$, but the reason why they would not have survived then is that it is *true* that $2 + 3 = 5$ [...] the truth of the belief explains why it is useful to believe it. In contrast, moral beliefs would be useful, so we would have evolved to believe them, even if they were not true. That is why the evolutionary explanation is supposed to undermine moral facts but not mathematical [...] facts.” (2006: 43)

Similarly, Allan Gibbard writes:

⁴⁸ For more on explanation by constraint see Lange (2016).

“Is there an explanation of why beings like us would tend to get mathematics right? [...] Simple counting and adding are needed for us to think about everyday experience. Our long-ago ancestors didn’t balance checkbooks, but the capacities later used in banking were needed in that earlier day for other tasks. Perhaps these were capacities for simple counting, addition, or subtraction [...] Like remarks would apply to simple geometry: good spatial visualisation must be important for hunters and warriors.”
(2003: 257)⁴⁹

Clarke-Doane (2012) has argued that this proposed vindication of our moral faculty fails. But his argument does not consist of simply pointing out that realistically construed mathematical facts are causally impotent, as it could if causally impotent facts were immediately barred from explaining our beliefs about them. Instead, he argues, the best evolutionary explanation of our mathematical beliefs actually implies that those beliefs are explained by first-order logical truths, or truths about the structure of our environment, rather than mathematical truths (2012: 327-333). But clearly, this does not imply that causally impotent facts are in principle incapable of explaining anything; it actually supports the contention that first-order logical truths can have some influence on our (mathematical) beliefs.

Obviously all this is highly speculative, and I do not mean to endorse any particular genealogy of our logical or our mathematical beliefs. My purpose is simply to establish that the causal impotency of a given realm of facts does not in itself entail that those facts play no role in explaining our beliefs about them⁵⁰. This is unsurprising under the assumption that non-causal explanations exist. In order to establish that some causally impotent realm of facts do not explain our beliefs about them, we need some reason for thinking that this is the case, and this is exactly what evolutionary debunking arguments provide for our moral beliefs. They claim that moral facts do not play a role in the explanation of our moral beliefs because there is good evidence for an evolutionary account of such beliefs, and (unlike evolutionary explanations for our other faculties for belief) there is no reason to think that the evolutionary benefit of such beliefs is dependent on their truth. Plainly, these considerations do not automatically apply to all beliefs in non-causal facts.

⁴⁹ For more examples of this kind of argument, see Sosa (2002) and Crisp (2006).

⁵⁰ This implies that we can agree with Nagel that the requirement that our moral beliefs be *causally* explained by the moral facts would beg the question against the realist (1986: 144). The requirement that they be explained by moral facts is not ruled out by their causal impotency, and so does not beg the question.

This implies that moral facts are not the kind of thing that are necessarily incapable of explaining our moral beliefs independently of any evolutionary considerations. This in turn implies that, if such an explanatory connection were required for our beliefs to be justified, our moral beliefs would not be automatically unjustified under the assumption of non-naturalism. An evolutionary debunking argument that relies on such a principle does not thereby imply that the non-naturalist's moral beliefs were necessarily unjustified prior to such arguments, or that evolutionary considerations are fundamentally irrelevant to the debunker's case.

Now consider Clarke-Doane and Klenk's further claim that non-naturalists explicitly deny that the absence of an explanatory connection between their moral beliefs and the moral facts undermines those beliefs, and will therefore be unphased by the debunker's claim that no such explanatory connection exists. Even if they do explicitly deny that the explanatory claim can undermine their beliefs, there seems to be nothing wrong in presenting an argument to the effect that their denial of this point is mistaken. If it is true that accepting such a non-truth-involving genealogy for our beliefs defeats those beliefs, then the realists who explicitly reject this claim are criticisable for their mistaken denial of this point. Their explicit acceptance that their moral beliefs are disconnected from moral facts would not insulate them from the sceptical implications of such a position if sceptical implications are indeed to follow. And establishing that the relevant explanatory disconnect does have these undermining implications would not mean that evolutionary considerations become irrelevant because, as I have just argued, evolutionary considerations are not irrelevant to the project of establishing that such a disconnect exists.

What Clarke Doane and Klenk's point does show is that we cannot assume it to be common ground, accepted by everyone, that such a genealogy would undermine our beliefs. We therefore cannot simply point out that such a genealogy exists and conclude that our moral beliefs have been undermined *without argument*. Granting that the non-naturalist already accepts the existence of such a genealogy, they are likely to object to this inference, and if we are to defend our debunking argument we have to provide some reason for thinking that the non-naturalist is thus mistaken.

Clarke-Doane raises this worry in objection to Joyce's debunking argument, and it might succeed if we were to view Joyce as simply presupposing that accepting an

explanatory disconnect between our beliefs and their truth can defeat those beliefs without argument. But I do not think we should accept such a reading, because Joyce argues for this position by analogy (and does so, to my mind, quite effectively). So Clarke-Doane has not shown that the debunker *cannot* take the non-naturalist's moral beliefs to be undermined by their acceptance of the not-truth-involving genealogy; he has just shown that doing so requires argument.

I therefore do not think there is any consideration stemming from the dialectic between moral realism and evolutionary debunking arguments that precludes the debunker from claiming that the absence of an explanatory connection between our moral beliefs and the moral facts undermines those moral beliefs directly, regardless of whatever further implications might follow such a disconnect. But this does not yet imply that debunking arguments, so interpreted, are successful. In order to defend this claim the debunker has to show that the explanatory claim can actually undermine; showing that the dialectic does not rule it out is not enough. I consider whether this is plausible in the next two chapters.

2.7: Conclusion:

In chapter one, I noted that debunking arguments typically make both explanatory and modal claims about the relationship between our moral beliefs and the moral facts, and that the relationship between these two claims (and their supposed role in debunking arguments) is not always clear. We have now seen that the success of minimalist responses to debunking arguments depends on the viability of interpreting those arguments as claiming the explanatory premise undermines our moral beliefs directly. In the following two chapters, I will investigate the plausibility of this claim, and will eventually defend an epistemological principle stipulating the kinds of attitude we could have towards the explanatory history of our belief that could undermine that belief. In the final two chapters I apply this principle to the debate surrounding evolutionary debunking arguments and minimalist responses.

Chapter 3: Explanationism vs Modal Security:

3.1: Introduction:

In this chapter I argue that accepting that our belief is explanatorily disconnected from the facts that make it true can defeat our belief regardless of the modal implications it has for that belief. That the debunker has not threatened the modal stability of our realistically construed moral beliefs therefore does not entail that they have failed to undermine those beliefs. This removes one way of resisting the claim that the explanatory disconnect implied by evolutionary debunking arguments defeats our realistically construed moral beliefs. In the next chapter, I will consider the specific attitude towards the explanatory history of our belief that can defeat that belief and whether or not evolutionary debunking arguments imply we should adopt that attitude.

I will use the term “explanatory concession” to describe an agent withholding belief in any explanatory relationship between their belief and the fact that makes it true⁵¹. I will use “explanationism” to describe the thesis that explanatory concessions can defeat our beliefs directly, independently of what they imply about the modal status of our belief⁵². An “explanatory constraint” is an epistemic principle that stipulates the specific kind of attitude we can have towards the explanatory history of our belief that defeats that belief⁵³.

The contrasting view, that whether or not a belief is defeated depends on what we believe (or have evidence to believe) about its modal stability, I call “modalism”. Clarke Doane’s Modal Security is the version of modalism that I will be primarily concerned with in this chapter, and it is this version of modalism with which I will be contrasting explanationism. But, as I will outline further below, other modalist positions are possible. A “modal concession” is an agent withholding belief that their belief is modally secure.

⁵¹ The term “explanatory concession” borrowed from Korman and Locke, who use it to mean something slightly more specific. See, for example, (2020: 316).

⁵² This is not to be confused with related kinds explanationism, such as explanationism about justification (McCain, 2014, 2015, 2017) or explanationism about epistemic coincidence (Faraci, 2019). I am specifically concerned with explanationism about epistemic defeat.

⁵³ This terminology is also borrowed from Korman and Locke (2020).

I begin by clarifying what I mean by “explanation” and by justifying my methodology for the rest of the chapter. I then draw from the literature to provide three examples in which explanatory concessions undermine beliefs even though they do not imply that the belief is modally insecure, which supports explanationism over Modal Security.

Next, I consider an objection from Clarke-Doane and Baras (2020) according to which these examples fail because the beliefs in question are in fact modally insecure, and their being undermined is therefore consistent with Modal Security. I respond by presenting an example of a belief being explanatorily disconnected from the facts that is modelled on the debate between debunking arguments and minimalist responses. I argue that the belief in question is undermined, and that Clarke-Doane and Baras are committed viewing this belief as passing Modal Security. As such, this example implies that an explanatory concession can defeat our beliefs even if that belief passes Modal Security, and this gives us reason to accept explanationism over Modal Security.

I end the chapter by considering two further objections to explanationism from Clarke-Doane and Baras. These objections arise specifically from the modalist camp; if they are successful, they imply that Modal Security, rather than explanationism, can best account for when a belief gets defeated. I argue that both of these objections fail and that, as a result, we ought to reject Modal Security in favour of explanationism.

This concludes my discussion of explanationism when it is contrasted with Modal Security. But it does not serve as a complete defence of explanationism; there are potential issues with explanationism that have nothing to do with modalism. I will therefore complete my defence of explanationism in the next chapter, where I address some of these independent concerns.

3.2: Explanation:

What do I mean by “explanation”? I will be understanding explanation in line with explanatory realism, according to which a correct explanation of some event describes what metaphysically determines that fact or event.

What is metaphysical determination? Canonical examples include causation and grounding. As Audi describes it, for X to metaphysically determine Y is for X to “*bring about or be responsible*” for Y (2012:690). Shaffer notes that both causation

and grounding feel as if they could be described as instances of “generation” or “production” (2016: 54). If X causes or grounds Y then X exists in virtue of Y, and this existing-in-virtue-of relation is central to the concept of metaphysical determination. It may be that determination is a primitive notion that resists definition⁵⁴, but I hope these brief comments help to illuminate the relation in question.

According to explanatory realism, to explain Y is to attempt to give information about what metaphysically determines Y. To explain X in terms of Y is to commit oneself to Y metaphysically determining X. Whether or not explanations are accurate will therefore depend on whether they reflect mind-independent determination relations between the facts cited in that explanation. When we speak of some explanans *explaining* an explanandum, this can be read as the claim that the event described by the explanans metaphysically determines the event described by the explanandum⁵⁵. To deny that some fact P explains our belief in P is to deny to P goes any way to metaphysically determining our belief in P.

Though explanatory realism is a popular position, it is not uncontroversial⁵⁶. Elanor Taylor has recently argued against explanatory realism by citing a number of apparent examples in which some fact Q is cited in an explanation of P but Q is not meant to metaphysically determine P. I do not want to engage with the debate about explanatory realism here; what is most important is to clarify that I am using the term explanation in line with explanatory realism. If explanatory realism is false, then my comments ought to be read as referring to a specific kind of explanation (i.e. that which describes metaphysical determination relations). Nobody, as far as I am aware, denies that this kind of explanation exists; the controversy concerns whether any other kind of explanatory relation exists as well.

One reason for not engaging with the explanatory realism debate is because it takes us too far afield from the purposes of this thesis. Another is that, as far as I can tell, it is irrelevant to the debate between evolutionary debunking arguments and minimalist responses to those arguments. Minimalists do not insist that our moral beliefs can be explained by the moral facts in some sense that is not captured by explanatory realism.

⁵⁴ Audi suggests as much (2012: 690).

⁵⁵ Here, in line with Kim, I am thinking of explanans and explanadums as statements or propositions that make up explanations (1984).

⁵⁶ For endorsements of explanatory realism, see Lewis (1986), Kim (1988), Ruben (1990) Skow (2014) and Audi (2015)

They either deny that explanatory considerations of any kind matter for our beliefs (as with the modal security response), or they defend the position that our moral beliefs are in fact connected to the moral facts by metaphysical determination relations (as with third-factor replies). None, as far as I am aware, defend the position that justification for moral beliefs is salvaged because of some anti-realist explanatory relation that might obtain between our moral beliefs and the moral facts.

Though I lack the space to do so, I think it would be possible to construct examples in which our belief in P is defeated because it bears no metaphysical determination relations with P itself, even though it may have *other* kinds of potentially (anti-realist) explanatory relations with P. I therefore think that the kind of explanatory relations that matter for the question of undermining defeat are metaphysical determination relations, and this is why I opt for this particular understanding of explanation. Again, though, whether this implies that all explanatory relations matter for defeat or whether it implies that only a particular subset of those relations matter depends on whether explanatory realism is viable, and I will not be engaging with that debate in this thesis.

3.3: Standards of Justification:

Before I make my case for explanationism over Modal Security, I need to explain and justify a methodological point that will guide the discussion of the next two chapters. In these chapters I will be assuming that the standards that govern the justification of our moral beliefs are not fundamentally different from the standards that govern our non-moral beliefs; our moral beliefs do not enjoy their own special standards of justification just in virtue of being about morality. More specifically, I will assume that conditions that are sufficient for defeating a non-moral belief are equally sufficient for defeating a moral belief. This latter claim follows from the first; if the standards of justification are the same for both our moral and non-moral beliefs, then the conditions under which previously justified beliefs become unjustified (and get defeated) will be the same for both moral and non-moral beliefs. To deny this would be to claim that there are situations in which our moral beliefs but not our non-moral beliefs could be justified, thus implying different standards of justification for the two.

Hence the following methodological point: when I am justifying an epistemological principle that I will ultimately apply to our moral beliefs in the context of evolutionary

debunking arguments in chapters five and six, I will often rely on examples and intuitions involving non-moral beliefs. I assume that these arguments and intuitions can support an epistemological principle that can be legitimately applied to our moral beliefs as well.

In saying this, I do not mean that I will be assuming the obviously false position that the sources of justification are the same for our moral and non-moral beliefs alike. My assumption is just that, if there are conditions according to which our perceptual beliefs, our mathematical beliefs and our other non-moral beliefs get defeated, then our moral beliefs are defeated in suitably similar conditions as well. Our moral beliefs are not let off the hook just because their subject matter is morality.

There are a number of reasons why I feel entitled to make this assumption. Firstly, the position that there might be fundamentally different standards of justification for moral and non-moral beliefs seems particularly ill-suited to the kind of non-naturalist moral realism that minimalist responses hope to defend. Part of the package of this position is the idea that moral facts exist in importantly similar, bog-standard ways to other, non-moral facts. Thus, Fitzpatrick (2008) outlines one of the first (and least controversial) tenets of non-naturalist realism as follows:

“Ethical claims purport to state facts (attributing ethical properties to actions, persons, policies, etc.), and so are straightforwardly true or false in the way that other purportedly fact-stating claims are, by accurately representing the facts or not” (2008: 161).

And, in characterising his robust metanormative realism, Enoch explains

“The crucial point is that, in whatever sense there are physical facts, there are normative ones; in whatever sense there are truths in biology, there are in normative discourse; in whatever sense in which there are mathematical properties, there are normative ones.” (2011: 5).

The idea that there are genuine moral facts, and that these moral facts are not of a fundamentally different kind to non-moral facts (for example, by being in some way just constituted by our attitudes towards them) is a major part of what sets non-naturalist realism apart from other meta-ethical positions. And, given that non-naturalism is committed to moral facts being, in a sense, just like non-moral facts, it would be surprising if our standards for justified beliefs about moral facts diverged

significantly from the standards that govern beliefs about non-moral facts. Perhaps this is why, when discussing how standards of justification apply to our moral beliefs, Enoch writes:

“on no theory of epistemic justification I am aware of do normative beliefs constitute an interesting particular instance of beliefs, an especially problematic class of beliefs [...] Whether you are a coherentist or a foundationalist (or perhaps hoping for some middle ground between them), whether you are an internalist or externalist about epistemic justification (or perhaps hoping for some middle ground between them), whether or not you think that epistemic justification is conceptually tied to epistemic responsibility, whether or not you like working with a conception of epistemic virtue - whatever your theory of epistemic justification, it is hard to see any special difficulties applying it to normative beliefs.” (2010: 416)

Here Enoch is stressing that the standards for our moral beliefs being justified are no more stringent than for our non-moral beliefs, but his comments (and his commitment to non-naturalism) imply that those standards are also no *less* stringent. If they were, then normative beliefs would in fact constitute “an interesting particular instance of beliefs”, *contra* his comments above.

It would at least require substantive argument for the non-naturalist to show how the standards of justification for beliefs about moral and non-moral facts differ, despite there being no such fundamental difference in the nature of the facts that these beliefs are attempting to accurately represent. This justifies a presumption that, under the assumption of non-naturalism, there is no such divergence.

I am only aware of one moral realist who explicitly defends a fundamental difference in the standards of justification for our moral and non-moral beliefs. Ronald Dworkin argues that our moral beliefs are uniquely immune to defeat by explanatory considerations because we cannot evaluate counter-factuals in which moral facts are different, and in order for some fact X to feature in the best explanation of another fact Y, we have to be able to make sense of and consider a counter-factual scenario in which X does not exist. (1996: 119; 2011: 73-4). Because we cannot understand the relevant counter-factual scenario for our moral beliefs (as there are no possible worlds in which moral facts are different), this implies that moral facts cannot figure in their best explanation. And presumably, if moral facts can *never* feature in the best

explanation of our moral beliefs, then our moral beliefs cannot be undermined when we find out that they are not explained by those facts.

The problem with this argument is that some fact can feature in the best explanation of that belief even if there are no possible worlds in which that belief is false. My discussion of logical and mathematical facts at the end of the last chapter implies as much. By showing how logical or mathematical facts could theoretically explain our beliefs about them, I have shown that the truth of our beliefs can feature in their best explanation even if there are no possible worlds in which those facts do not hold. In her discussion of Dworkin's argument, Sarah McGrath provides the following example:

“Imagine a mathematician who initially has no opinion about whether some mathematical conjecture is true or false. Suppose that she subsequently succeeds in proving the conjecture and thus comes to believe the relevant proposition on the basis of the proof. In this case, the explanation of why the mathematician currently believes the proposition is that she succeeded in proving the theorem. But of course, that explanation entails that the relevant proposition is true.” (2019: 180-181).

It therefore seems that we need not imagine a possible world in which the fact does not hold in order for that fact to feature in the best explanation of our belief. As this was the central premise in Dworkin's argument, Dworkin's argument fails.

Another reason I feel entitled, in what follows, to presume the same standards of justification for our moral and non-moral beliefs is that the minimalist responses I am assessing at no point appeal to a fundamental difference between these standards. They appeal to epistemological claims that could, presumably, be applied to our moral and non-moral beliefs alike. So even if there *are* different epistemological standards for our moral and non-moral beliefs, the strength of the responses that I will be considering do not rely on or exploit these differences. These responses do not claim that, in virtue of the unique epistemological standards that apply to our moral beliefs, our moral beliefs do not fall prey to debunking arguments. They seem instead to argue that, according to regular standards of justification (that can be applied to both our moral and non-moral beliefs alike) evolutionary debunking arguments do nothing to diminish justification for our moral beliefs. So I can show that these responses fail just by showing that, according to regular generally acceptable standards of justification,

evolutionary debunking arguments do threaten our moral beliefs under the assumption of non-naturalism.

Now, I cannot establish definitely that none of the authors employing minimalist replies implicitly assume some such kind of fundamental difference, and that this informs the epistemological claims they make in the process of arguing for their position (though, as I pointed out above, when these authors are non-naturalists, this seems a difficult position to maintain). Maybe their position is that the standards of justification to which they appeal (and in virtue of which moral beliefs can remain justified) apply specifically to our moral beliefs, implying that an analogous debunking argument made against, say, our perceptual or mathematical beliefs would in fact be successful.

However, given the aforementioned presumption *against* a difference in these standards, forcing them to make this explicit will be a useful step forward in the direction of evolutionary debunking arguments. It will show that their arguments work *only* under the controversial assumption that the standards for justification for our moral beliefs are crucially different and, in a sense, weaker, than our non-moral beliefs. And then the relevant to authors will have to explicitly defend this claim in order to fully defend their position. This is particularly pressing as these authors do not even mention, let alone defend, this controversial presupposition. I will therefore have shown that, at the very least, these responses are importantly incomplete and that there is much more work to be done before they are viable.

A final consideration is that the intuitions supporting my epistemological arguments are not sensitive to the subject matter of the beliefs in question. I can consider various non-moral beliefs with a particular property, and I feel intuitively that the belief is unjustified and should be withheld. And when I consider cases of a *moral* beliefs having this property, nothing changes; I still feel intuitively that the belief is unjustified. I therefore feel that the moraliser would be just as unjustified in maintaining beliefs with this property as the non-moraliser. This supports my assumption that the considerations I am about to canvass apply equally to both moral and non-moral beliefs.

This closes my reasons for proceeding as though moral beliefs are not subject to different fundamental epistemological standards as our non-moral beliefs. Obviously,

whether or not this is true is unlikely to be settled decisively in an introductory section of a chapter of a PhD thesis. But I do hope to have justified working with this assumption going forwards.

3.4: Counterexamples to Modal Security:

In order to arbitrate between explanationism and Modal Security, we need cases in which an agent makes an explanatory concession but has no good reason to think that their belief, if true, is modally insecure⁵⁷. If, in such cases, the belief is defeated, then this implies that an explanatory concession can undermine our belief directly regardless of what modal implications it has for that belief. We should therefore accept explanationism over Modal Security.

One relevant example is given by Lutz:

“Mischievous Lab Assistant: An early chemist is running an experiment to determine the nature of water. A pile of experimental data confirms—water is H₂O! But then the chemist learns that his lab assistant played a practical joke on him: all of the data that was supposed to confirm the hypothesis that water is H₂O was actually falsified by the lab assistant.” (2018:1116).

The fact that water is H₂O is generally taken to be metaphysically necessary, which implies that the chemist’s belief that water is H₂O is trivially sensitive. There are no possible worlds in which his belief is false, and therefore no possible worlds in which his belief is false but she has the belief. Still, when the chemist learns that the evidence has been fabricated, he should obviously give up his belief.

Lutz presents this as a counterexample to the position that undercutting defeaters work by giving us reason to doubt the sensitivity of that belief⁵⁸. As such, it is not yet a counterexample to Modal Security, which entails that information can also defeat our belief by giving us reason to doubt its safety. It seems as though the information that the lab assistant falsified their data is good reason to doubt the safety of that belief, and so it is consistent with Modal Security that the belief in question is defeated. If the

⁵⁷ A belief being “modally insecure” means that it is defeated according to Modal Security. If a belief not defeated according to Modal Security, the belief is modally secure. Similarly, the “modal security” of a belief refers to its property of being undefeated according to Modal Security, and its “modal insecurity” refers to its being defeated according to Modal Security.

⁵⁸ This is how defeaters work according to Pollock (1987: 485).

lab assistant had never falsified their data, then chemist would not believe that water is H₂O even though water would, of course, still be H₂O⁵⁹. So the chemist's belief could easily have been wrong, and the fact that his belief is defeated is consistent with Modal Security.

What the case needs, in order to challenge Modal Security, is some constraining influence on the content of chemist's belief such that chemist could not easily have failed to believe that water is H₂O even though his assistant falsified the data. We can therefore amend the example as follows:

Mischievous Lab Assistant 2: An early chemist is running an experiment to determine the nature of water. A pile of experimental data confirms—water is H₂O! But then the chemist learns that his lab assistant played a practical joke on him: all of the data that was supposed to confirm the hypothesis that water is H₂O was actually falsified by the lab assistant. The chemist learns that the lab assistant has a track record of playing this same trick on a number of other unsuspecting chemists; it is something the assistant feels compelled to do. This is frustrating because the scientist cannot fire the lab assistant; the scientist cannot conduct their research without an assistant, and the lab assistant is the only qualified person in town.

In this example, I have tried to implement a constraining influence on the content of t chemist's beliefs. It seems that t chemist could not have easily failed to hold his belief in water being H₂O; he could not easily have done without a lab assistant, he could not easily have done without *this specific* lab assistant, and this specific lab assistant could not easily have failed to falsify the data. Assuming it is no accident that the chemist forms their beliefs on the basis of the available evidence, the only way he might *not* have had this belief seems to be if he had never been a chemist to begin with, but this possibility is presumably too distant to matter⁶⁰.

In this example, then, the information that the lab assistant has been falsifying data gives the chemist no reason to think that his belief, if true, is modally insecure. The fact that water is H₂O is metaphysically necessary, and so obtains in all possible

⁵⁹ This assumes that the belief is not overdetermined, such that if the fake evidence had not explained the belief, then genuine evidence would.

⁶⁰ If not, then by all means construct another example in which this difficulty is removed; perhaps the chemist was in some way *destined* to become a chemist. For reasons outlined at the end of this section, I am confident that our resulting intuitions will still support explanationism over Modal Security.

worlds. The fact that the lab assistant falsified the data does not imply that the chemist could easily have had a different belief about water, because it seems that the lab assistant was *bound* to falsify this data. As such, when the chemist accepts the absence of an explanatory connection between his belief and the fact that makes it true, he does not have reason to doubt the modal security of that belief, under the assumption that the belief is true. The content of his belief and the relevant fact were bound to align, and there are no close possible worlds in which they do not.

Still, just as with the first example, the chemist's belief that water is H₂O is defeated when he accepts that his lab assistant falsified the data. By accepting that his lab assistant falsified his data, the chemist is accepting that the data (and in turn his belief) bears no explanatory relation with the fact that water is H₂O. If my assessment of the example is correct, this supports explanationism, according to which explanatory concessions can undermine beliefs directly regardless of what they imply for the modal security of our beliefs. Correspondingly, it counts against Modal Security, according to which information has to give us reason to think our beliefs are modally insecure in order to defeat those beliefs.

Another example of this kind is presented by Korman and Locke (2020):

Neora:

“On the basis of clear and distinct intuitions, Neora believes in an all-powerful deity. Later, Agent Smith convinces her that she is part of a computer simulation. He tells her that the designers had a terrible time getting consciousness and cognition to arise in the simulation, but—through endless trial and error—found that they could achieve this result only by programming the inhabitants to be strongly disposed to believe in an all-powerful deity. Without such beliefs, the simulations would break down before they even got going. Neora believes everything he tells her. And she believes that the deity (if it does exist) had nothing to do with her religious intuitions and associated beliefs. Despite having now accepted all this, she doesn't abandon her belief in an all-powerful deity.” (2020: 322)

It seems clear that Neora ought to give up her belief in an all-powerful deity once she accepts that her belief in God was implanted by designers for reasons that had nothing to do with God's existence. In other words, her belief is defeated when she accepts that her belief has no explanatory relationship with the fact that makes it true.

Importantly, however, this information does not give Neora reason to think that her belief is modally insecure. If God exists, then presumably Their existence is metaphysically necessary such that there are no possible worlds in which God does not exist. For reasons already rehearsed, this implies that her belief is trivially sensitive regardless of its explanatory history.

Furthermore, it appears Neora could not have easily been wrong about her belief. Her existence is predicated on her being disposed to have the belief, and so there are no close possible worlds in which she falsely believes that God does not exist. The information about the origin of her belief does not give Neora reason for thinking that her belief is unsafe.

But despite not giving Neora reason to think her belief is either insensitive or unsafe, the information still seems to remove Neora's justification for her belief. When Neora finds out that she only believes that God exists because of a practical issue to do with the possibility of programming life (rather than anything to do with the existence of God) she is no longer justified in holding this belief. As such, accepting an explanatory disconnect between the belief and the fact that makes it true must be able to defeat Neora's belief without giving her reason to think that her belief is modally insecure. This implies that information can defeat our belief without giving us reason to think that our belief is unsafe or insensitive, and we should accept explanationism over Modal Security.

For our third and final counter-example to Modal Security we can draw on David Faraci's work on epistemic coincidence (2019). According to Faraci's explanationist account of epistemic coincidence:

“For any true belief B and truth it concerns T, B is coincidentally true if (a) there is no unified explanation for B and T; (b) T does not explain B; and (c) B does not explain T”. (2019:7).

In order to motivate this account of epistemic coincidence over an alternative modal account, Faraci constructs examples in which an agent's beliefs are modally secure but lack any explanatory connection with the facts that make them true. By showing that the belief in question is coincidentally true, Faraci hopes to support his view that the absence of an explanatory connection is sufficient for one's belief to be coincidentally true.

Because Faraci is concerned primarily with epistemic coincidence rather than epistemic defeat, his examples do not directly imply that our beliefs can be defeated even if modally secure. But they can be easily amended such that they can help support a judgement on whether or not explanatory concessions can undermine directly. Here is one such amended example:

Eula:

Eula is defeasibly justified in forming beliefs about which numbers are prime by consulting the Source, and there is no available evidence that the Source is untrustworthy. In fact, the Source is outputting numbers at random. Eula consults the Source to form beliefs about which numbers are prime. The numbers the Source outputs at random are all prime numbers. Eula's resultant beliefs are therefore true; and there is no good evidence to the contrary. At every possible world, Eula's counterpart forms beliefs about which numbers are prime by consulting the Source's counterpart, and at every possible world, the Source's counterpart delivers the same answers as at the actual world. Eula then finds out that the source has been outputting numbers at random, and that her counterpart forms beliefs via the same method, and gets the same results, in every possible world.

In this case, then, Eula's beliefs are maximally deterministic. She forms the very same beliefs about prime numbers in all possible worlds (using the very same method) and the truth of those beliefs are metaphysically necessary. There are therefore no possible worlds in which the content of her prime beliefs and the prime facts fail to align. But, when she accepts that the source is outputting numbers at random, she thereby accepts the absence of an explanatory connection between her beliefs and the facts that make them true. And, it seems that Eula ought to give up her prime beliefs when she accepts that the source is outputting numbers at random, regardless of the fact that this does not threaten the modal security of her beliefs. This implies that explanatory concessions can defeat beliefs directly, even if they do not imply that those beliefs are modally insecure.

We therefore have three examples in which an explanatory concession defeats our belief even though the modal security of that belief is not threatened. This provides good evidence for explanationism over Modal Security. In these examples the defeating force of the explanatory concession cannot come from some further troubling modal implications because there are no such implications. We therefore

cannot explain away the apparent defeating force of an explanatory concession with appeal to modal conditions like sensitivity or safety. We need an explanatory constraint on belief to account for this fact.

I think that there is also a further important feature of these examples. This is that, once we have stipulated whatever fact that ensures an explanatory disconnect between belief and truth, modal considerations strike us as *irrelevant* to the question of whether or not the belief in question remains justified. When explanatory considerations are separated from any influence they might have on the explanatory status of our beliefs, they feel intuitively like the *wrong sort of thing* to affect the justificatory status of a belief.

Take Korman and Locke's Neora example. The consideration that the existence of God is necessary if true, rather than contingent, does not nudge our intuitions in the direction of Neora's belief being justified rather than unjustified. Swap out the belief in a necessary truth and replace it with a belief in a contingent truth and our intuitions on whether the belief in question is justified remains unchanged. If Neora found that her belief in the metaphysically contingent existence of Napoleon had a similar genealogy, would our intuition that her belief is defeated be any more robust? I think not. Similar comments apply about the other examples; if the chemist's belief had been in some contingent scientific fact, or if Eula had formed beliefs about the weather rather than prime numbers, the contingency of the relevant facts would not factor into our judgement of whether her belief is ultimately justified.

Furthermore, the fact that Neora could not have existed without being primed for her belief in God (thus modally-securing the content of her belief) also seems totally irrelevant to whether or not we judge Neora's belief is justified. That she now takes the content of her belief to be modally fixed in this way does not give Neora any greater justification for retaining this belief than if the content of her belief was *not* guaranteed by her very existence. If, for example, the programmers had imbued her with these beliefs because a large multi-sided die unexpectedly dictated they do so, this information would not make her beliefs less justified than finding out that the programmers were *bound* to imbue her with these beliefs for practical reasons. Questions about how easily she might have failed to have a belief with this content seem irrelevant to whether her belief is now justified. As long as these reasons were

totally independent of the belief's truth, how likely it is that she would have beliefs with these particular contents becomes irrelevant. Likewise, the differences between the first and second mischievous assistant example seem to have no impact on our intuitions about whether or not the subsequent beliefs are justified.

These considerations imply that, once we have specified that a belief in P is explanatorily disconnected from P, tightening or loosening the modal stability of that belief does not affect our intuitions about whether the belief in question is justified. The modal considerations seem beside the point. This is exactly what the explanationist would expect.

We can see this point even more clearly if we imagine the agent in question attempting to retain justification for their belief by citing considerations about its modal security. Imagine that the chemist in the second example finds out that his lab assistant had falsified the data but then continues to hold his belief that water is H₂O. When asked to justify why he still believes that water is H₂O, he cites the fact that he could not easily have failed to have the belief; after all, he was *bound* to have a lab assistant that falsified the data in this way. What's more, he points out, water being H₂O is metaphysically necessary if true, and so there are no possible worlds in which it is false.

Upon hearing this, we would presumably conclude that the chemist was severely misguided about what constitutes a good reason for continuing to hold a belief. Considerations like these, about how easily the chemist might have had a belief that diverged from the facts (under the assumption it is true to begin with) are not the right sort of consideration if he is looking to successfully justify his continuing to hold this belief. What he needs to do is establish that the data on which he basis his belief does in fact bear an actual relationship with the fact that water is H₂O. It is by arguing for this point that the chemist could justifiably retain his belief. Again, this is exactly what the explanationist would expect.

I think this can explain the frustration of those sympathetic to debunking arguments when they are exposed to the modal security response. The modal security of our beliefs seem beside the point, and we can argue for the modal security of such beliefs, from either end, without thereby changing our intuitions about whether such beliefs can remain justified. It may well be the case that, if true, our moral beliefs could not

easily have been false, but this in itself irrelevant to the justificatory status of those beliefs. It may be difficult to describe or argue for the absence of the relevant explanatory connection without making clumsy and objectionable modal comments, but once the absence of such a connection is fully understood its epistemological ramifications become clear.

I have argued, then, that the examples I have considered in this section strongly imply we should accept explanationism over Modal Security. As we will see, however, Clarke-Doane and Baras have objected that examples like these do not, and potentially cannot, have this implication. I turn now to these objections.

3.5 These Examples Fail to Support Explanationism:

(i) In a recent article, Clarke-Doane and Dan Baras (CD&B) consider the kind of counter-example to Modal Security that I have been pressing in the last section (2021). Before assessing their response, it is important to get clear on what we are trying to establish, and what conditions an example will have to satisfy in order help us achieve this goal. We are trying to determine whether an explanatory concession can undermine our belief directly even if it does not imply that our belief fails Modal Security. In order to establish whether this is the case, we need an example in which (i) there is an explanatory concession and (ii) Modal Security is satisfied. Once we have an example that fulfils these requirements, we need to then establish whether the belief in question has been undermined. If the belief is defeated then this will imply that explanatory concessions can undermine directly, and if the belief remains undefeated this will imply that explanatory concessions cannot undermined directly and this will support Modal Security.

We now need to outline two clarifications that CD&B make regarding Modal Security. As a reminder, Modal Security states:

“Modal Security: If evidence, E, undermines our belief that P, then E gives us reason to doubt that our belief is sensitive or safe.” (2021: 162)

CD&B first clarify that sensitivity and safety should be understood relative to methods for forming beliefs. A belief in P is sensitive iff, were it the case that not-P, we would still have believed that P “had we used the method that we actually used to determine whether P” (2020: 163). Similarly, our belief in P is safe iff we could not easily have

been wrong about P “using the method we actually used to determine whether or not P” (2020: 163)⁶¹.

The second clarification is that our belief is not safe just if we could not easily have been wrong about P using the method that we actually used to arrive at our belief in P. In order for our belief to be safe, it must also be the case that we could not have easily had false beliefs about any proposition *similar enough to P*. If we calculate tips by using rules that are unreliable in general, but that happen to give correct outputs about “what is 20% of \$9.98”, then our belief in this proposition unsafe because we could have easily had false beliefs about propositions that are closely related to the content of this belief (2021: 164).

With these clarifications in place, Clarke-Doane and Baras consider two examples of the kind given in the previous section. The first is as follows:

“Jan is defeasibly justified in forming beliefs about which numbers are prime by consulting a computer. Perhaps its results to date have been verified by direct computation, or he has been told by any otherwise trustworthly mathematician that this computer is reliable. Jan consults the computer to form beliefs about which numbers are prime. Jan then gets evidence, E, that the machinery has been stuck in the last 10 cases. As it happens, the last 10 numbers considered have been prime.” (2021: 173)

In this example, we appear to have a case that satisfies conditions (i) and (ii) laid out above. That is, (i) Jan’s finds out that her belief does not have an explanatory relationship with the fact that makes it true, and (ii) she has no reason to think that her belief, if true, fails Modal Security. We can therefore attend to the status of her belief to determine whether an explanatory concession can undermine directly, and it seems that it can: her belief has been undermined.

But CD&B point out that Jan’s belief actually fails modal security, because she could have easily have had false beliefs about propositions that are very close to the content of her actual beliefs: namely, about whether composite numbers were prime (2021: 173). If, after the machine had gotten stuck, Jan had inputted composite numbers rather than exclusively prime numbers, she would have had false beliefs about whether these

⁶¹ For now, ignore the generality problem about the correct way to individuate methods. I will address this below.

numbers we prime, and this means that her beliefs about the last ten numbers are not actually safe. As a result, this first example does not satisfy condition (ii), and it therefore cannot be used to establish whether an explanatory concession can undermine directly. Though Jan's belief is undermined, it could be undermined by its failure to satisfy Modal Security.

Here is Clarke Doane and Baras' second example, based on Faraci's example that I used above:

“Aya is defeasibly justified in forming beliefs about which numbers are prime by consulting a computer. Aya consults the computer to form beliefs about which numbers are prime. As it happens, the numbers the computer outputs are all prime numbers. Moreover, these events are *maximally deterministic*. At every metaphysically possible world, Aya's counterpart forms beliefs about which numbers are prime by consulting the computer's counterpart and at every possible world, the computer's counterpart delivers the same answers as the actual world. Aya then gets evidence, E, that the computer has been generating numbers using a random number generator.” (2021: 173)

In this case, it seems that Aya's belief is counter-factually stable in an extremely strong sense, and that her belief therefore satisfies Modal Security. Aya could not have easily had false beliefs about which numbers are prime by checking the computer. She could not easily have had false beliefs about whether *any* numbers are prime, because the machine outputs only prime numbers in all possible worlds. Aya's use of this infallible method is also stipulated as being metaphysically necessary, and so she could not have easily used another method for forming her prime beliefs.

CD&B do not provide any account of how to individuate the relevant method for forming beliefs. As such, they note that it is possible that the relevant method is in fact “trusting a random number generator to answer mathematical questions” and this is not a reliable method for forming beliefs (2021: 174). This would imply that this example also fails (ii) and so cannot be used to argue for explanationism over Modal Security. But they grant, for the sake of argument, that Aya's method for forming beliefs should be individuated in such a way that her belief passes Modal Security.

CD&B also note that this example is difficult to imagine, and that this in itself gives us reason to doubt our intuitions about it (2021: 173). Still, they acknowledge that

some commentators claim to have the intuition that Aya's beliefs are defeated. If so, we appear to have a case that satisfies (i) and (ii), and in which the belief is defeated. This supports explanationism over Modal Security.

The example fails, according to Clarke Doane and Baras, because it is incoherent (2021: 175). How could it be metaphysically necessary that someone form their beliefs in prime numbers by consulting a computer? Surely this fact is contingent. Furthermore, how could a *random* number generator give the same outputs in all possible worlds? It could not, and that it is stipulated as doing so in the above example implies that the above example is incoherent. Because the example is incoherent it cannot be used to support explanationism over Modal Security⁶². This implies that we should add a third condition that an example needs to satisfy in order to help us arbitrate between explanationism and Modal Security. It must:

- (i) involve an explanatory concession
- (ii) satisfy Modal Security and
- (iii) be coherent.**⁶³

So the first example appeared to be a coherent case in which there was an explanatory concession, modal security was not threatened, and our belief was undermined. In other words, it appeared to satisfy conditions (i)-(iii) and to support explanationism over Modal Security. On closer inspection, however, it became clear that the example actually involved evidence that threatened the modal security of the belief, and so therefore failed (ii) and did not support explanationism. The second example also appeared to satisfy conditions (i)-(iii) above, but it turns out that the example is incoherent and therefore fails (iii), and so also cannot be used to support explanationism.

Following the discussion of these two examples, CD&B conclude that examples of this kind cannot threaten Modal Security because "They are either unimaginable, not undermining, or may, in fact, involve evidence of lack of safety" (2020: 175). This implies that any coherent example of an explanatory concession for a belief that passes modal security (i.e any case in which conditions (i)-(iii) are satisfied) will be a case in

⁶² Clarke Doane and Baras do not provide explicit argument in favour of this claim, but I am happy to grant it going forward.

⁶³ This condition is usually implicit in any potential use of counter-examples, but for the sake of the foregoing discussion it is helpful to make it explicit.

which that belief is not undermined. Above I noted that if we could give an example in which (i)-(iii) are satisfied but the relevant belief remains undefeated, this would imply that explanationism is false and Modal Security true. It appears we should therefore accept Modal Security over explanationism.

But this conclusion too quick because CD&B have not actually provided an example that satisfies these three conditions. They first present an example in which conditions (i) and (iii) are satisfied, but condition (ii) is not. They then present an example in which conditions (i) and (ii) are satisfied, but condition (iii) is not. They then conclude that, in an example that satisfies (i), (ii) and (iii) our belief will not be undermined. But they have not provided an example in which (i), (ii) and (iii) are satisfied, and have therefore not licenced any opinion on whether our belief would be undermined in an example in which all three conditions are satisfied. They have just shown that two examples that appear to satisfy these conditions do not actually satisfy these conditions, rather than showing what would follow in an example that does.

In order to use examples such as these to support Modal Security over explanationism we need a case in which (i)-(iii) are satisfied and the belief in question is not undermined, and so showing that the two examples fail to satisfy (i)-(iii) shows just as much that they fail to support Modal Security over explanationism as it shows that they fail to support explanationism over Modal Security. To conclude on the basis of these examples that an example that succeeded in satisfying (i)-(iii) would result in an undefeated belief and would therefore vindicate Modal Security is unwarranted.

CD&B do comment that there is an inherent tension involved examples of the kind we have been discussing:

“Either the worlds in which we have different, and so false, beliefs are “distant” (or nonexistent), or not. If so, then it is hard to see why, epistemically, we should care about them. If not, then there is nothing to preclude us from saying that the subject’s beliefs are undermined by evidence of their lack of safety.”
(2021: 175)

The point seems to be that, in these examples, either there are close possible worlds in which our belief is wrong and then the example fails (ii), or there are not close possible worlds in which the belief is false, and then the example may satisfy conditions (i)-(iii), but the belief in question remains undefeated. Perhaps it is this consideration,

rather than anything to do with their considered examples, on the basis of which CD&B conclude that beliefs that satisfy all three conditions would not be defeated

The problem is that, when we are trying to determine whether an explanatory concession can defeat our beliefs independently of what it implies for the modal security of those beliefs, this reasoning begs the question against the explanationist. The debate between explanationism and Modal Security is a debate about whether information can undermine even if it gives us no reason to think that there are close possible worlds in which our belief is different and therefore false. To assume that there need to exist such possible worlds (or that such possible worlds need to be close) is to assume that information can only defeat to the extent that it implies our beliefs are modally insecure, and this is to assume the very issue at hand. This consideration therefore cannot be the basis on which they conclude that Modal Security is true and an example that satisfies (i)-(iii) would result in an undefeated belief. What we need are coherent examples in which beliefs are modally secure, but there is no explanatory connection, in order to judge whether we *should* care about the closeness or existence of possible worlds in which our belief is false.

Perhaps CD&B take their second example to imply that there are *no* coherent examples in which there is an explanatory concession for a belief that fails Modal Security. In other words, examples that pass conditions (i) and (ii) will necessarily fail (iii). It is therefore impossible to construct examples that fulfil the requirements necessary to support explanationism over Modal Security, or vice versa. This would undermine the explanationist's use of those examples discussed in section 3.4.

But this cannot be CD&B's position because they are committed to there being cases in which all three of these conditions are satisfied. According to CB&D, various genealogical debunking arguments (including evolutionary debunking arguments against moral realism) imply that the relevant beliefs are not explained by the facts that make them true, but the beliefs in question are undefeated because they do not fail Modal Security. And, presumably, there is nothing incoherent about the situation that they themselves argue we find ourselves in once we accept the debunker's evolutionary genealogy of our moral beliefs.

This shows that there is at least one coherent example in which our belief in P is explanatorily disconnected from P but our belief satisfies Modal Security. To deny

that this is possible, and to accept the explanatory disconnect typically argued for by debunkers, would be to accept that our moral beliefs fail Modal Security and are therefore undermined by evolutionary debunking arguments. But the fact that our moral beliefs can pass Modal Security and remain justified even if there is such an explanatory disconnect is precisely the point of Clarke Doane's minimalist response to debunking arguments⁶⁴.

So Clarke Doane and Baras are committed to the possibility of there being cases that satisfy (i) to (iii), but do not actually consider any examples of such cases. They claim that their second, incoherent example is difficult to imagine and that this threatens the validity of the intuitions that it invokes. Perhaps they mean to imply that *all* cases that satisfy conditions (i) and (ii) are similarly difficult to imagine (coherent or not) and that this is why they do not consider any examples that satisfy (i)-(iii).

But even this can't be true; there is nothing particularly difficult or complicated about the claim that our moral beliefs have some complete evolutionary explanation that does not involve their truth. And, as CD&B have argued, this situation is one in which our moral beliefs pass Modal Security and conditions (i) and (ii) are therefore satisfied. If there is a worry that understanding the relevant evolutionary explanation might render this situation irredeemably complicated, there seems to be no reason to think we could not remove the evolutionary details and provide a much simpler non-truth involving explanation of our moral beliefs. In this simpler scenario, our moral beliefs would presumably be explanatorily disconnected from the facts that make them true while passing modal security, and we could use our intuitions about such a case as evidence for or against Modal Security⁶⁵.

Summarising, CD&B give us no reason to think that we cannot construct examples that satisfy (i)-(iii) in order to establish whether or not the explanationist is correct, and they have given no good reason for thinking that, when we do so, the example will confirm that explanationism is false. The obvious next step, then, is to amend the discussed examples to construct a coherent case in which Modal Security is satisfied

⁶⁴ This implication is noted in Clarke Doane and Baras (2020: 166-169) as well, implying that Clarke Doane has not changed his position on this issue since he presented his minimalist response to evolutionary debunking arguments in (2015).

⁶⁵ Such a case would have to ensure that our moral beliefs could not easily have been different in whatever way that the evolutionary explanation of our moral beliefs ensures this. It would not matter that such a case is false or empirically unsupported; we are investigating the possibility of hypothetical examples that will illustrate whether an explanatory disconnect could defeat all by itself.

despite the absence of any explanatory connection. This example need be no more complicated than the situation that we find ourselves in, according to CD&B, if we accept that our moral beliefs can be given a complete evolutionary explanation. Our intuitions about such a case can then help us determine whether an explanatory disconnect can defeat our moral beliefs by itself, or whether a failure of Modal Security is required for our beliefs to be undermined.

Now, there is a difficulty involved in presenting examples of beliefs that uncontroversially satisfy Modal Security. Because CD&B give us no way of individuating methods for forming belief, nor any way of determining the class of propositions that are close enough to the actual belief to be relevant to our belief's safety, it is always possible that an example will be deemed unsafe because it is formed using an unreliable method or because the method could have produced false beliefs about propositions that are sufficiently close to the content of our actual belief. This makes it difficult to construct an example that uncontroversially satisfies condition (ii).

Luckily, we do have a case in which we know that our beliefs pass Modal Security: the case of our moral beliefs under the assumption they can be given a complete evolutionary explanation. We can therefore construct an example that satisfies Modal Security by reflecting the modal stability of our moral beliefs under the assumption that the debunker's evolutionary genealogy of those belief is accurate. The following example, which is based on Korman and Locke's Neora example, is designed with this goal in mind:

Neo:

Neo has a number of beliefs about which numbers are prime on the basis of clear and distinct intuitions. Later, Agent Smith convinces him that he is part of a computer simulation. He tells him that the designers found they could only bring him into existence by programming him to have these (and only these) intuitions about which numbers are prime. Without being programmed to have such intuitions, Neo could have never existed. Neo is also told that he could not have been programmed into existence using any program other than the one via which he was actually programmed into existence. Finally, he is told that the program requires imbuing him with these particular prime intuitions for reasons that have nothing to do with the relevant numbers actually being prime; they are just the only set of prime intuitions

that allow for life in the simulation in which he finds himself. Despite having now accepted all this, Neo doesn't abandon his prime beliefs.

I think that, in this case, Neo's beliefs pass Modal Security. The beliefs are about facts which, if true, are metaphysically necessary and therefore trivially sensitive. Furthermore, he could not easily have had false beliefs using the method he actually used, which is to trust his intuitions about what numbers are prime. He also could not easily have had any false beliefs about which numbers are prime using this method, because he could not easily have had intuitions that encouraged the belief that alternative (composite) numbers are prime. So the case is not like the Jan example, in which she could easily have used the same method and come up with false beliefs about relevantly similar propositions.

The modal stability of Neo's beliefs reflects the modal stability of our moral beliefs under the assumption that the debunker's evolutionary genealogy is accurate. Just as evolution could not have easily encouraged us to have dispositions to form wildly different moral beliefs, so could the programmers not have given Neo different intuitions about prime numbers. And just as we could not easily have been the product of anything other than evolution, neither could Neo have been the product of any program other than the one actually responsible for his existence. And facts about which numbers are prime are just as metaphysically necessary as moral facts. This implies that, if our moral beliefs pass Modal Security given the debunker's evolutionary genealogy of those beliefs, then Neo's beliefs also pass Modal Security.

I therefore think we now have what was left wanting in CD&B's discussion: an example that satisfies conditions (i)-(iii). Neo accepts there is no explanatory connection between his beliefs in prime numbers and prime numbers. But his belief passes Modal Security (or so I have argued above). Finally, I do not see anything *incoherent* about the above example; it seems perfectly coherent to imagine a world in which only one computer program is capable of programming consciousness, and that, in order to do so, it needs to imbue that consciousness with particular intuitions about prime numbers.

Is this the example *difficult* to imagine? I think it may be in some ways, but that this difficulty is irrelevant for our purposes. It might be difficult to imagine *why* only one computer program would be capable of creating life. Similarly, it might be difficult to

imagine why it can only do so by programming the agent to have a particular set of intuitions about prime numbers. But it is not difficult to imagine *that* either of these things is the case. And once we have understood that both these things are true in this scenario (a relatively easy task) we can start to examine our intuitions about the status of Neo's beliefs.

This too reflects the situation of evolutionary debunking arguments against moral realism. We do not need to understand why evolution encouraged moral beliefs with particular contents in order to generate robust intuitions about what would follow, philosophically, if this were the case. Very little of the philosophical discussion surrounding evolutionary debunking arguments turn on the intricacies of group selection. Similarly, it may be difficult to understand why we could not have arisen from any process other than evolution, but that is irrelevant when it comes to determining what would follow philosophically if this were true. *That* natural selection encouraged beliefs with particular contents, and *that* we could not easily have arisen out of any other process is easy to imagine. And this is all that we need to imagine in order to generate intuitions about what follows from the status of our beliefs in this case. Similarly, then, our intuitions about what would follow if the Neo scenario were true are not threatened by the fact that it may be difficult to imagine why it is true, just so long as we can (fairly easily) imagine that it is true.

And, in this case, I think that our intuitions are fairly robust. When Neo accepts that he only has his prime-beliefs as a result of a feature of the computer programme that (somewhat inevitably) made him, and that this feature had nothing to do with the truth of those prime beliefs, his prime beliefs become unjustified. He ought to withhold those prime beliefs because they have been defeated by this explanatory concession. This implies that an explanatory concession can defeat our belief even if we have no reason to think that the belief is either insensitive or unsafe. Explanatory concessions can undermine directly, and Modal Security is false.

Now, one lesson from Clarke Doane and Baras' discussion is that there are different kinds of modal stability. There is the modal stability had by those examples considered in the Jan case, and then there is the more robust modal security that is enjoyed by our moral beliefs given the debunker's story and by Neo's beliefs in the example above. It may therefore be possible for a persistent modalist to claim that there is some modal

constraint according to which our moral beliefs and Neo's prime beliefs *are* in fact defeated. This would allow them to insist that Neo's example is not an indication that explanatory concessions can undermine directly, because Neo's beliefs actually fail whatever (presumably extremely robust) modal conditions are specified by this constraint.

Obviously, though, this move is only possible if the modalist in question is prepared to accept that our moral beliefs are defeated by debunking arguments, because it involves arguing that Neo's similarly stable beliefs are defeated. If I were only looking to defend evolutionary debunking arguments against minimalist replies, this open possibility would not threaten me because it would involve conceding that debunking arguments succeed in undermining realistically construed moral beliefs. But, in addition to this goal, I am looking to defend explanationism independently of this dialectic, so it is worth pausing briefly to consider this possibility.

The first thing to note is that, without any sense of what this alternative modal constraint might look like, it is impossible for me to consider it. But if one finds such a constraint plausible, then I would suggest constructing an example that satisfies conditions (i)-(iii) in order to determine whether it, or explanationism, is true⁶⁶.

Though it would obviously be impossible for me to engage with every conceptually possible form of modalism, I think that the examples that I have considered imply that, when it comes to defeating beliefs, explanatory considerations are what really matters. We can fix the modal stability of the belief as tightly as we like, but as long as the explanatory relationship with truth is threatened, then so is the justification of the belief. I am therefore somewhat confident that, in an example in which (i)-(iii) are satisfied, the belief will be undermined and that will vindicate explanationism. But until any such actual constraint is offered, the best I can do is speculate. The ball is in the modalist's court if this is the line that they want to pursue.

Secondly, we can note that this strategy would involve endorsing a fairly stringent modal constraint on belief. It would have to imply that the modal stability of Neo's beliefs are not strong enough for those beliefs to be justified, and Neo's beliefs are pretty darn stable. As such, I think it would be difficult to formulate a constraint that

⁶⁶ Clause (ii) makes reference to Modal Security, and so this would have to be swapped out for an analogous condition that references this new, alternative modal constraint.

satisfies this condition without making that constraint implausibly strong, such that it falsely implies that actually justified beliefs are unjustified. This is a difficulty that will have to be faced by the modalist hoping to pursue this line of response.

I therefore conclude that the examples I have considered strongly support explanationism over Modal Security. In every case that I have considered, an explanatory concession undermines the belief even though the modal security of the belief remains unthreatened. As long the minimalist is correct in claiming that cases like these are coherent, then such cases support the contention that explanatory concessions can undermine directly, without implying anything about the modal insecurity of those beliefs.

3.6: Modal Security via Explanationism:

If my arguments in the previous section are sound then I will have successfully responded to CD&B's argument that the examples typically employed in support of explanationism over Modal Security do not succeed. But CD&B and provide two additional arguments that are supposed to threaten explanationism, and in the following two sections I respond to each of these arguments in turn.

(ii) The first such argument begins with the claim that a plausible reason why explanatory concessions typically defeat is that they imply our beliefs lack reliability (2020: 180). But one plausible way of analysing unreliability is in terms of a lack of modal security. If this is the correct account of why explanatory concessions undermine, then it implies that explanatory concessions only undermine to the extent that they imply that our beliefs are modally insecure.

I think the suggestion that explanatory concession undermine because they imply our beliefs lack reliability is somewhat plausible. However, I do not think that we can analyse a lack of reliability in terms of modal insecurity. As such, even if it is true that explanatory concessions defeat because they imply our beliefs lack reliability, this does not imply Modal Security in the way that CD&B suggest.

Consider Neo. He finds out that he believes that certain numbers are prime just because this was a pre-requisite for his existence, and this has nothing to do with the truth of those beliefs. It seems that, when he finds this out, he therefore has a good reason to think that his beliefs were not formed reliably; given the fact that his intuitions are in

no way explained by the facts, relying on those intuitions is probably not a reliable way of forming beliefs about which numbers are prime. Thus, it is plausible that Neo's explanatory concession gives him reason to think that his beliefs about which numbers are prime are unreliable.

But as we have seen, Neo has been given no reason to doubt the modal security of his beliefs. This implies that we can be given reason to doubt that our beliefs were reliably formed without being given reason to believe that they are modally insecure. So the explanationist can accept CD&B's claim that explanatory concessions defeat because they give us reason to doubt that our belief was formed reliably whilst denying that explanatory concessions undermine by giving us reason to doubt modal security; it seems we can have reason to doubt the reliability of our beliefs without having reason to doubt they are modally secure.

3.7: Defeat Without Explanatory Concessions:

(iii) Finally, CD&B provide the following example of a case in which the modal security of our belief is threatened even though there is an explanatory connection between our belief in P and P:

“One's (defeasibly justified) belief in any given logical truth will be insensitive and unsafe despite the fact that, as a logical truth, it is implied by every explanation at all (and so certainly by the explanation of your coming to believe it) if you formed it by consulting an authority who decided on its truth by flipping a coin, for instance.”
(2020: 181).

Here, then, we have an example in which someone consults a supposed authority about logical truths, and we find out that they have been deciding the answers by flipping a coin. In this case, then, we have reason to think our beliefs are modally insecure, and so our belief fails Modal Security. However, this explanation of our beliefs will ultimately presuppose the truth of our logical beliefs. This, CD&B suggest, implies that logical facts do feature in the best explanation of our logical beliefs, and so this information does not give us reason to make an explanatory concession. But our logical beliefs *would* be undermined by this information, which implies that when there is a modal concession but no explanatory concession, our belief is undermined.

The disagreement between explanationism and Modal Security is about whether an explanatory concession is *sufficient* for undermining a belief. According to the explanationist, an explanatory concession is sufficient for undermining a belief and, according to the modalist, it is insufficient (because in order to undermine it is necessary to show that our belief is modally insecure and an explanatory concession is not sufficient for showing this).

If it is successful, CD&B's example implies an explanatory concession is not *necessary* for undermining a belief, because a modal concession is sufficient for doing so and one need not make an explanatory concession when making a modal concession. But even if the example successfully establishes this point (more on this later) this would not entail that an explanatory concession is insufficient for undermining beliefs, or that showing our beliefs to be modally insecure is necessary for undermining our beliefs. The explanationist could accept both that accepting an explanatory disconnect is sufficient for undermining our beliefs directly, and that there are also *other* ways in which our beliefs get undermined (i.e, by showing that they are modally insecure). Once it is granted that explanatory concessions are sufficient (but not necessary) for undermining beliefs, this implies that our moral beliefs could be defeated by the debunker's genealogy of our moral beliefs and that the minimalist cannot rely on substantive moral claims in response to debunking arguments.

Still, it might be argued that CD&B's example is important because it supports a general view of undermining defeat according to which what matters are modal rather than explanatory connections. If, in cases of a modal concession without explanatory concession our belief is defeated, this implies that what is really essential to defeat is the modal security of our belief. And this provides indirect support for the view that explanatory considerations cannot defeat by themselves, independently of modal considerations.

And so, granting that an example of this kind could weaken the explanationist's position, can the explanationist respond? I think that there are two plausible responses available. The first is to bite the bullet and claim that, in the given scenario, our logical belief is *not* undermined when we find out that the so-called authority has been deciding what to tell us by flipping a coin.

I do not think that this response is as objectionable as it might first appear. Presumably there *would* be something amiss about giving up our logical beliefs when we find out that they were formed because of an authority flipping a coin, precisely because those logical beliefs have to be presupposed in order to form the explanation that is ultimately supposed to do the defeating. Having given up our logical beliefs we would have no reason to accept the explanation that supposedly defeated those beliefs, and therefore no reason to think our beliefs had been defeated. The position becomes self-defeating, and I think it is unclear that our logical beliefs would therefore be unjustified. This implies that, in cases where the modal security of our belief is threatened but we make no explanatory concession, our belief is not undermined, and supports a view of undermining defeat according to which what matters for defeat is explanatory, rather than modal, relations.

Secondly, I think it is open to the explanationist to argue that, in the case just described, there actually *is* an explanatory concession (or good reason for making one). Though logical facts are presupposed by the explanation, it is unclear whether, according to that explanation, the logical facts metaphysically determine our logical beliefs in the sense required to satisfy the kind of explanation that I have been working with. Presumably the behaviour of the coin is not determined by logical facts in anything like the way that I suggested reality may be determined by the logical facts in chapter two. And I think it is precisely the absence of this explanatory connection that generates the intuition in this case; if we took the behaviour of the coin to be in some way determined by the relevant logical truths, then finding out that the logical authority was consulting the coin would not threaten to undermine our logical beliefs.

I think this indicates that CD&B's example is flawed in two significant ways. It is not clearly an instance of a belief being undermined, and it is not clearly an instance of a modal concession without an explanatory concession. So even if we allow that such examples would provide indirect support for the claim that an explanatory concession is not sufficient for undermining our beliefs, CD&B's proposed scenario is not a successful instance of such an example.

3.8: Conclusion:

If what I have argued in this chapter is correct, then accepting an explanatory disconnect between our moral beliefs and the facts that make them true can defeat those beliefs regardless of whether we have reason to think that the beliefs in question are modally insecure. This does not yet imply that the explanatory disconnect implied by debunking arguments, and granted by minimalists, defeats our moral beliefs. There remains the open possibility that the specific explanatory disconnect argued for by the debunker does not defeat our moral beliefs, and in the next chapter I will argue that this is in fact the case. Evolutionary debunking arguments do not rule out the possibility that our moral beliefs are explained by some fact that *also* explains their truth, and this allows our moral beliefs to remain justified according to the explanatory constraint that I will defend in that chapter.

As such, the arguments in this chapter do not entail that the modal security response, outlined in chapter two, fails because it utilises defeated beliefs without arguments. This is because the arguments of this chapter do not imply that evolutionary debunking arguments defeat our realistically construed moral beliefs. However, the arguments of this chapter do imply that the modal security response is fundamentally misguided. It attempts to ensure justification for our moral beliefs by showing that they are modally secure but, as I have argued, establishing that our beliefs are modally secure does not establish that those beliefs are undefeated. So even though the realist, for all I have said so far, may not need to vindicate their moral beliefs (because evolutionary debunking arguments may not imply that those beliefs are defeated) the modal security response does fail in the sense that it does nothing to secure the justification for those beliefs either. In chapters 5 and 6, I will argue that modal security responses (and third-factor replies) do in fact fail because they crucially rely on defeated beliefs without argument, but even independently of this point we can see that the modal security response is on the wrong track. In order to establish that our moral beliefs remain undefeated, we need to show that we need not make an explanatory concession that defeats those beliefs, rather than showing that those beliefs are modally secure. In the next chapter, I try to establish the specific sort of attitude we can have towards the explanatory history of our belief such that the belief in question is defeated.

Chapter 4: Objections to explanationism and (IEC):

4.1: Introduction:

In the last chapter I argued in favour of explanationism over Modal Security. Accepting an explanatory disconnect between our belief in P and P can defeat that belief regardless of what it implies for the modal security of that belief. In this chapter, I defend an explanatory constraint that pins down the specific attitude we can have towards the explanatory history of our belief that defeats that belief.

I begin by considering an obvious first attempt at an explanatory constraint, which I point out is subject to an important kind of counterexample. This prompts me to endorse my own “inclusive” explanatory constraint (IEC), according to which our belief is defeated if we withhold belief that it is explained by the fact that makes it true *and* we withhold belief that there is some fact that explains both our belief and the fact that makes it true. (IEC) avoids counterexamples that faced the less sophisticated constraint, and is also consistent with our intuitive judgements about defeat. I then argue that (IEC) can successfully account for justified beliefs about the future, which some commentators have suggested is a challenge for explanationism. Finally, I consider and reject a potential objection to (IEC) raised by Korman and Locke (2020).

I then relate (IEC) back to the debate between evolutionary debunking arguments and minimalist replies. I argue that (IEC) does *not* imply that the minimalist’s realistically construed moral beliefs are defeated by their acceptance of the explanatory disconnect implied by debunking arguments.

This leads me to consider various alternative explanatory constraints that have been offered by commentators who argue that the minimalist’s beliefs are in fact defeated by their acceptance of this explanatory disconnect. I argue that these constraints are subject to counter-examples because they mistakenly attribute epistemic significance to a feature of our beliefs that is actually epistemically irrelevant. This reveals a further asset of (IEC): it does *not* mistakenly attribute significance to this feature of our

beliefs, and thereby avoids the counterexamples that sink these alternative constraints.

If I am successful, then by the end of this chapter I will have defended (IEC) as a legitimate explanatory constraint on belief. I will also have shown that EDAs, as they are traditionally construed, do not imply that minimalist responses fail according (IEC). I conclude by noting that this does not necessarily imply that minimalist responses are successful. It is possible that traditionally construed evolutionary debunking arguments can be supplemented with additional arguments that show that minimalist responses *do* fail given (IEC). Presenting and defending these additional arguments will occupy chapters five and six of this thesis.

4.2: Tf-connected Beliefs:

In general, it seems as though accepting that our belief is not explained by the fact that makes it true renders that belief unjustified. For example, suppose I have a visual experience of leaves on my bed and thereby form the belief that my bed is covered with leaves⁶⁷. Later I find out that I was actually hallucinating when I “saw” the leaves on the bed. I therefore accept that my belief in the leaves on the bed is not explained by there actually being leaves on the bed. In this case, I can no longer continue to justifiably hold my belief in the leaves.

Cases such as these might motivate the following explanatory constraint on belief:

(EC): If S believes that her belief that P is not explained by the fact that P, then S’s belief in P is unjustified⁶⁸.

But (EC) is subject to important counter-examples. Beliefs can be justified when we take our belief to be explained by some fact that *also* explains the fact that makes it true, even if we do not take our belief to be explained by the fact that makes it true.

Here is an example of such a case⁶⁹:

Sarah’s Fireplace:

⁶⁷ This example is taken from Barry Stroud (2000: 54-5).

⁶⁸ This is similar to a constraint considered in Korman and Locke (2020). Korman and Locke do not endorse this constraint; they use it, as I do, in order to assess its flaws with hopes of articulating a more plausible constraint.

⁶⁹ This is adapted from an example given in (Goldman, 1967: 366).

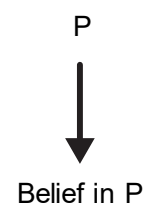
Sarah sees a fire in her fireplace and forms the belief that there is smoke coming out of her chimney. She forms this belief on the basis of her belief that her fireplace is lit in addition to her belief that active fireplaces cause smoke to come out of chimneys.

A philosopher then informs her that her belief is not explained by the fact that smoke is coming out of the chimney. She has the belief because of the fire in her fireplace. The smoke coming out of the chimney does not explain the presence of the fire, and therefore does not explain her belief that there is smoke coming out of the chimney. Thus, her belief is not explained by the fact that makes it true.

Sarah accepts the philosopher's point that her belief in the smoke is not explained by the fact that makes it true, but continues to believe that there is smoke coming out of the chimney anyway on the basis of her knowledge that there is a fire in her fireplace, and that fires cause smoke.

In this case, then, Sarah accepts that her belief in P is explained by some third-factor (the fire) that explains both her belief in the smoke and the smoke, and she accepts that the smoke does not explain her belief in the smoke. It seems clear that Sarah, having accepted as much, remains justified in believing that the smoke is emerging from her chimney. This implies that we need not believe that our beliefs are explained by the facts that make them true in order for those beliefs to be justified and (EC) is false.

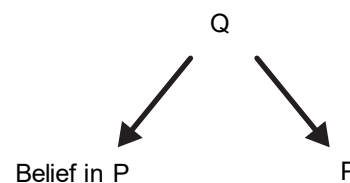
Figure 3:



Let us say that when our belief in P is explained by P, our belief is “e-connected”⁷⁰. The explanatory structure for e-connected beliefs can be represented in figure 3. When our belief in P is explained by some fact that also explains P, we can say our belief is “tf-connected” (for third-factor). The explanatory structure of tf-connected beliefs is represented in figure 4. The correct explanatory constraint must allow for tf-connected beliefs in order to be plausible.

The (EC) advocate could deny that e-connected beliefs and tf-connected beliefs are genuinely distinct. Maybe, when our belief in P is explained by some fact Q which also explains P, this is actually an instance of P explaining our belief, and so beliefs of this kind can be

Figure 4:



⁷⁰ Again, the term “e-connected” is taken from Korman and Locke (2020).

justified according to (EC). I think that this response fails, and showing why it does will be useful in getting clear the distinction between e-connected beliefs and tf-connected beliefs going forward.

It is generally accepted that events explained by a common cause do not explain each other. Hausman uses the following example to illustrate this point:

“drops in barometer readings do not explain the onset of storms, for such drops are effects of the atmospheric phenomena that cause storms, not them selves causes.” (1993: 228).

There is no sense in which the drops in a barometer reading explain the storm, and this is because the drops in the barometer reading are caused by some fact that also causes the storm (rather than, say causing the storm itself).

In the example of a tf-connection used above, Sarah’s belief in the smoke shared a common cause with the smoke itself. The common assumption that effects of a common cause do not explain each other therefore that implies that her belief in the smoke does not explain the smoke and vice versa. In other words, Sarah’s tf-connected belief in the smoke is not e-connected.

In order to see this more clearly, consider again Sarah’s belief in the smoke emerging from her chimney which she forms on the basis of her belief in the fire. Given that her belief in the smoke is explained by the fire, and the fire explains the smoke, her belief is tf-connected. If tf-connected beliefs are actually just a particular kind of e-connected belief, then it must be the case that, in addition to the fire, the smoke also explains her belief.

But it is hard to see *how* the smoke might explain her belief. She does not see the smoke or have any other kind of “direct” access to it. The smoke therefore does not explain her belief in the smoke by, say, her seeing it and forming a belief in the smoke. Furthermore, the smoke does not explain the fire, and is instead explained *by* the fire. It is therefore not the case that the smoke explains the fire which explains her belief, which would imply that the smoke ultimately explained her belief⁷¹. This implies that

⁷¹ This is clearly true for the current example, and is necessarily true for all tf-connected beliefs if explanations are asymmetric.

it is not in virtue of its explanatory relationship with the fire that the smoke explains her belief.

It seems that the only remaining possibility is that the smoke might explain some other fact that goes some way towards explaining her belief. But it is hard to see what this could be, and even harder to see why this is necessarily the case for every possible tf-connected belief. Maybe the smoke emerging from the chimney is what, unbeknownst to Sarah, drew her into the room and explained her looking at the fire in the first place. This would mean that, in addition to the fire, the smoke does in fact explain her belief in the smoke⁷². But it seems equally possible (and in fact more likely) that no such connection holds. As such, it seems perfectly possible to describe the case such that her belief in the smoke is tf-connected but not e-connected, thereby vindicating the possibility of tf-connected beliefs that are not e-connected.

As Korman and Locke emphasise, *other* instances of smoke that she has observed in the past may well go some way to explaining her belief (2020). But her belief is about the smoke currently emerging from the chimney, rather than past instances of smoke coming out of chimneys. As a result, this observation does not imply that tf-connected beliefs need also be e-connected, because it does not imply that Sarah's tf-connected belief is necessarily explained by the fact that makes it true (as is required for e-connection). It therefore seems that the most plausible interpretation Sarah's belief in the smoke is one according to which it is tf-connected but *not* e-connected.

This implies that e-connection and tf-connection represent two distinct ways in which our belief might have some kind of explanatory relationship with the fact that makes it true. And, as the example we have been discussing indicates, accepting that our belief is not e-connected because it is instead tf-connected can allow that belief to remain justified. Realising that her belief in the smoke was not explained by the smoke did not render Sarah's belief unjustified because she did not also deny that it was explained some fact that did explain the smoke. (EC) is therefore flawed, and we need

⁷² This possibility depends on how we chose to individuate facts. If the smoke emerging from the chimney at t1 is what draws her into the room, and she forms a belief about smoke emerging from the chimney at t2, and these two facts are distinct, then this will not be a case of her belief being explained by the fact that makes it true. Still, it seems that with enough inventiveness one could come up with an alternative example that doesn't have this drawback. If not, this further supports my contention that tf-connected beliefs need not be e-connected.

to amend it in order to account for the possibility of justified beliefs that we take to be tf-connected.

4.3 An Inclusive Explanatory Constraint:

I endorse the following explanatory constraint on belief:

Inclusive explanatory constraint (IEC): If (i) S withholds belief that her belief that P is explained by the fact that P and (ii) S withholds belief that there is some fact that explains both her belief that P and the fact that P, then S's belief that P is unjustified.⁷³⁷⁴

“Withholding” a belief can be understood as considering a proposition and opting not to believe it, either by believing that it is not true or by remaining agnostic (Korman 2019b: 357). (IEC) can avoid the counter-examples faced by (EC), which involve justified beliefs that we took to be tf-connected but not e-connected. Because (IEC) stipulates that we need to withhold belief that our belief in P is e-connected *and* tf-connected in order for that belief to be defeated, it does not imply that beliefs we take to be tf-connected but not e-connected are unjustified. It can therefore account for the possibility of justified beliefs we take to be tf-connected but not e-connected.

Now, one might have the following worry about (IEC). We are presented with counter-examples to (EC) in the form of justified beliefs that we take to be tf-connected rather than e-connected, and in response we amend (EC) by adding a clause that allows

⁷³ Some explanatory constraints allow beliefs to remain justified if we take our belief in P to explain the fact that P (see, for example, Korman (2019b)). (IEC) does not include such a clause, because, for reasons I lack the space to get into, I doubt that such beliefs can be justified. Fortunately, this possibility is irrelevant when applied to evolutionary debunking arguments. Anti-realist meta-ethical positions take moral facts to be explained by our evaluative attitudes, but they invariably take the evaluative attitudes that explain the moral facts to be states *other* than beliefs (see, for example, Street (2008) & (2010)). As a result, moral beliefs themselves do not explain the moral facts even according to these positions, and so the possibility of our belief in P explaining P is irrelevant even if we take evolutionary debunking arguments to ultimately support some form of meta-ethical anti-realism. According to these positions our moral beliefs, when they are justified, can be explained by the other evaluative attitudes that determine moral truth, and the relationship between moral beliefs and the moral facts thereby resembles something like a tf-connection, as is argued for by Selim Berker (2014). As such, even if our beliefs can be justified if they explain the facts that make them true, nothing relevant to the purposes of this thesis depends on it.

⁷⁴ In chapter one I said that evolutionary debunking arguments attempt to provide an undercutting defeater for our moral beliefs. Beliefs that fail (IEC) are not necessarily defeated, because they may have never been justified to begin with. Still, in the current context I describe beliefs that fail (IEC) as being defeated because the evolutionary debunker is granting that, independently of evolutionary debunking arguments, our realistically construed moral beliefs are *prima facie* justified.

beliefs to remain justified if we take them to be tf-connected. This move might strike some as too easy or *ad hoc*. What is the relationship between taking our belief to be e-connected and taking it to be tf-connected? If there is no independent reason for thinking the two might be related, then the addition of clause (ii) in (IEC) seems like an unrelated patchwork addition to (EC) designed entirely to deal with these specific counterexamples.

In order to see why this is not a problem, consider Faraci's account of epistemic coincidence:

“It is a coincidence that the members of Γ are all true if (a) there is no unified explanation for the members of Γ and (b) no explanatory chain links all members of Γ .” (2019: 7)

Faraci's previous formulations of this principle omitted (b), and Faraci adds (b) after considering another possible way in which coincidence might be avoided. On this amendment, Faraci notes:

“The addition of (b) is well-motivated, because (b) is clearly recommended by the same underlying thought as (a). They represent the two ways in which there can be an explanatory connection between the members of a set: (a) through a shared relation to something outside the set or (b) through internal relations within it”. (2019: 7)

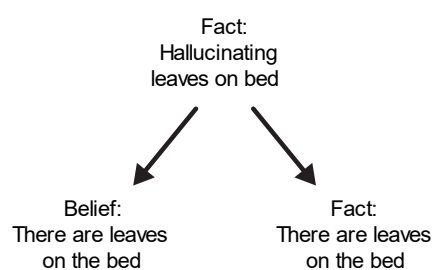
I think similar reasoning applies to the addition of the second clause of (IEC). The addition of clause (ii) represents the same underlying explanationist thought that (EC) was attempting to capture: if we take our belief to be explanatorily disconnected from P, then that belief is defeated. It is well-motivated because a tf-connection represents the *other* way in which our belief might be explanatorily connected to the fact that makes it true⁷⁵. Given our view that defeat is primarily about the explanatory relationship between our belief and the relevant fact, we should expect to expand our explanatory constraint when we are confronted with another possible way in which the two might be explanatorily related. I therefore do not think that we should view a constraint like (IEC) as *ad hoc*.

⁷⁵ In the chapter 6, I consider whether there are any further kinds of explanatory relationships and whether taking our beliefs to be related to the facts via these connections might allow those beliefs to remain undefeated.

I think that (IEC) can be further supported by considering the examples that typically motivate explanationism and that we initially took to imply (EC). We have a visual experience of leaves on our bed and then find out that we were hallucinating. The information that we were hallucinating rules out the possibility that our belief in the leaves is e-connected. But if we amend the case such that we do not withhold belief in there being a tf-connection between our belief in the leaves and the leaves, then it seems our belief in the leaves can remain undefeated. If we believe, for example, that our hallucinating leaves on the bed somehow actually explains the fact that there are leaves on the bed, then finding out that we are hallucinating does not defeat our belief in leaves on the bed.

Imagine we believe that a (fairly benign) demon has access to our mental states, and whenever we are hallucinating he endeavours to re-create the hallucinated scene in the actual world. Upon finding out that we are hallucinating, we would have reason to believe that our belief in the

Figure 5:



leaves is tf-connected in the way outlined in figure 5. And it seems as though, with these stipulations in place, the information that we are hallucinating (and that our belief therefore cannot be e-connected) does not defeat our belief. It seems we are justified in continuing to hold our belief in the leaves because we now believe in something that we take to explain the leaves being on the bed.

I think the same goes for all cases in which information that our belief is not e-connected defeats that belief⁷⁶. If we specify the details of the case such that the agent takes their belief to be tf-connected but not e-connected, then we get the intuition that their belief remains justified.

The reason that these examples might initially seem to support (EC) is just because they are typically examples in which it is natural to assume that the agent already withholds belief in (or has good reason to withhold belief in) a tf-connection between their belief in P and P, prior to whatever information prompts them to withhold belief

⁷⁶ This includes those examples in the previous chapter in which a modal disconnect does not threaten to disrupt the modal stability of those beliefs. The cases are described such that it is natural to assume the agent in questions withholds (or has good reason to withhold) belief that their belief is tf-connected, as well as e-connected. If we amend the examples such that this is not the case, then our intuitions about whether the beliefs are justified changes.

in its being e-connected. So their belief is defeated when they accept that their belief is not e-connected, but this does not show that their belief is defeated just if they withhold belief that it is e-connected; it is defeated because they withhold belief in their belief being e-connected *and* they withhold belief in it being tf-connected. This just draws our attention to the fact that, if we already withhold belief that our belief is tf-connected, all we need to do to achieve the state of withholding belief that our belief is *both* e-connected and tf-connected is to start withholding belief in its being e-connected. And it is withholding belief in both of these possible explanatory connections between our belief in P and P that defeats our belief.

4.4: Future Beliefs:

A possible objection to explanationist accounts of defeat relates to the possibility of justified beliefs about the future. Obviously we have many justified beliefs about the future, and a plausible account of defeat cannot imply that all our beliefs about the future are unjustified. Can explanationism account for justified beliefs about the future?

Various commentators have thought not. Consider an example from Byerly⁷⁷:

“Suppose I'm on the golf course on a sunny, calm day. My putting stroke has been working for me most of the day, and I'm now on the sixteenth green. It's not a long putt—just six feet. I'm fairly confident. I rotate my shoulders, pulling the putter back, and then accelerate through the ball. It rolls toward the cup. The speed looks good. The line looks on. Yes, I believe it's going in!” (2013: 235).

I see no reason to think that Byerly's belief that the ball will go in the hole is unjustified. But, as Byerly points out,

“Surely the ball's rolling into the cup at some later time doesn't explain why right now I have the evidence that I do. My current evidence consists in my memories of my puts earlier in the day, my visual awareness of the ball's rolling, my sense of the temperature and lack of wind, and the like. What explains this evidence is a body of current and perhaps past propositions.” (2013: 235-6).

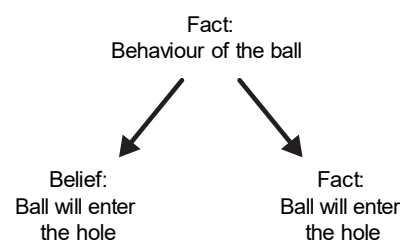
⁷⁷ Byerly's example is actually presented as a counterexample to explanationism about *justification* rather than about defeat. But if it succeeds against explanationism about justification it also succeeds against explanationism about defeat, so I assume Byerly would also take his example to be relevant to explanationism about defeat.

It is therefore unlikely that Byerly's belief that the ball will enter the hole is explained by the future fact that the ball will enter the hole. It is instead explained by previous instances of the ball rolling in the hole, and the present conditions that seem to imply it will happen again. This implies that Byerly's belief that the ball will go in the hole is not e-connected. Generally speaking, it is difficult to see how any of our beliefs about the future could be explained the relevant future facts, rather than by an array of present and past facts. But it seems like we can accept this point for any one of our future beliefs without our future beliefs thereby becoming unjustified.

I think that (IEC) can account for the fact that our belief that the ball will roll into the hole is undefeated. We may doubt that it is e-connected, but unless we also withhold belief that it is tf-connected we have no reason to think that (IEC) implies that it is unjustified. So could Byerly's belief be tf-connected?

According to Byerly, his belief that the ball will enter the hole is based on the visual experience of the ball rolling at a certain speed in a certain direction, and his awareness of a lack of wind and other interfering factors. Furthermore, there is no

Figure 6:



reason to think that Byerly's experience of the ball rolling at that speed in that direction cannot be explained by the fact that the ball *is* rolling at that speed in that direction. Byerly's awareness of a lack of wind and other interfering factors may also be explained by the lack of wind and other interfering factors. Finally, it is highly plausible that the ball's behaviour and the lack of wind might explain the fact that the ball will enter the hole. If so, then the explanatory structure outlined in figure 6 can obtain for Byerly's belief, and Byerly's belief can be tf-connected. Generally speaking, even if our beliefs about the future are explained entirely by past and present factors, there is nothing to stop these factors from explaining the fact that the future event will occur, and so nothing to stop our beliefs about future events being tf-connected. Because (IEC) implies that we have to withhold belief in our belief being e-connected and tf-connected, we have no reason to think that (IEC) implies our beliefs about the future cannot be justified.

What (IEC) *does* imply is that, if Byerly were to withhold belief in his future belief being tf-connected his future belief would be defeated. How plausible is this?

It is plausible that, in order for these his visual experience of the ball's behaviour and his awareness of an absence of intervening factors to justify his belief, Byerly cannot deny that these mental states are explained by the facts that they represent. Byerly would not be justified in basing his belief that the ball will go in the hole on his visual representation of the ball travelling at a certain speed and direction if he took this visual experience to be the result of a hallucination, rather than to be in response to the ball genuinely travelling at that speed and in that direction.

Furthermore, it is plausible that, in order for his belief to be justified, Byerly could not deny that the fact that the ball is rolling at the specific speed and at the specific direction explains the fact that the ball will go in the hole. If he takes the ball rolling at that speed to play no role in determining whether or not the ball will enter the hole then Byerly's visual representation of the ball travelling at that speed seems not to constitute a good basis for his belief that the ball will enter the hole.

This supports the view that, in order for Byerly's belief that the ball will go in the hole to be justified, he cannot deny that his belief is explained by factors that *also* explain the truth of that belief. In other words, Byerly cannot deny that his belief is tf-connected, which (given that it cannot be e-connected) is exactly what is implied by (IEC). (IEC) can therefore account for our intuitive judgements regarding justified beliefs about the future, and this counts in favour of (IEC)⁷⁸.

4.5: Not Necessary For Defeat?

Korman and Locke argue against the following explanatory constraint on belief:

(EC**): If S believes that (i) her belief that p neither explains nor is explained by the fact that p and (ii) there is no single fact that (at least partially) explains both her belief that p and the fact that p, then S is thereby rationally committed to withholding belief that p.

The difference between (EC**) and (IEC) is that (EC**) stipulates that if we *believe* that our belief in P is neither e-connected nor tf-connected then we ought to withhold

⁷⁸ For reasons of space, I do not consider how (IEC) might allow justification for other areas of belief that might seem problematic for explanationism. For a discussion of these cases, see McCain (2017). McCain's arguments imply that the best explanation of our justified inductive beliefs and mathematical beliefs can explain their truth. This implies that these beliefs can be tf-connected, and the possibility of justified inductive and mathematical beliefs is therefore consistent with (IEC).

belief in P. By contrast, (IEC) stipulates that if we *withhold* belief that our belief in P is e-connected or tf-connected to P then our belief in P is unjustified. Because one way of withholding belief in P is to disbelieve P, if one disbelieves that one's belief in P is e-connected or tf-connected to its truth, one will also withhold belief that one's belief is e-connected or tf-connected to its truth. This means that if (IEC) is true then so is (EC**) ⁷⁹. Given their close relationship, it is worth considering whether (IEC) can avoid the following objection that Korman and Locke's raise against (EC**).

Consider the following scenario:

Martian:

Shaquira believes that

B1: protons cause streaks when they pass through cloud chambers.

She believes this because Martians have told her this was the case. She then finds out that the Martians told her this not because it is true, but because they like the sound of the word "proton". Despite this, she continues to believe that streaks in cloud chambers are caused by protons passing through them.

She is then informed by a reliable source that a proton has been fired through a cloud chamber in the next room that she cannot see. Using her belief that protons cause streaks in cloud chambers, she forms the belief

B2: there is a streak in the chamber.

(EC**) fails, according to Korman and Locke, because it fails to account for the irrationality of (B2) (2020: 326). Shaquira can reason her way to a belief that there is a third factor (the proton being fired) that explains both her belief in the streak and the streak itself. According to (EC**), there is no requirement to withhold this belief. But clearly (given her acceptance of the genealogy of the Martian's information) Shaquira ought to withhold this belief. As (EC**) fails to correctly imply as much, (EC**) must be false. What's more, Shaquira does not *withhold* belief that her B2 belief is TF-connected to its truth, and so this objection applies equally to both (IEC) and (EC**).

⁷⁹ This assumes that we should withhold beliefs that are unjustified.

One problem with this objection is that (EC**) presents a necessary condition on rational belief. It says that, in order for our belief to be rational, we cannot have certain attitudes towards that belief. Korman and Locke then respond by presenting an instance of irrational belief that is not deemed irrational by (EC**). But this is perfectly consistent with the truth of (EC**), because (EC**) makes no claims to provide a sufficient condition on rational belief that would imply *all* beliefs which pass (EC**) are thereby rational. A similar point can be made for (IEC), which also attempts to provide a necessary, and not a sufficient, condition for justified belief.

To support this point, imagine that we are inclined to accept the following constraint on belief:

Hallucination: if we withhold belief that our perceptual belief in P is not the result of a hallucination, our belief in P is unjustified.

The truth of this principle cannot explain why Shaquira's (B2) belief is unjustified; B2 is not a perceptual belief that she thinks might be the result of a hallucination. But surely the failure of this principle to account for the fact that Shaquira's (B2) belief is unjustified does not give us reason to reject that principle, because it does not imply that all beliefs that we accept are *not* the result of a hallucination are thereby justified. Similarly, the fact that (EC**) and (IEC) cannot account for the fact that Shaquira's belief is unjustified and ought to be withheld does not imply that they are false, because they do not claim that all beliefs that we accept are e-connected or TF-connected are automatically rational or justified. Thus, the proponent of either (EC**) or (IEC) is not obliged to give an account of how their constraint correctly diagnoses Shaquira's (B2) belief as irrational or unjustified⁸⁰. Korman and Locke's objection to (EC**) therefore gives us no reason to reject either (EC**) or the closely related (IEC).

But, as it happens, I think (IEC) does play a role in explaining why Shaquira's (B2) belief is unjustified. I think that (IEC) entails that Shaquira's (B1) belief is unjustified and, as Shaquira's (B2) belief is based on (B1), this implies Shaquira's (B2) belief is unjustified as well.

⁸⁰ For these reasons, I do not discuss another worry that (IEC) would fail to deem certain cases of defeated beliefs defeated that is found in Byerly and Martin (2015). For an account of how (IEC) could in fact deem such beliefs defeated, see McCain (2017).

(B1) is the belief that protons cause streaks when they pass through cloud chambers. This belief is unjustified, according to (IEC), if Shaquira withholds belief that this belief is e-connected and Shaquira withholds belief that it is tf-connected. Presumably, Shaquira would withhold both these beliefs. She accepts that her belief is fully explained by the Martian's testimony, and that this has nothing to do with protons. She therefore withholds belief that this belief is e-connected.

Because her belief is fully explained by the Martian's testimony, in order for this belief to be tf-connected, the Martian's testimony would have to explain the fact that protons cause streaks in cloud chambers. We can stipulate that she does withhold belief that whatever the Martians tell her somehow manifests itself in reality, and she thereby withholds belief that her belief (B1) belief is tf-connected.

If so, Shaquira's belief (B1) that protons cause streaks in cloud chambers is unjustified according to (IEC). This is because she withholds belief that it is either e-connected or tf-connected. Furthermore, Shaquira's (B2) belief is based on (B1). As (B1) is unjustified, so is (B2). (IEC) therefore plays some role in explaining why (B2) is unjustified, by entailing that the belief on which (B2) is based fails (IEC) and is thereby unjustified. Under the assumption that we cannot justifiably rely on unjustified beliefs, this implies that Shaquira's (B2) belief is unjustified.

All this is consistent with the claim that (B2) itself passes (IEC). (B2) is unjustified because a belief on which it is based fails (IEC), and it is there based on an unjustified belief. (B2) is therefore unjustified even though Shaquira takes it to be tf-connected, which entails it satisfies (IEC). Fortunately, because (IEC) does not attempt to provide sufficient condition for justified belief, it is perfectly consistent with the claim that some unjustified beliefs will satisfy its requirements.⁸¹ The claim that (B2) is both unjustified and passes (IEC) therefore presents no challenge to the proponent of (IEC).

Given their closeness, it is worth saying something about why I endorse (IEC) over (EC**). Recall that the main difference between these two principles is about the kind of attitude that we might take towards a given belief being e-connected or tf-connected

⁸¹ Because (IEC) does not attempt to provide a necessary condition on defeat, it is also consistent with the position that our belief is defeated by our having good reason to withhold belief that our belief in P is e-connected or tf-connected, regardless of whether or not we actually adopt this attitude. I do not explicitly consider this possibility because minimalist replies explicitly accept the explanatory disconnect implied by debunking arguments.

that could render that belief unjustified or imply that we ought to withhold that belief⁸². According to (EC*), disbelieving in any such connection is required in order for it to be the case that we ought to withhold our belief in P. By contrast, (IEC) stipulates that if we disbelieve in such a connection or *even* consider but remain agnostic about it, our belief in P is unjustified. (IEC) is therefore stronger and (EC**), because it does not allow a belief to remain justified if we remain agnostic about whether it is e-connected or tf-connected.

In order to see why I have opted for the stronger (IEC), consider the following case. If we find out that we are only having the visual experience of the cup on a table because we are dreaming, then it seems we ought to withhold our belief that there is a cup on the table. But it also seems that, if we are genuinely *uncertain* about whether or not we are dreaming, then our belief in the cup also becomes unjustified.

It seems odd to accept that our visual experience of the cup *may* be the result of a dream then to continue to affirm our belief in the cup on the table on the basis of our visual experience of that cup. If we have considered the possibility that our visual experience is explained by our dreaming, and we cannot bring ourselves to believe either that we are or we are not dreaming, then it seems we cannot justifiably continue to hold our belief in the cup basis of our visual experience; we ought to withhold this belief in the cup⁸³. This implies that we should accept (IEC) over (EC**). If we *withhold* belief that our belief in P has the appropriate explanatory relationship with P, our belief in P is unjustified.

Let me summarise the reasons that I have given in favour of (IEC) so far. In the previous chapter I argued against Modal Security and in favour of explanationism about defeat. At the beginning of this chapter I offered an obvious first attempt at an explanatory constraint, but argued it was subject to an important class of counter-examples. I then offered my own explanatory constraint on belief, which involves amending (EC) to allow beliefs to be justified if we take them to be tf-connected. This amendment allows (IEC) to avoid the counter-examples faced by (EC), and I also indicated why (IEC) might be motivated independently of such counter-examples. I

⁸² A less significant difference between them is that (EC**) stipulates that the resulting belief should be withheld, while (IEC) stipulates that the relevant belief is unjustified. Obviously there is an intimate connection between these notions, and, for reasons of space, I therefore focus on the more significant difference between the two.

⁸³ Bergmann (2006: 426-7) presents a similar argument.

then noted that (IEC) can account for our intuitions regarding those cases that might, at first glance, be taken to support (EC). I then argued that (IEC) can account for justified beliefs about the future, and, finally, I considered and rejected a potential objection to (IEC), and justified my endorsement of it over the closely related (EC**).

In the next section, I will pause my case for (IEC) to consider how it relates to the debate between evolutionary debunking arguments and minimalist replies. This will naturally lead me to consider some other, alternative explanatory constraints that have been offered in the literature, and my critique of these replies will shed further light on the plausibility of (IEC).

4.6: (IEC) and Minimalist Replies:

As I outlined in chapter two, minimalist replies to debunking arguments cannot work if the explanatory concession implied by debunking arguments defeats our realistically construed moral beliefs. Does (IEC) imply that this is the case?

I think not. EDA's work by arguing that the complete explanation of our moral beliefs nowhere includes the moral facts. If this is true, then it successfully rules out the possibility that our moral beliefs are e-connected, because our beliefs being e-connected requires the facts that makes them true explaining (and therefore appearing in the explanatory history of) our belief. But this does not rule out the possibility that they are tf-connected. For our belief to be tf-connected, one of the facts in the explanatory history of our beliefs must (in addition to explaining our moral beliefs) also explain the fact that would make those beliefs true. Establishing that the complete explanatory history of our beliefs does not include the moral facts cannot by itself rule out the possibility that some of the facts that do appear in this explanatory history might also explain those moral facts. It is exactly this possibility that the third-factor replier exploits when they rely on moral claims that show how our moral beliefs might be explained by some fact that also explains the truth of their beliefs *even given* the debunker's evolutionary history of our moral beliefs.

Debunking arguments therefore only establish that, under the assumption of moral realism, our moral beliefs are not e-connected. They leave open the possibility that they might be tf-connected. Given (IEC), the minimalist's moral beliefs therefore

remain undefeated by such arguments, because they need not withhold belief in their moral beliefs being tf-connected.

In the next section, I consider alternatives to (IEC) that have been used to argue that evolutionary debunking arguments do in fact defeat our realistically construed moral beliefs. I focus on these constraints specifically because I think they all share a flaw that is not shared by (IEC), and this will provide additional support for (IEC) as a legitimate constraint on belief⁸⁴.

4.7: Alternative Explanatory Constraints:

(i) Locke (2014) endorses the following explanatory constraint on belief:

“Cognition Defeat: Where D is one of S’s cognitive dispositions to form beliefs involving the concept X, if S withholds belief that the explanatory history of D involves X-facts, then S is not default entitled to believe that D is reliable.” (2014: 232).

According to Cognition Defeat, we cannot withhold belief that the explanation of our disposition to form moral beliefs involves moral facts and retain our default entitlement to those beliefs. If debunking arguments establish that the explanation of our disposition to form moral beliefs does not involve moral facts, then minimalist replies cannot rely on such beliefs in response to EDA’s because those beliefs have been defeated.

But what does Locke mean by the explanatory history of one’s dispositions “involving” a given set of facts? Does this require that our dispositions be e-connected to the facts, or are the moral facts also appropriately “involved” in the explanatory history of those dispositions if the two are tf-connected?

Locke describes Enoch (and other proponents of third-factor replies) as defending the view that “it doesn’t matter whether you believe that normative facts were involved in the explanatory history of your normative dispositions, provided that you believe that

⁸⁴ I therefore do not consider Lutz’s (2018) explanatory constraint which is employed for this same purpose on the grounds that I do not think that it has the flaw shared by the other constraints that I will consider. See also Joyce (2016) for an argument that seems to imply the need for some sort of explanatory constraint on belief.

there is *some* explanation of why you would come to have reliable normative dispositions” (2014: 228). This implies that, according to Enoch’s third-factor reply, the moral facts are *not* “involved” in the explanatory history of our moral beliefs. We should therefore interpret Cognition Defeat as requiring that we do not withhold belief that our disposition to form moral beliefs are e-connected with the moral facts.

With this constraint in mind the argument against third-factor replies makes itself available. Locke construes evolutionary debunking arguments as proceeding from empirical claims which naturally imply that “normative facts were nowhere involved in the explanatory history of our normative dispositions” (2014: 223).

Our moral beliefs therefore fail Cognition Defeat. For the minimalist, this means that her “default entitlement to rely on her normative dispositions [...] were all defeated the moment they stopped believing that those dispositions had explanatory histories that involved the relevant kind of facts” (2014: 232). As their default entitlement to such beliefs have been defeated, they cannot be relied on in response to evolutionary debunking arguments. As minimalist replies rely crucially on our moral dispositions, such replies fail.

(ii) Korman (2019b) uses a similar strategy, but employs the following, alternative explanatory constraint:

“(EC5) S is justified in believing p only if: either (i) S doesn’t withhold belief that [S’s belief that p is e-connected to p], or (ii) S’s belief that p is based on beliefs Bp1 . . . Bpn that prima facie justify S in believing p and S doesn’t withhold belief that [S’s beliefs that p1 . . . pn are e-connected to p1 . . . pn].”
(Korman 2019b: 357)

(EC5) therefore states that, in order to be justified in believing P, we cannot withhold belief that our belief in P is e-connected, or inferred on the basis of other beliefs that are e-connected.

Like Locke, Korman characterises evolutionary debunking arguments as proceeding from empirical claims that expose “an explanatory disconnect between the reactions and the associated facts” (2019b: 339). More precisely, the empirical basis of debunking arguments establishes that our moral beliefs are not e-connected, and

presumably we can assume that they are not inferred on the basis of non-moral beliefs that are e-connected. Thus, evolutionary debunking arguments defeat our moral beliefs by showing that they fail (EC5). Once one concedes this much, Korman claims that “you lose any entitlement you had for relying on those faculties [...] and bootstrapping from them is no longer an option” (2019b: 353).

Thus, by showing that our moral beliefs are not e-connected (or inferred on the basis of beliefs that are e-connected), evolutionary debunking arguments defeat the realist’s moral beliefs. As these beliefs are defeated, they cannot be assumed in response to EDAs. Minimalist replies, which utilise such moral claims in order to argue for the reliability of our moral beliefs, thereby fail.

(iii) Finally, in Korman and Locke (2020) we again see the same general strategy with a slightly different explanatory constraint:

(EC*): If P is about domain D, and S believes that her belief that P is neither explained by nor explains some D-facts, then S is thereby rationally committed to withholding belief that P.

According to Korman and Locke, evolutionary debunking arguments first establish that “realists are rationally committed to believing that their moral beliefs are not e-connected” to any moral facts (2020: 27). Evolutionary debunking arguments establish that our moral beliefs are not explained by moral facts which, given (EC*), entails that the minimalist “is rationally required to withhold on her moral beliefs” (2020: 325).

Korman and Locke therefore point out that “it is neither here nor there” that the minimalist can utilise substantive moral claims to argue for the reliability of our moral beliefs, “for those beliefs have already been undermined by minimalists’ explanatory concessions” (2020: 326). Minimalist responses therefore fail because they utilise defeated beliefs without argument.

4.8: An Objection to These Explanatory Constraints:

I want to argue that the explanatory constraints outlined in the previous section are flawed because they imply that equally justified beliefs are in fact unequally justified. Consider the following two cases (the first of which should be familiar from section 4.5 above)⁸⁵:

Shaquira:

Shaquira believes that

B1: protons cause streaks when they pass through cloud chambers.

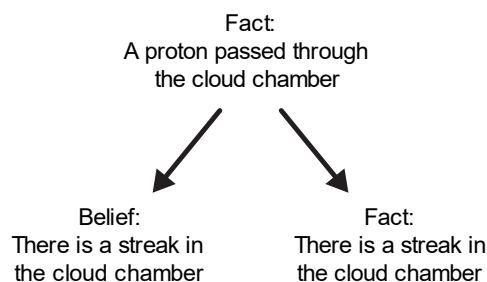
Shaquira believes this entirely because Martians have convinced her this is so. Furthermore, Martians do not actually have any good reason to believe that protons cause streaks when they pass through cloud chambers; they have told Shaquira this fact entirely because they like the way the English word “proton” sounds. Shaquira then finds out and accepts this fact about the Martian’s reasoning, but maintains her belief that protons cause streaks when they pass through cloud chambers. Later, Shaquira is informed by a reliable source that a proton was fired through the cloud chamber, and thereby comes to believe that

B2: there is a streak in the cloud chamber.

Shaquira believes that her (B2) belief is tf-connected, because the proton passing through the cloud chamber explained the streak, and also explained the reliable source informing her about the streak, which explained her belief. The proton passing through the cloud chamber therefore explains the streak and her belief about the streak.

⁸⁵ These examples are borrowed and adapted from Locke (2014).

We can represent the explanatory structure that Shaquira takes to hold for her B2 belief in figure 7. Plausibly, Shaquira's (B2) belief is unjustified despite her correct belief that it is tf-connected. Now consider a parallel, e-connected version of this scenario:



Jack:

Jack believes that

B1: protons cause streaks when they pass through cloud chambers.

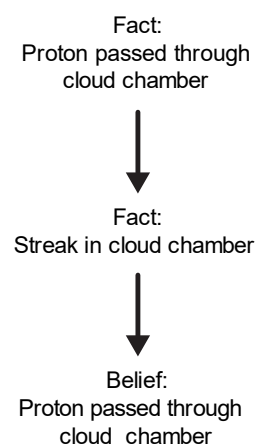
Jack believes this entirely because Martians have convinced him this is so. Furthermore, Martians do not actually have any good reason to believe that protons cause streaks when they pass through cloud chambers; they have told Jack this fact entirely because they like the way the English word "proton" sounds. Jack then finds out and accepts this fact about the Martian's reasoning, but maintains his belief that protons cause streaks when they pass through cloud chambers. Later Jack is informed by a reliable source that there is a streak in the cloud chamber, and thereby comes to believe that

B2*: a proton just passed through the cloud chamber.

Jack believes that his belief is e-connected, because the proton explained the streak, and explained the reliable source informing him about the streak, which then explained his belief in the proton.

We can represent the explanatory structure that Jack takes to hold for his B2* belief in figure 8. Like Shaquira's (B2) belief, Jack's (B2*) belief is unjustified. I think it is also clear that (B2*) is not even slightly more justified than (B2), and that they are both unjustified for exactly the same reason: because they are based on a belief (B1 and B1* respectively) that is unjustified. The two cases are structurally analogous and they exhibit the same kinds of flaws when it comes the agents' unjustified beliefs in (B2) and (B2*).

Figure 8:



Now recall that Korman (2019b) suggests the following explanatory constraint on rational belief:

(EC5) S is justified in believing p only if: either (i) S doesn't withhold belief that [S's belief that p is e-connected to p], or (ii) S's belief that p is based on beliefs $Bp_1 \dots Bp_n$ that prima facie justify S in believing p and S doesn't withhold belief that [S's beliefs that $p_1 \dots p_n$ are e-connected to $p_1 \dots p_n$].

Thus, according to (EC5), any belief that we do not take to be e-connected itself must be based on beliefs that we take to be e-connected. This feature of (EC5) leads to problems. It means that Jack's (B2*) belief passes (EC5) while Shaquira's (B2) belief does not, and is therefore unjustified. Jack believes his belief in a proton passing through the chamber is ultimately explained by a proton passing through the chamber. He therefore takes his belief to be e-connected and this belief is allowed by (EC5).

By contrast, Shaquira takes her belief to be tf-connected. Thus, the requirements for her belief to be justified (according to EC5) are greater than the requirements for Jack's (B2*) belief to be justified. Shaquira has to believe (B2) on the basis of beliefs that she takes to be e-connected. But she does not; (like Jack) she bases this belief on her belief that protons cause streaks in cloud chambers, which (like Jack) she accepts is not e-connected. Unlike Jack, this renders her (B2*) belief unjustified according to (EC5). Jack and Shaquira seem equally unjustified in their respective beliefs, and the fact that (EC5) fails to account for this indicates that there is something wrong with this explanatory constraint.

Now consider Korman and Locke's (EC*):

(EC*): If P is about domain D, and S believes that her belief that P is neither explained by nor explains some D-facts, then S is thereby rationally committed to withholding belief that P.

(EC*) implies that Shaquira's (B2*) belief ought to be withheld, but does not imply that same about Jack's (B2*) belief. Jack's belief escapes defeat because (B2*) is his belief that there is a proton in the cloud chamber, and he believes that this belief is ultimately explained by a proton passing through the cloud chamber. As such, Jack believes that (B2*) is explained by some fact in the same domain as the fact that makes

it true because he believes it is explained by that very fact, and all facts are in the same domain as themselves.

Shaquira, on the other hand, believes that (B2) is tf-connected to the fact that makes it true, rather than being e-connected to this fact. As a result, Shaquira's belief does not pass (EC*) in virtue of the explanatory relationship that she accepts between her belief and the fact that makes it true, as is the case with Jack. What's more, it is difficult to see how any *other* belief in the same domain as streaks forming in cloud chambers could explain her belief. She forms (B2) on the basis of a belief that streaks are caused by protons, and she accepts that the genealogy of this belief has nothing to do with streaks in cloud chambers (and has everything to do with Martians and their proclivity for the word "proton"). Shaquira therefore accepts that no fact in the relevant domain explains her belief, and this implies she ought to withhold her belief according to (EC*)⁸⁶.

(EC*) therefore entails Shaquira ought to withhold her (B2) belief but does not imply the same about Jack's (B2*) belief. This is because Jack believes his (B2) belief is e-connected to its truth, whereas Shaquira believes (B2*) is tf-connected to its truth. As the latter is not enough to pass (EC*), and Shaquira's belief is not explained by any other facts in the relevant domain, Shaquira's belief should be withheld. As Jack and Shaquira's beliefs are in fact equally unjustified, this is a problem for (EC*).

Now consider the following example, which focuses on the explanatory history of our *disposition* to form beliefs, rather than the beliefs themselves. Benacerraf questioned how we can be sure that our mathematical beliefs are reliable given mathematical Platonism (1973). One potential answer is as follows: God pre-determined the nature of our mathematical dispositions and ensured that those dispositions are correct. This story might be fleshed out in two ways:

Pat:

⁸⁶ This is true under the assumption that facts about protons and facts about streaks are in different domains. But this is an assumption that Korman and Locke are committed to; they reject an alternative explanatory constraint on the basis that it *cannot* render (something structurally identical to) Shaquira's (B2*) belief unjustified, and so it must be the case that their proposed principle *can* render her belief unjustified (2020: 325-326). This is only true if proton facts are different to streak facts, and so Korman and Locke must intend their use of "domain" to be understood as such.

Pat believes that mathematical truths exist independently of God's will, and God attended to those truths and based our mathematical dispositions on them. Pat therefore believes his mathematical dispositions are e-connected to the mathematical facts; his moral beliefs are based on the fact that God endowed him with these dispositions, and he believes this fact is itself explained by mathematical truths.

Amanda:

Amanda believes that it is within God's power to decide upon mathematical truths, and God therefore willed that whatever dispositions They endowed us with would thereby be true. Thus, the fact that God created us to be disposed to form the mathematical beliefs of a particular kind explained the fact that the corresponding mathematical facts exist⁸⁷. Amanda therefore believes that her disposition to form mathematical beliefs are tf-connected to the mathematical facts, because both are explained by God's endowing us with those dispositions.

Depending on one's opinion about the necessity of mathematical truths and the omnipotence of God's will, either Pat or Amanda's account might be preferable. But let us stipulate that Pat and Amanda are equally justified in all relevant supporting beliefs. That is, Pat is as justified in believing that God attended to independent mathematical truths as Amanda is in believing that God created those truths. Assuming this is the case, it seems that Pat and Amanda are equally justified in their mathematical beliefs. Nothing about the difference in the structure of the explanatory relationship that they believe holds between their disposition to form moral beliefs and the facts that make them true has any impact on the justificatory status of their mathematical beliefs.

Now Recall that Locke (2014) suggests the following explanatory constraint on belief:

⁸⁷ Importantly, this is different from the claim that mathematical facts are actually reducible to (or identical with) God's decision to endow us with certain mathematic dispositions. In the account under consideration, God's decision to endow us with certain mathematical dispositions explain the distinct fact that the corresponding mathematical facts exist.

Cognition Defeat. Where D is one of S's cognitive disposition to form beliefs involving the concept X, if S withholds belief that the explanatory history of D involves X-facts, then S is not default entitled to believe that D is reliable.

Cognition Defeat requires one to not withhold belief that one's disposition to form beliefs involving a particular concept are *e-connected* to the relevant facts. As a result, Cognition Defeat entails that Amanda is not default entitled to believe that her mathematical beliefs are reliable because she denies that the explanatory history of her mathematical dispositions involves mathematical facts⁸⁸. She thinks there is a third factor that explains both her dispositions to form mathematical beliefs and the mathematical facts, and that one does not explain the other. Amanda therefore withholds belief that the explanatory history of her belief involves mathematical facts, and the explanatory structure that she believes holds between her moral dispositions and the facts is insufficient according to Cognition Defeat. She is therefore not default entitled to believe that those dispositions are reliable.

By contrast, Pat is default entitled to assume that his mathematical dispositions are reliable. This is because he believes that the mathematical facts are involved in the explanatory history of those dispositions, as they determined which dispositions God gave us. I have suggested that the justificatory status of Pat and Amanda's mathematical beliefs are the same, and this implies that Cognition Defeat gives incorrect pronouncements on whether either Pat and Amanda's mathematical beliefs are justified.

I therefore think that we have reason to reject these explanatory constraints on belief. They falsely imply that equally justified beliefs have different degrees of justification and are thereby flawed.

Now, a proponent of these constraints might point out that they only provide a *necessary* condition for justification. As a result, they could be supplemented with a further condition for justification that would render the justificatory status of each pair of beliefs equal. This is especially true for (EC5) and (EC*), which would presumably have to be supplemented with a further condition on belief in order to render Jack's

⁸⁸ In the sense that Locke is understanding "involved", as outlined above.

(B2) belief unjustified and remain plausible. Given this additional condition, one could no longer claim that their account gives different judgements about the justificatory status of equally justified beliefs, because both beliefs would have been rendered unjustified.

But Shaquira's (B2) and Jack's (B2*) are not only both unjustified; they are also that they are unjustified for the same reason. It is implausible to claim that Shaquira accepting her belief to be tf-connected implies that her belief is unjustified in some further way that Jack's belief, which is taken to be e-connected, is not. They are both for exactly the same reason; they are based on a defeated belief. The justificatory status of their (B2) and (B2*) beliefs seems to be unaffected by the fact that one is taken to be tf-connected and the other e-connected.

Adding another condition to (EC5) and (EC*) would be able to render the two beliefs unjustified, but not for the same reason. Jack's belief would be unjustified in virtue of failing to meet whatever extra condition is added to (EC5) or (EC*) while Shaquira's belief would be unjustified because it fails (EC5) or (EC*). Thus, the constraints would still imply that Shaquira's belief is unjustified *because* the explanatory structure she accepts between her beliefs and the relevant facts is somehow inadequate in a way that Jack's accepted explanatory structure is not.

The constraints I have considered therefore cannot account for the fact that each of the pair beliefs are equally justified, and that the difference between them does not bear on the extent to which each are justified. In the next section, I will argue that (IEC) does not have this flaw.

4.9 Back to (IEC):

The reason that the constraints just considered are subject to counter-examples of this sort is that they imply there are different standards of justification for beliefs we take to be e-connected and beliefs we take to be tf-connected. According to Korman's (EC5), any belief that we take to be e-connected can escape defeat. All other beliefs (including those we take to be tf-connected) are unjustified if we do not take them to be ultimately inferred on the basis of beliefs we take to be e-connected. (EC5) therefore attaches different standards of justification for e-connected and tf-connected beliefs; the latter need to be based on beliefs we take to be tf-connected but the other

does not. It is this disparity that allows for counter-examples in which equally justified beliefs we take to be e-connected and tf-connected have different justificatory statuses; though all supporting beliefs are equally justified across both examples, the standards that (EC5) implies each belief needs to meet are different.

Korman and Locke's (EC*) implies that a belief can escape defeat if we take that belief to be e-connected, because we thereby take the belief to be explained by the fact that makes it true, and any fact is trivially in the same domain as itself. A belief that we take to be tf-connected is not trivially explained by some fact in the same domain as the fact that makes it true, and so in order for these beliefs to be justified they must have some additional explanatory relationship with some fact in the same domain. Once again, there are different standards of justification for beliefs we take to be e-connected and tf-connected, and this leads to possible examples in which beliefs we take to be e-connected and beliefs we take to be tf-connected are differently justified despite being otherwise identical.

Locke's *Cognition Defeat* is slightly different in that it implies different standards of justification for those beliefs that result from *dispositions* that are e-connected and those that result from dispositions that are tf-connected. Default entitlement for Amanda's beliefs is therefore ruled out because she takes her mathematical dispositions to be tf-connected to the relevant facts, whereas Pat's are allowed because he takes his dispositions to be e-connected.

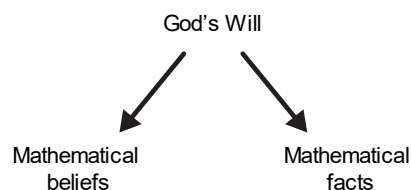
This implies that, in order for a constraint to avoid counter-examples of this kind, it must not attach different standards of justification for beliefs we take to be e-connected and those we take to be tf-connected, or those beliefs that result from dispositions that we take to be e-connected and tf-connected. (IEC) satisfies this condition. Recall that, according to (IEC):

(IEC): If S (i) withholds belief that her belief that P is explained by the fact that P and (ii) S withholds belief that there is some fact that explains both her belief that P and the fact that P, then S's belief that P is unjustified.

In order for our belief to avoid being defeated by (IEC) we can take that belief to be either e-connected or tf-connected. (IEC) does not specify any conditions that a belief we take to be tf-connected must satisfy in order to be justified, and therefore does not

posit additional conditions that a justified tf-connected belief must satisfy that beliefs we take to be e-connected need not.

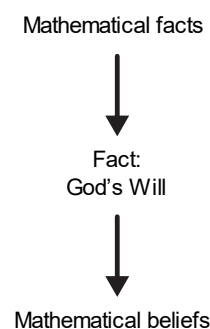
We should therefore expect that (IEC) is not Figure 9:



subject to the same counter-examples that are faced by those alternative constraints considered above. And this expectation is borne out when we apply (IEC) to those examples. According to (IEC), both Shaquira's B2 belief and Jack's B2* belief do not fail (IEC) because Shaquira takes her belief to be tf-connected and Jack takes his belief to be e-connected. But (IEC) *can* help explain why both these evidently unjustified beliefs are unjustified; they are both unjustified because they rely on (B1) and (B1*), and both of these beliefs fail (IEC) and are thereby defeated. Thus, (IEC) does not falsely imply that Shaquira and Jack's belief have different degrees of justification, and it can also explain why both of these beliefs are unjustified.

Under the assumption that our disposition to form a belief in P in some way explains our belief in P, (IEC) implies that both Amanda's and Pat's mathematical beliefs are undefeated. Amanda takes her mathematical beliefs to be tf-connected, as outlined in figure 9, and Pat takes his beliefs to be e-connected, as outlined in figure 10. Because (IEC) requires our withholding belief that our belief is e-connected and tf-connected in order to be defeated, (IEC) implies that neither belief is defeated. Although, as indicated above, this does not mean that we are committed to either being ultimately justified if we accept (IEC). (IEC) expresses a sufficient, rather than a necessary, condition on defeat. As such, these beliefs might be defeated for reasons that have nothing to do with (IEC), or they may be defeated in the same way that Shaquira and Jack's beliefs were defeated (i.e, because a belief on which they rely fails (IEC)).

Figure 10:



(IEC) therefore avoids the counter-examples expressed above. It does so because it does not attach different standards of justification to e-connected beliefs or tf-connected beliefs. Now, some may view this as a flaw of (IEC); surely beliefs we take to be e-connected must be in some sense prior to, or more easily justified than those

we take to be tf-connected. Surely the standards for each kind of belief being justified cannot be exactly the same.

I think that the motivation behind attaching different standards of justification to e-connected and tf-connected beliefs might be as follows. If we take a belief in P to be tf-connected, this might presuppose that we already have some beliefs about P. Specifically, perhaps we must already believe that there is an explanatory relationship between the fact on which we base our belief in P and P itself. We cannot base our belief in P on Q if we do not already believe in some relationship between P and Q. So if we base our belief in P on some distinct fact Q which we take to explain P, this rationally commits us to some prior belief in a relationship between P and Q⁸⁹.

This implies that a belief we take to be tf-connected can only be justified if we have the correct background beliefs. Specifically, justification for some tf-connected belief depends on our already believing that Q explains P. By contrast, when we have “direct” access to P, we need have no such background beliefs about P. Our access to P justifies our belief in P regardless of whatever background beliefs we have about P. So the conditions under which beliefs we take to be tf-connected can be justified are different from those we take to be e-connected; the former cannot be justified without these supporting background beliefs whereas the latter can.

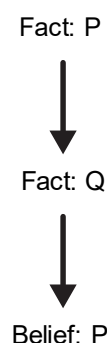
This line of thought might also motivate the claim that our beliefs about a given fact, or in a given domain, cannot be tf-connected “all the way down”. If our beliefs about some fact were exclusively tf-connected, there would be at least one background belief (about the relationship between P and Q) presupposed by such beliefs that we would fail to have, and our beliefs about P would be ultimately without foundation. At some point we need to have a belief about P that is not based on something that is distinct from P, but which is based on direct contact with P itself. In other words, our beliefs about P must ultimately be justified by beliefs that we take to be explained by P, and so tf-connected beliefs are only justified to the extent that they are ultimately based on e-connected beliefs.

⁸⁹ I do not mean to endorse this line of thought. I just mean to express it so I can comment on its significance to the current discussion further down.

I will not comment on the success of these arguments, and will instead point out that it applies just as much to a certain class of e-connected beliefs as it does tf-connected beliefs. Let us say a belief is “indirectly” e-connected if there is some fact Q that explains our belief and is itself explained by P. This can be represented in figure 11. Because P explains Q and Q explains our belief in P, our P explains our belief in P and so our belief is thereby e-connected⁹⁰.

When we take our belief to be indirectly e-connected, we take ourselves to be forming a belief in P on the basis of some distinct fact Q. Just as we did for tf-connected beliefs, we could therefore reason that those beliefs we take to be indirectly e-connected presuppose certain background beliefs, about the relationship between the basis and the object of our belief. As such, beliefs we take to be indirectly e-connected can only be justified if we have the appropriate background beliefs, whereas having direct contact with P justifies our belief in P regardless of whatever background beliefs we have. Furthermore, this might motivate the claim that our beliefs about P cannot be indirectly e-connected all the way down. At some point we must have direct contact with P in order to justify our other subsequent indirectly e-connected beliefs about P.

Figure 11:



The considerations that appeared to support different standards of justification for e-connected beliefs over tf-connected beliefs therefore actually supports different standards of justification for a particular kind of “direct” e-connected beliefs over and above both indirectly e-connected beliefs and tf-connected beliefs. As the Shaquira/Jack and Pat/Amanda examples indicate, indirectly e-connected beliefs and tf-connected beliefs do not have fundamentally different standards of justification, and the fact that (IEC) is consistent with this fact is my reason for preferring it to Cognition Defeat, (EC5) and (EC*). Because these latter constraints implied different standards

⁹⁰ I have assumed that, in this particular case, explanation is *transitive*. Whether or not explanation is always transitive is a matter of some debate. There may be cases in which A explains B and B explains C but A does not explain C (Lehrer, 2020). But nobody, as far as I am aware, doubts that explanation is often transitive. Indeed, the counter-examples to explanatory transitivity are typically viewed as having a very specific structure that is not present in *all* cases of indirect explanation (see, for example, Stevens (2008) and Paul and Hall (2013)). We therefore have no reason to think that all chains of e-connections will result in a breakdown of explanatory transitivity.

of justification to e-connected and tf-connected beliefs, they implied that there are different standards of justification for indirectly e-connected and tf-connected beliefs, and they were therefore subject to the counterexamples that I have presented in this section.

4.10: Conclusion:

In this chapter, I have outlined and defended an explanatory constraint on belief. I have argued that the constraint correctly accounts for our intuitions about defeat and avoids difficulties that are faced by alternative explanatory constraints. I have also pointed out that, according to this constraint, our moral beliefs are not defeated by EDA's. EDA's establish that those beliefs are not e-connected, but there remains the possibility that they are tf-connected. So long as we can avoid withholding belief in this alternative explanatory connection, our realistically construed moral beliefs are not defeated by EDA's and minimalist replies can rely on substantive moral claims when responding to such arguments.

That being said, this does not imply that minimalist replies are successful, or that their reliance on substantive moral claims is unproblematic. I think there are reasons, independent of evolutionary debunking arguments, for thinking that our "basic" realistically construed moral beliefs cannot be tf-connected. If this much is established, then we can combine these considerations with evolutionary debunking arguments to show that our basic moral beliefs can be neither e-connected nor tf-connected under the assumption of moral realism. According to (IEC), once we accept this point, our realistically construed basic moral beliefs become unjustified, and any reliance on those beliefs is illegitimate. It is true that EDAs only establish that our moral beliefs are not e-connected, but there are in fact other, independent reasons for withholding belief in our basic moral beliefs being tf-connected. And once basic moral beliefs are off the table, I argue, the rest are soon to follow. I make this argument in the following chapter.

Chapter 5: An Expanded Evolutionary Debunking Argument:

5.1: Introduction:

Thus far, I have argued for the following explanatory constraint on belief:

(IEC): If S withholds belief that (i) her belief that P is explained by the fact that P and (ii) S withholds belief that there is some fact that explains both her belief that P and the fact that P, then S's belief that P is unjustified.

I have also argued that evolutionary debunking arguments do not imply that our realistically construed moral beliefs are defeated according to this constraint. As such, minimalist replies can accept the explanatory disconnect implied by debunking arguments and then rely on moral claims because accepting that explanatory disconnect does not defeat our moral beliefs.

In this chapter, I show that traditional evolutionary debunking arguments can be combined with additional arguments to imply an explanatory disconnect that does in fact defeat our moral beliefs given (IEC). There are reasons, independent of evolutionary debunking arguments, for thinking that our basic moral beliefs cannot be *tf*-connected. Assuming that evolutionary debunking arguments already establish that our moral beliefs are not *e*-connected, the combination of traditionally construed debunking arguments and these additional arguments imply that our basic moral beliefs are defeated according to (IEC). This, I will argue, ultimately implies that minimalist replies to debunking arguments fail.

Because the target of this chapter is minimalist replies to debunking arguments I will write as though evolutionary debunking arguments successfully establish that our moral beliefs are not *e*-connected. My argument is that, if we grant this point (as minimalist replies do) then there are good reasons to think that our realistically construed moral beliefs are defeated by the conjunction of EDA's and the independent arguments that I am about to present. I will therefore assume that evolutionary debunking arguments establish that our moral beliefs are not *e*-connected because my gripe with the minimalist concerns what follows once we make this assumption. I therefore write as though evolutionary debunking arguments successfully establish

that our moral beliefs are not e-connected because I am writing in a context where this point is taken for granted in order to determine what then follows from this concession.

I start this chapter by outlining a particular account of the structure of normative reality and justifying my working with this account going forward. Next I argue that, given this structure of normative reality, the realist's beliefs in basic moral facts cannot be tf-connected. Assuming that EDA's have already established our moral beliefs cannot be e-connected given moral realism, this implies that our realistically construed beliefs that we take to be about basic moral facts are defeated under the assumption of (IEC). I will then provide two separate arguments to show that, if our beliefs in facts that we take to be basic are defeated, then so are our beliefs in facts that we take to be non-basic. I will argue that this implies that *all* of our moral beliefs have been defeated under the assumption of moral realism, and minimalist replies therefore fail when evolutionary debunking arguments are combined with the additional arguments presented in this chapter. I end by considering two potential responses from the minimalist.

In chapter four I considered an objection to a principle like (IEC) raised by Korman and Locke. I take it that Korman and Locke felt the need to rule out this constraint because of the worry that, if it were true, it would not imply that the minimalist's moral beliefs are defeated by debunking arguments. As their goal is to rule out all minimalist responses to debunking arguments on the grounds that the explanatory concession implied by evolutionary debunking arguments defeats our realistically construed moral beliefs, they had to show that a constraint like (IEC), which could not ground this claim, should be avoided.

Indeed, in setting up and arguing for third-factor replies, commentators sometimes make comments that seem to support something in the ballpark of (IEC). Enoch views the basic epistemological challenge posed to moral realism by EDAs as the challenge of explaining the correlation between our moral beliefs and moral facts. As EDA's rule out direct explanatory link between the two, how else might this correlation be explained? Enoch claims that

“The thing to look for is a third-factor explanation. For it is possible that the explanation of a correlation between the two factors A and B is in terms of a third factor, C, that is (roughly speaking) responsible both for A-facts and for B-facts [...]

the realist-friendly explanation of the correlation I am about to offer is exactly such a third-factor explanation” (2010: 429).

If we grant (IEC), we therefore seem to sanction these replies to debunking arguments. But, if what I am about to argue is true, then (IEC) can, somewhat surprisingly, be used to establish that minimalist replies to debunking arguments fail⁹¹.

5.2: The Structure of Normative Reality:

In what follows, I will be working with a particular account of the structure of normative reality⁹². This account of normative reality will feature crucially in my argument that minimalist replies to debunking arguments fail. As I will outline below, part of my justification for assuming this picture of normative reality is that various non-naturalist realists seem to either explicitly or implicitly endorse it and, as I will argue, I think that they have good reason to do so. Furthermore, minimalist responses actually presuppose all or part of this account of normative reality. As such, I take this account to be common ground when arguing against the realist and/or minimalist.

According to this account, there are basic moral facts. These moral facts are unexplained; they are not explained by any other moral truth, nor by any contingent natural fact, nor by anything else⁹³. There may be just one of these facts, or there may be a number of these basic moral facts. These facts might be “snappy” and simple enough to play a role in moral reasoning. For example, among these basic moral facts might be the principle that causing unnecessary pain is wrong. Alternatively, the basic moral facts might assign a moral property to a potentially infinite chain of natural conditions, and therefore be too complex to ever play a role in moral reasoning⁹⁴. For

⁹¹ As will become clear, other arguments that are independent of (IEC) are also necessary for establishing this point, but (IEC) itself plays a crucial role in this process.

⁹² This structure has been explicitly outlined by Rosen (2020), but as I outline below, other commentators also endorse part or all of it.

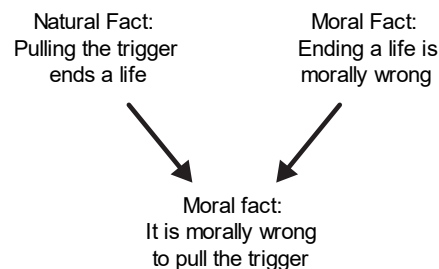
⁹³ These basic moral facts are traditionally construed as being metaphysically necessary. However, some commentators have denied this, taking basic moral facts to be *normatively*, but not metaphysically, necessary (Hattiangadi, 2018; Rosen, 2020). These commentators still construe such basic moral facts as being unexplained, and so these positions are consistent with the account of normative reality that I am describing.

⁹⁴ This account is therefore consistent with those particularists who accept the existence of basic moral facts but deny that they take the form of snappy, action-guiding moral principles. As Rosen notes, the metaphysical particularist who denies the existence of any moral principles whatsoever may be fictional (2020: 207). Secondly, there is reason for thinking that this kind of metaphysical particularism is incoherent (2017: 167).

the sake of simplicity, in what follows I will assume that these basic moral facts are snappy, but nothing in my argument depends on this claim.

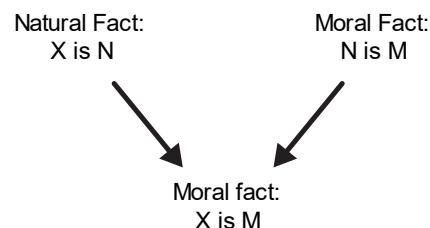
These basic moral truths form the foundation for the rest of morality. More specifically, all non-basic moral truths are explained by basic moral truths. The fact that it is morally wrong to pull the trigger of this gun might be *explained* by the fact that it is morally wrong to end a life. As this example suggests, non-basic moral facts are explained by natural facts as well as being explained by basic moral facts. In the example just given, the natural fact that pulling the trigger will end a life *also* explains the fact that pulling the trigger is morally wrong (as well as the fact that ending a life is wrong). We can represent this explanatory structure in figure 12.

Figure 12:



Generally speaking, a non-basic moral fact of the form X is M (where X is denotes something natural like an event or an action, and M denotes a moral property) is explained by a natural fact in combination with a more basic moral fact. The natural fact will take the form X is N (where N denotes some natural property possessed by the event or action) and the more basic moral fact will take the form N is M (where M is denotes the same moral property ascribed to X according to the explained, non-basic moral fact that X is M). This structure is represented in figure 13. When I refer to some natural fact explaining a moral fact, I am specifically referring to whatever natural fact that occupies the position occupied by X is N in figure 13.

Figure 13:



This “more basic” moral fact might itself be non-basic, and so might be explained by a corresponding natural and moral fact. Or it might be basic, and therefore unexplained by facts of any kind. This explanatory structure is supposed to fully account for all moral facts. Either they are basic, in which case they are unexplained, or they are non-basic, and are explained by some more basic moral fact in combination with the relevant natural fact.

The features of this account of normative reality are not unrelated. It is hard to see how one can coherently accept the existence of basic moral truths without accepting that less basic moral truths are explained by these basic moral truths in combination with the relevant natural fact. Basic moral truths posit a necessary connection between moral properties and natural properties; pain is bad, or survival is good. Once we accept the existence of these truths, it is hard to see how we could deny that the fact that some specific action causes pain, in combination with the fact that pain is bad, is responsible for the fact that the action in question is bad. How could we deny this without also denying the basic moral fact that pain is bad? So, by accepting the existence of basic moral truths that posit a necessary connection between natural and moral properties, we thereby seem bound to accept that less basic truths will be explained by the relevant natural fact and the more basic moral truth. As a result, the full account of normative reality that I have been describing follows naturally from the claim that there are basic moral facts of the relevant kind.

In the next section I will show that some of the non-naturalists that I have been discussing explicitly accept this account, and argue that there are good reasons for non-naturalists to do so.

5.3: Non-naturalism and the Structure of Normative Reality:

Non-naturalists are often sympathetic to the picture of normative reality just described. Thus, Wielenberg explains:

“my version of non-theistic robust normative realism [...] implies the obtaining of substantive, metaphysically necessary, brute facts. Some ethical facts fall into this category; I call such facts basic ethical facts. Such facts are the foundation of (the rest of) objective morality and rest on no foundation themselves. To ask of such facts, “where do they come from?” or “on what foundation do they rest?” is misguided in much the way that, according to many theists, it is misguided to ask of God, “where does He come from?” or “on what foundation does He rest?” The answer is the same in both cases: they come from nowhere, and nothing external to themselves grounds their existence; rather, they are basic features of the universe that ground other truths.” (2014: 38).

Here we get two of the features of the structure of normative reality that I have been outlining: the existence unexplained moral facts, and the claim that these facts explain all other moral facts. Elsewhere, Wielenberg writes:

“Explanation, as they say, must come to an end somewhere. Why does being an instance of torturing someone just for fun entail moral wrongness? Because being an instance of torturing someone just for fun makes an act wrong. But why does being an instance of torture just for fun make an act wrong? Perhaps further explanation is available: for example, perhaps torturing just for fun never maximizes utility and failing to maximize utility makes an act wrong. But why does failing to maximize utility make an act wrong? Eventually we hit bottom; no further explanation is available. But I don’t see why possessing this sort of explanatory bottom is a problematic feature for a view to have.” (2014: 24)

Here we get the third feature of the described account of normative reality: non-basic moral facts, in addition to being explained by more basic moral facts, are also explained by the relevant natural fact. These three features make up the structure of normative reality that I have endorsed, and it therefore seems Wielenberg endorses this account.

Enoch makes comments that seem to naturally imply this account of normative reality as well. Firstly, Enoch takes there to be “basic” or “ultimate” moral norms that are (at least) metaphysically necessary (2011: 146). The fact that these norms are metaphysically necessary implies that they cannot be explained by contingent natural phenomena. Furthermore, the “basic” and “ultimate” terminology implies that these moral facts do not depend for their existence on some other moral fact. It is hard to see how these basic moral facts could be explained by anything *other* than moral or natural facts and, from the fact that they are explained by neither, it is reasonable to conclude that they are unexplained.

What is the relationship between these basic moral facts and the rest of normative reality? The following quote will be useful here, in which Enoch clarifies what he means by a basic moral norm:

“There may be some problems in making the idea of a basic or an ultimate norm fully explicit. But the idea is, I take it, intuitive enough – we just backtrack, so to speak, practical syllogisms to their major premises, until we reach such a major premise that

is not itself the conclusion of a practical syllogism, or that is (roughly speaking) free of empirical content.” (2011: 150).

I take it that the major premise is a moral principle (i.e, survival is good), the minor premise a particular instance (i.e, looking after our offspring promotes survival), and the some less basic moral fact (looking after our offspring is good).

Strictly speaking Enoch is describing premises and conclusions in arguments, rather than the mind-independent metaphysical relations between these facts. But, according to Enoch, by “backtracking” these syllogisms to their ultimate major conclusion we arrive at those moral facts that are *metaphysically* the most basic. This seems to imply a relationship between the inferential relations of practical syllogisms and the metaphysical relations in mind-independent normative reality; if by backtracking these syllogisms as far as they can go we arrive at the most metaphysically basic truths, it is not unreasonably to suppose that the position of facts in this inferential structure reflects their metaphysical position in the structure of normative reality. If this is correct, then Enoch takes the basic moral truths (N is M) to explain less basic moral truths (X is M), in combination with some natural fact (X is N) (because this reflects the inferential relations in practical syllogisms)⁹⁵. What’s more, this non-basic moral truth can then explain even less basic moral truths in the same way that the major premise of a practical syllogism can itself be the conclusion of another practical syllogism.

Putting all this together, we get an account of normative reality according to which basic, unexplained moral facts explain all other moral facts, and they do so in conjunction with the relevant natural facts. This is the structure or normative reality outlined above.

I think it is no accident that the non-naturalists I have been discussing endorse this structure of normative reality. I think that it is difficult to see how the realist could deny this picture of normative reality without being forced to give up non-naturalism in favour of some form of naturalism.

⁹⁵ This interpretation is strengthened by the fact that Enoch presents a third-factor response to debunking arguments and, as I will outline below, third-factor replies assume this feature of the structure of normative reality.

The argument for this claim stems from the claim that we should all accept (naturalists and non-naturalists alike) that the moral properties of some action or state of affairs are at least partially explained by natural properties. It is highly plausible to think that the fact that my action is wrong is (at least partially) explained by the fact that the action promotes pain, or that it involves breaking a promise⁹⁶. To deny that the natural features of some action go any way to determining the moral properties of that action is highly counter-intuitive.

One prominent specification of how the natural features of some action or scene explain its moral properties is that they ground those moral properties (Rosen, 2010). But the non-naturalist cannot claim that moral properties are *fully* grounded in natural properties, because this would imply ethical naturalism⁹⁷. In order to see this, we can note that the grounding relation is supposed to be a “maximally intimate explanatory relation” such that, if X grounds Y then, though X and Y are distinct, Y is not “a separate, free floating fact” (Rosen, 2017, 156). This implies that, if moral facts are *fully* grounded in natural facts, then it is difficult to see how moral facts are not themselves just natural facts (albeit natural facts that are distinct from any of the natural facts that ground them). As such, if we accept that natural properties exhaustively ground moral properties, it seems that we ought to accept ethical naturalism according to which normative properties are natural properties. Obviously, this is not an option for the non-naturalist.

The non-naturalist should therefore accept that the rightness or wrongness of particular actions are only partially grounded in natural facts and are *also* grounded in some moral fact as well. This preserves their hypothesis that moral properties are non-natural. Call this moral fact that partially grounds the rightness or wrongness of particular actions the “more-basic” moral fact.

This more-basic moral fact cannot itself be completely grounded in natural facts because, once again, this would imply ethical naturalism. Natural facts would fully ground this more-basic moral fact, and the less-basic moral fact would then itself be fully grounded in natural facts (because it would be fully grounded in natural facts as

⁹⁶ See, for example, Caj Strandberg, (2009: 129), Rosen (2017: 63) and Hattiangadi (2018: 599).

⁹⁷ According to Rosen (2017), ethical naturalism just *is* the thesis that normative facts can be entirely grounded in natural facts. Berker (2014) defines naturalism as the claim that normative properties are identical with *or* entirely grounded in natural properties.

well as this more-basic moral fact, and the more-basic moral fact would itself be fully grounded in natural facts)⁹⁸.

So the non-naturalist can either take this more basic moral fact to be *unexplained*, or they can take it to be grounded in both natural and moral facts⁹⁹. If they go for the latter option, then they can ask what grounds this even-more-basic moral fact. And we can see that the very same reasoning will apply concerning the grounds of this even-more-basic moral fact as well.

If we ignore the possibility that normative reality continues indefinitely, with every moral fact being itself explained by a more-basic moral fact (more on this possibility below), we have to posit a *most basic* moral fact that is not itself explained by any other more basic moral fact. This most-basic moral fact cannot itself be completely grounded in natural facts because, once again, this would imply naturalism. But it also cannot be explained by any more moral facts because we are supposing that this is the *most basic* moral fact (if it were explained by other moral facts, then this would just imply that we are not actually at the top of the explanatory ladder as we had previously stipulated). As such, this most basic moral fact can be neither explained by natural nor moral facts¹⁰⁰. There are therefore no plausible candidates for what might explain this most basic moral fact, and we should view it as unexplained.

In short, in order to avoid the implication that all moral facts are ultimately entirely grounded in natural facts, (thereby avoiding naturalism) the non-naturalist has to posit unexplained basic moral facts that ground all other moral facts. If the non-naturalist then hopes to acknowledge that natural facts do in fact partially ground particular moral facts, they seem forced to accept that account of normative reality I have been

⁹⁸ This assumes that grounding is transitive. For reasons to think it is not always transitive, see Shaffer (2016). There is no reason to think that transitivity breaks down in this case, even if it does break down in the examples that he gives.

⁹⁹ They might also take this moral fact to be fully grounded in other moral facts. This makes no difference to the argument that I am making, and ultimately implies the existence of unexplained basic moral facts in the same way.

¹⁰⁰ These basic moral facts cannot be explained by anything *else*, such as the natures of the natural and moral properties involved in this fact, because this involves positing too intimate a connection between natural and moral facts and threatens to collapse the position into naturalism. See, for example, (Hattiangadi, 2018) (Rosen, 2020: 223). I am also ignoring the possibility that these basic moral facts are explained by some supernatural fact like the essence of God; none of the non-naturalists that I am concerned with endorse this as a feature of their account.

describing, according to which there are unexplained basic moral facts that explain all other moral facts in combination with the appropriate natural facts.

One way for the realist to object to this argument is to deny that natural properties explain moral properties by grounding them; perhaps some alternative explanatory relation holds between the two. However, in order for this move to effectively block the argument just given, the alternative explanatory relation between natural and moral properties cannot be so intimate that it furnishes the very same conclusion for which I have just argued. For example, Enoch's suggestion that the natural properties of some act *constitutes* its moral properties (2010: 431) will not allow the realist to deny my argument that they are committed to the relevant account of normative reality. This is because the idea that natural properties fully constitute moral properties implies naturalism just as much as the claim that natural properties fully ground moral properties. As such, in order to avoid collapsing into naturalism, the non-naturalist cannot accept that the moral properties of particular actions are fully constituted by natural facts, and must claim that there is also some moral constituent of moral properties. But we can then ask about the constitution of this moral constituent, and an analogue of the argument I have just given is set in motion.

If the non-naturalist wants to avoid endorsing this structure of normative reality, then, they need to defend an alternative explanatory relation that holds between natural and moral properties such that natural facts can fully explain moral facts without this position collapsing into some form of ethical naturalism. If they cannot do this, then there is good reason for them to accept the structure of normative reality that I have been defending¹⁰¹.

5.4: Minimalist Replies and the Structure of Normative Reality:

Finally, we can note that the minimalist responses that I will be arguing against actually rely on some or all of this account of normative reality. The modal security response first points out that basic moral facts are metaphysically necessary. Thus,

¹⁰¹ Though I lack the space to defend this claim, I also think that the non-naturalist is well-advised to accept the structure of normative reality that I have outlined in order to account for the supervenience of the moral on the natural. See Stranberg (2008) and Enoch (2011) for the relationship between basic moral principles and supervenience. See Fitzpatrick (2008) and Hattiangadi (2018) for further reasons for thinking the non-naturalist ought to accept this account of normative reality.

they argue, there are no possible worlds in these moral facts fail to obtain, and so our beliefs about them are trivially sensitive.

They then argue that our beliefs in non-basic moral facts could not easily have been wrong, because the closest worlds in which these beliefs are false are presumably those in which the natural facts that are responsible for these non-basic moral facts do not obtain. Here, then, we get the idea that natural facts combine with basic facts to explain non-basic facts, such that in possible worlds without the relevant natural facts the corresponding non-basic facts also do not obtain. This implies that there are unexplained basic moral facts which, in combination with the relevant natural facts, explain non-basic moral facts. The modal security response therefore presupposes the structure of normative reality that I have described.

Third-factor replies take the moral fact that, say, looking after our offspring is good (X is M) to be explained by the natural fact that looking after our offspring promotes survival (X is N). Furthermore, in the process of making this claim, they crucially presuppose a moral claim like survival is good (N is M), implying that the relevant natural fact could not explain the moral fact without it also being the case that some other more basic moral fact *also* exists.

All we need to add to this set-up in order to get the account of normative reality that I have been describing is that it bottoms out in unexplained basic moral facts. Now, there is nothing in the third-factor reply that essentially makes reference to basic moral facts¹⁰², but it is difficult to see how the third-factor replier *could* deny the existence of such facts.

One potential way to do so would be to claim that is that there is an infinite series of non-basic moral facts that are themselves explained by a more basic moral fact as well as the appropriate natural fact. But, worries about infinity and regresses aside, this suggestion is unable to account for the fact that some moral facts are most plausibly interpreted as not being explained by any other moral fact. What more basic moral fact could possibly explain the fact that pain is bad? What natural property *other* than pain might instances of pain instantiate that would explain why causing pain is wrong?

¹⁰² Although, as I have just outlined, in practice both Enoch and Wielenberg do endorse the existence of such facts.

The absence a plausible answer to these questions would also imply that explanation does not go in a big circle, in the sense that moral facts are explained by some fact that, ultimately, they themselves explain. It is plausible that pain does not derive its moral status from some other moral fact, regardless of whether this fact is itself explained by a circular chain of metaphysical dependence relations that loops back to the fact that pain is bad. Besides, a big circle of explanatory relations would presumably imply that explanations are reflexive, which is something that most commentators deny¹⁰³.

Some moral facts, then, seem not to depend on other moral facts. The third-factor replier therefore cannot suggest that the explanatory structure they endorse for some moral facts, such that they are explained by the combination of a more basic moral fact and a natural fact, continues indefinitely. It has to end somewhere because we must eventually arrive at a moral fact that is not explained by some other moral fact.

Now this does not yet imply the existence of unexplained basic moral facts. It could be the case that this explanatory structure bottoms out in the most basic moral facts that are not themselves explained by any moral fact but are instead entirely explained by natural facts. Causing pain might be bad because doing so instantiates some further natural property, even though there is no general moral principle such that instantiating that further moral property is wrong.

Firstly, note that this positions arguably entails moral naturalism, for the reasons given above. Moral facts would ultimately be entirely grounded in natural facts, and it is difficult to see how the non-naturalist could accept this point. Secondly, we can note that the third-factor replier has accepted that generalisable moral principles do exist, in the form of the moral fact that they assume and in virtue of which our beliefs can be tf-connected. Survival is generally good, and this explains why particular instances of promoting survival (i.e, when we care for our offspring) are also good. So there can be nothing wrong with general moral principles *per se*, so why should we deny that they exist at this most basic level? Once we have allowed that generalisable moral principles exist, it is hard to see why we should not allow that they act as unexplained

¹⁰³ For an overview of the consensus that metaphysical dependence is not reflexive (and an assessment of this assumption) see Jenkins (2011). For a positive argument that explanation cannot be reflexive, see Guigon (2015).

basic moral truths, especially if by doing so we can allay the worries about collapsing into naturalism that were outlined above.

This concludes my justification for assuming the structure of normative reality that I have described. This picture of normative reality is endorsed by robust realists, and there seems to be good principled reasons for their doing so. Furthermore, the minimalist replies that I am arguing against also presuppose at least part of it, and it is hard to see how they could deny the part that they do not presuppose. I think this justifies me in assuming this account when arguing against minimalist replies to debunking arguments.

Corresponding to my distinction between basic and non-basic moral facts, we can distinguish between basic and non-basic moral beliefs. Basic moral beliefs are beliefs in those moral facts that the believer takes to be basic, and non-basic moral beliefs are beliefs in those moral facts that the believer takes to be non-basic. Importantly, our basic moral beliefs need not be inferred on the basis of non-basic moral beliefs. The basicness of basic moral beliefs refers to the supposed metaphysical status of the facts in question, and not any kind of foundational role that our belief about such facts might play in moral reasoning. For all I have said so far, a non-basic moral belief of ours might serve as the cornerstone around which we construct a normative theory. I will say more about this possibility in the next chapter.

5.5: The Basic Argument:

I will now argue that, given the account of normative reality that I have been discussing, our basic moral beliefs cannot be tf-connected. When we combine this argument with evolutionary debunking arguments (which, we are assuming, establish that none of our moral beliefs are e-connected) this implies that we should withhold belief that our basic moral beliefs are e-connected and we should withhold belief that our basic moral beliefs are tf-connected. This implies that those beliefs fail (IEC) and are unjustified. In the next section I will argue that if our basic moral beliefs are unjustified then so are our non-basic beliefs. This implies that all of our moral beliefs are defeated by the combination of those debunking arguments and the arguments that I present in the following two sections, and minimalist replies to debunking arguments fail.

In order for some belief in P to be tf-connected, there must be some fact that explains both our belief in P and P itself. That there is such a fact is just what it means for our belief in P to be tf-connected. And, in order for there to be some fact that explains both our belief in P and P, there must be some fact that explains P. If there is no fact that explains P, then there can be no fact that explains both our belief in P and P itself.

This implies that beliefs about unexplained facts cannot be tf-connected. If our belief is about some fact that is unexplained, then there can be no fact that explains the fact that makes it true, and so there can be no fact that explains both one's belief about it and the fact that makes it true. Because this is required in order for a belief to be tf-connected, such beliefs cannot be tf-connected.

According to the structure of normative reality that I have been outlining above, basic moral facts are unexplained. As basic moral facts are unexplained, our beliefs in basic moral facts cannot be tf-connected. Our beliefs about basic moral facts cannot be explained by some fact that also explains the basic moral facts, because there is no fact that explains those basic moral facts. According to the normative structure of reality that I have outlined, then, it follows that our basic moral beliefs cannot be tf-connected. We should therefore withhold belief that our basic moral beliefs are tf-connected.

As I outlined above, we can assume that evolutionary debunking arguments establish that we should withhold belief that all of our moral beliefs are e-connected. When pre-existing debunking arguments are fortified with the argument I have just presented that beliefs in basic moral facts cannot be tf-connected, we see that the minimalist is bound to accept that our beliefs in basic moral facts can be neither e-connected nor tf-connected. Recall that according to (IEC):

(IEC): If S (i) withholds belief that her belief that P is explained by the fact that P and (ii) S withholds belief that there is some fact that explains both her belief that P and the fact that P, then S's belief that P is unjustified.

Because (IEC) stipulates that S's belief in P is unjustified if S withholds belief that it is e-connected or tf-connected, rather than if the belief in question actually *is* e-connected or tf-connected, (IEC) does not imply that all our beliefs that are *actually about* basic moral facts are defeated. Instead, it implies that all beliefs that *we take to*

be about basic moral facts are e-connected or tf-connected. If some moral belief of ours is about a basic moral fact but we do not realise it, then we may not have any reason to think that it cannot be tf-connected, and so we may justifiably not withhold belief that it is tf-connected. Conversely, if we mistakenly believe that a belief of ours is in some basic moral fact then we should withhold belief that it is e-connected or tf-connected, and that belief is defeated according to (IEC). This true even if, unbeknownst to us, the belief is not about a basic moral fact and is in fact tf-connected. My argument that beliefs that are actually about basic moral facts cannot be e-connected or tf-connected therefore implies, when we combine it with (IEC), that we should withhold belief in those moral facts that we take to be basic.

This implies that, as long as we grant that EDAs establish our moral beliefs are not e-connected (as minimalist replies do) beliefs that one takes to be about basic moral facts are defeated by the conjunction of EDAs and this new argument. In the context of the debate between evolutionary debunking arguments and minimalist replies, this has the implication that the minimalist cannot rely on beliefs in moral facts that they take to be basic, because those beliefs have been defeated.

Importantly, this argument does not presuppose anything about our reasons for forming beliefs in moral facts that we take to be basic. It is consistent with the claim that we infer them from moral beliefs in other facts that we take to be non-basic, or on the basis of moral beliefs about whose basicness we have no opinion. That our beliefs in basic moral facts cannot be tf-connected is entailed by the fact that basic moral facts are unexplained. So even if we have inferred our beliefs in facts that we take to be basic on the basis of other moral beliefs, our beliefs in basic moral facts still cannot be tf-connected because the basic moral facts still are not explained by any other moral fact. So *however* we form our moral beliefs, if we take them to be about basic moral beliefs we ought therefore withhold belief that they are tf-connected (assuming the structure of normative reality outlined in 5.2). If we also grant that they are not e-connected, as minimalist replies do, then we ought to withhold those beliefs according to (IEC).

One might be inclined to insist that if our basic moral beliefs are inferred on the basis of other moral beliefs they can still remain justified *even if* I am correct in claiming that we ought to withhold belief that they are e-connected and tf-connected. In the next

chapter, I will investigate the claim that (IEC) is false because we can remain justified in beliefs that we deny are either e-connected and tf-connected if we infer those beliefs from other beliefs. For now, I just hope to have established the conclusion that given the validity of (IEC), then, because our beliefs in basic moral facts cannot be tf-connected, then the debunker's genealogy of those beliefs implies that those beliefs are unjustified (assuming that this genealogy successfully implies that those beliefs cannot be e-connected).

Even if I have successfully established that the combination of evolutionary debunking arguments and this new argument defeats those beliefs we take to be about basic moral facts, it would leave those moral beliefs that we do not take to be about basic moral beliefs to be unaffected. But showing that we should withhold belief just in those moral facts that we take to be basic is a worthwhile conclusion in and of itself. The modal security response first points out that *if* our basic moral beliefs are true then they could not easily have been false, even assuming the evolutionary genealogy of our moral beliefs. *Under the assumption our beliefs about basic moral facts are true*, this implies that our beliefs in basic moral facts are trivially sensitive. The first step of the modal security response therefore assumes that all our beliefs in basic moral facts are true. If those beliefs that we take to be basic have in fact been defeated, then this first step of the modal security response is void.

The moral facts assumed by third-factor repliers are not explicitly basic, but some of them seem like textbook candidates for basic moral facts. For example, Skarsuane's claim that pain is bad and Wielenberg's claim that certain cognitive capacities entails having rights are both examples of moral facts that may be most plausibly viewed as basic. If Skarsuane and Wielenberg take these facts to be basic, then (as a result of the argument just given) they cannot be assumed in response to debunking arguments, and these third-factor replies fail. If they do not take them to be basic, there is some pressure on these commentators to explain which *more* basic moral facts could possibly explain these ones. If no candidates are forthcoming, then these commentators ought to take these facts to be basic and they therefore ought to withhold belief in those facts. By showing that third-factor repliers cannot rely on beliefs in moral facts that they take to be basic, we at least make their job harder; we restrict the moral beliefs on which they can rely, and it may be difficult to find moral facts that

can play the appropriate role of ensuring a tf-connection between our moral beliefs and the moral facts whilst also being plausibly viewed as non-basic.

So by arguing that the evolutionary debunking arguments can be supplemented to show that they defeat our beliefs in facts that we take to be basic, we deal a considerable blow to minimalist responses to debunking arguments. In the next section, I will argue for the more ambitious claim that, if belief in facts that we take to be basic are defeated, beliefs in facts that we take to be non-basic are also defeated. This implies that minimalist responses fail, regardless of whether they rely on moral beliefs that are taken to be basic or non-basic.

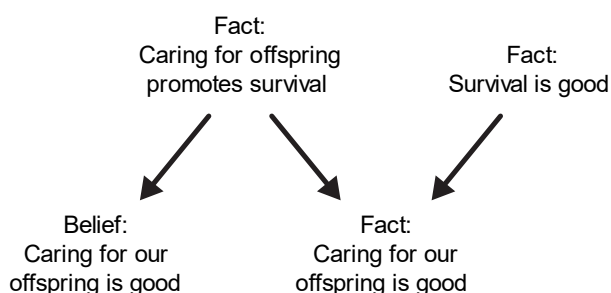
5.6: The First Non-Basic Argument:

In the next two sections I will provide two separate arguments that both imply that, if our belief in moral facts that we take to be basic are defeated, then so are our beliefs in moral facts we take to be non-basic. I take both these arguments to be independently sufficient for establishing this conclusion, and I include them both in case the reader is left unpersuaded by either one taken individually.

In the previous section I argued that there is good reason to think that our beliefs in realistically construed basic moral facts cannot be tf-connected. If we accept that evolutionary debunking arguments do in fact establish that those beliefs cannot be e-connected, our beliefs in basic moral facts are therefore unjustified because they fail (IEC). This implies that the minimalist ought to withhold her basic moral beliefs¹⁰⁴.

In order to justifiably believe that some non-basic moral fact is tf-connected, we cannot withhold belief in some more-basic moral fact that explains this fact in combination with the relevant

Figure 14:



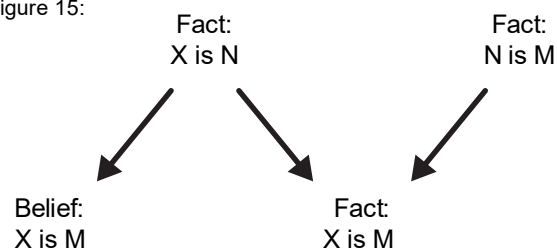
natural fact. For example, in order to believe that our belief in the goodness of caring for our offspring is tf-connected *via* the fact that caring for our offspring promotes

¹⁰⁴ It may be possible that there are cases in which our belief in P is not justified but it is not the case that we ought to give up our belief in P. See, for example, Bergman (2006: 424-7). Still, these are special cases, and in the current context there is no reason to think that we ought not give up our basic moral beliefs even if those beliefs are unjustified.

survival, as outlined in figure 14, we cannot withhold belief that survival is good. If survival is not good, then the fact that caring for our offspring promotes survival would not explain the fact that caring for our offspring is good, and so our belief could not be tf-connected *via* the fact that caring for our offspring promotes survival. If we withhold belief that survival is good, we should therefore withhold belief in the explanatory structure outlined in figure 14.

A parallel point applies to *all* beliefs in non-basic moral facts that we take to be tf-connected. In order to believe that it is tf-connected via some natural fact, we cannot withhold belief in the more

Figure 15:



basic moral fact that joins the relevant natural property with the relevant moral property. Relating this to figure 15, in which X and N denote natural properties and M a moral property, we can only justifiably believe that the fact that X is N explains our belief that X is M and the fact that X is M if we do not withhold belief that N is M. If N were not M, then the fact that X is N would not explain the fact that X is M. Call the fact that N is M the “bridging principle”¹⁰⁵. If what I have just claimed is true, then we can only justifiably take our moral belief to be tf-connected if we do not withhold belief in the relevant bridging principle that is necessary for that tf-connection to obtain.

This implies that, in order to justifiably believe that our non-basic belief is tf-connected, we cannot withhold belief in some other, more basic moral fact that acts as a bridging principle and that also explains this belief. If we take this bridging principle to be basic, then we ought to withhold belief in it for the reasons outlined in the previous section. If we believe that the goodness of survival is a basic moral fact, then we ought to withhold belief that it is e-connected and tf-connected, and this belief is unjustified. If, as I have just suggested, withholding belief that survival is good implies that we have no justification for believing that our belief in the goodness of looking

¹⁰⁵ This terminology is taken from Tersmann (2017: 765).

after our offspring is tf-connected, then we ought to withhold belief that our belief in the goodness of caring for our offspring is tf-connected¹⁰⁶.

What's more, we are taking for granted that none of our moral beliefs are e-connected due to the debunker's evolutionary genealogy. This implies that, when some moral fact we take to be non-basic can only be tf-connected given the truth of some bridging principle that we take to be basic, we ought to withhold belief that this non-basic moral belief is both e-connected and tf-connected. According to (IEC), these beliefs that we take to be about non-basic moral facts are therefore defeated. This shows that, if our beliefs in facts that we take to be basic are defeated then so are some of our beliefs in facts that we take to be non-basic.

So far, though, this line of argument only rules out justification for those non-basic moral beliefs that can we take to be tf-connected in virtue of some *basic* bridging principle. But there are some non-basic moral beliefs that can be tf-connected in virtue of some non-basic bridging principle. For example, it may be the case that the goodness of survival is actually a non-basic moral fact, itself explained by some more basic fact. If so, then the argument just given does not imply that we ought to withhold belief that the goodness of looking after our offspring is tf-connected, because the argument that we should withhold belief in moral facts that we take to be basic does not imply we should withhold belief in the goodness of survival. We therefore do not need to withhold belief in the bridging principle in virtue of which the explanatory structure represented in figure 14 can obtain, and so we do not need to withhold belief in this explanatory structure.

However, the sceptical consequences do ultimately spread to these lower rungs of non-basic moral beliefs as well. If the relevant bridging principle is non-basic, then we can only believe in this moral fact if we do not withhold belief it is e-connected and tf-connected. The debunker has established that we should withhold belief that it is e-connect; can this non-basic bridging principle be tf-connected? Only if we do not withhold belief in some further, more-basic bridging principle. And justified belief in this further, more basic bridging principle requires our not withholding it to be either e-connected or tf-connected. It is easy to see that eventually we will arrive at a non-

¹⁰⁶ Again, I assume that if our belief in this tf-connection is unjustified then we out to withhold belief in it.

basic bridging principle that can only be tf-connected via some basic bridging principle in which I have already shown we should withhold belief. The non-basic moral belief thereby gets defeated, and so does the less basic belief for which this one acted as a bridging principle. This process continues until justification for all those beliefs taken to be non-basic has been undone.

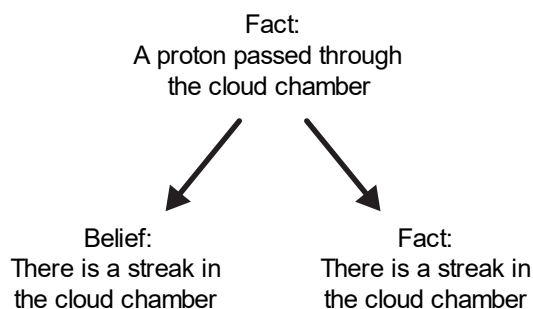
This implies that, if our beliefs in moral facts that we take to be basic are unjustified, then so are our beliefs in moral facts we take to be non-basic. This shows that minimalist responses fail if they rely on beliefs that the minimalist takes to be non-basic, as well as if they rely on beliefs that they take to be basic.

A crucial part of this argument is the assumption that, if we withhold belief in the relevant bridging principle, we should also withhold belief that our belief is tf-connected. If we withhold belief that survival is good, we should also withhold belief that our belief that we ought to care for our offspring is tf-connected *via* the fact that caring for our offspring promotes survival. One could therefore object to this argument by resisting this assumption; perhaps we could be justified in believing our belief is tf-connected in this way even if we withheld belief that survival is good.

This raises the following question: when *should* we withhold belief that our belief is tf-connected? Perhaps we only need to withhold belief that our belief is tf-connected if we actively *disbelieve* the relevant bridging principle. If we merely remain agnostic about whether survival is good then we need not withhold belief that our belief that caring for our offspring is good is tf-connected via the fact that survival is good, and it is only once we actively believe that survival is *not* good that we ought to withhold belief in this structure.

In order to see why this is false, consider Figure 16:

the following example: Shaquira is told that a proton has just passed through the cloud chamber. She forms the belief that there is a streak in the cloud chamber, because she believes that protons cause streaks when they pass through cloud



chambers. She takes her belief to be tf-connected in the way represented by figure 16.

But then she starts to doubt this belief; do protons genuinely cause streaks when they pass through cloud chambers, or has she made this up? She seems to remember being told they do, but she cannot say for certain whether this apparent memory is reliable. In the end she decides to remain agnostic about whether protons cause streaks when they pass through cloud chambers.

I think it is clear that, having withheld belief that protons cause streaks when they pass through cloud chambers, Shaquira ought to withhold belief that the explanatory structure outlined in figure 16 holds. Because she is uncertain that protons do cause streaks when they pass through cloud chambers, she should be equally uncertain whether her belief that there is a streak in the cloud chamber is tf-connected in the way represented by figure 16, because this can only occur if protons do cause streaks in cloud chambers.

In this example, Shaquira's initial belief that protons cause streaks in cloud chambers acts as a bridging principle in virtue of which her belief that there is currently a streak in the cloud chamber could be tf-connected via the proton passing through the cloud chamber. The fact that she ought to withhold belief that this belief is tf-connected when she withholds belief that this bridging principle holds implies that, when it comes to our moral beliefs, we ought to withhold belief that they are tf-connected if we withhold belief in the relevant moral bridging principle. Just as Shaquira did not need to actively believe that streaks do *not* cause streaks in cloud chambers, neither does the realist have to actively believe that the relevant moral bridging principle is false.

I think that, on reflection, this result is unsurprising. As I have explained, a moral belief can only be tf-connected if the relevant bridging principle is true. And, generally speaking, if we withhold belief in some fact F, and we take F to be necessary for some other fact H to obtain, then it seems we should also withhold belief in H. As it is necessary for our belief being tf-connected that the relevant bridging principle holds, we ought to withhold belief in that tf-connection if we withhold belief in the bridging principle.

This explains why third-factor repliers all assume that the relevant bridging principle is true in the process of arguing that our beliefs can be tf-connected. They do not

merely point out the survival *may* be good, or pain *may* be bad, but they assume that survival *is* good and pain *is* bad. If they could make their argument as effectively without these substantive assumptions they would surely strengthen their position, as a large portion of the criticism they face is due to their reliance on such moral claims.

All this indicates that, if we believe our moral belief is tf-connected then we cannot withhold belief in the relevant bridging principle. If we withhold belief in this principle, *either* by remaining agnostic about it *or* by disbelieving it, then we ought to withhold belief in the tf-connection. As I have argued, given (IEC), this implies that our non-basic moral beliefs are defeated if our basic moral beliefs are defeated.

5.7: The Second Non-Basic Argument:

I will now present my second argument for the claim that, if our beliefs in basic moral facts are unjustified then so are our beliefs in non-basic moral facts. This argument stems from the claim that, if we have a belief in some moral fact that we take to be non-basic, we must take the existence of this moral fact to depend on some more-basic moral fact. Non-basic moral facts just *are* those moral facts that are explained by other moral facts, and as such we cannot take some moral fact to be non-basic without taking the existence of that fact to depend on some other, more-basic moral fact. If we believe in some moral fact and we do not take its existence to depend on some more-basic moral fact then we do not take this moral fact to be non-basic.

Our moral beliefs in facts that we take to be non-basic therefore commit us to a belief in the further moral fact that explains the non-basic moral fact. We cannot justifiably believe (i: that looking after our children is morally good), (ii: that looking after our children is good because survival is good), if we withhold belief (iii: that survival is morally good). If we believe some moral fact obtains in virtue of another moral fact, then we are rationally committed to believing in that other moral fact. If we withhold belief in the more-basic moral this fact, we are no longer justified in believing in the non-basic fact that we took it to explain.

The further moral fact (which explains the non-basic fact) can be either basic or non-basic. Say Enoch takes the fact that survival is good to be basic. If we withhold belief

that survival is good, then our belief in the fact that we took to be explained by this fact (that caring for our children is morally good) also becomes unjustified¹⁰⁷.

If the more-basic moral fact that directly explains this non-basic fact is itself a non-basic fact, then this more-basic fact must also be explained by another moral fact (for this is what it means to be non-basic). We must continue up this explanatory ladder until we reach a basic moral fact that is ultimately responsible for all subsequent non-basic facts. As we should withhold beliefs in all moral facts that we take to be basic, this involves withholding belief in the very thing that we take to be responsible for all non-basic moral facts, and our beliefs in these non-basic moral facts therefore become unjustified.

This argument differs from the previous one I gave for this same conclusion. That argument depends on the truth of (IEC), as well as the claim that we can only justifiably take our non-basic moral beliefs to be tf-connected as long as we do not withhold belief in the relevant bridging principle. This argument depends on no such claims¹⁰⁸: it just assumes that, if we take P to depend for its existence on some other fact Q, then if we withhold belief in Q we must also withhold belief in P. This holds independently of the truth of (IEC), and independently of my claim about when we ought to withhold belief in a belief of ours being tf-connected.

Importantly, the argument of this section does not presuppose that our beliefs in non-basic moral beliefs are inferred on the basis of beliefs in moral facts that we take to be basic. In order to see this, consider the following example. I may believe you will arrive in London at 1pm. I may also believe that, in order for you to arrive at London

¹⁰⁷ The argument I presented in the previous section shows that we should withhold belief in all basic moral facts, *whatever* their content. We should therefore withhold belief in some basic moral fact even if we know nothing of its specific content. My argument for withholding any belief in a basic moral fact did not depend on the particular content of the basic fact in question, and so ignorance of this content should not shield our belief in that fact from defeat. This implies that we should withhold belief in moral facts we take to be non-basic even if we have no particular belief about *which* basic moral fact it is explained by. So long as we take a moral fact to be non-basic (and therefore explained by *some* basic moral fact), we know that we ought to withhold belief in the basic fact that explains it, and our belief in the subsequent non-basic fact is defeated.

¹⁰⁸ Here, I am referring to the arguments of this section specifically that are intended to achieve the conditional conclusion that if our belief in facts that we take to be basic are defeated then so are our beliefs in non-basic moral facts. But clearly, my previous argument that our beliefs in basic moral facts are defeated does depend on (IEC). So I do not mean to imply that we can ultimately do without (IEC) when arguing against minimalist replies.

at 1pm, you must have left Birmingham at 11:30am. On the basis of these two beliefs, I might form the belief that you left Birmingham at 11:30am. If I then withhold belief that you left Birmingham at 11:30, my belief that you will arrive in London at 1pm becomes unjustified (as long as I still believe that your arrival in London at 1pm is dependent on your leaving Birmingham at 11:30am). This is true even though I did not infer my belief that you will arrive in the station at 1pm from my belief that you left Birmingham at 11:30.

This implies that if I take P to depend on Q, then if I withhold belief in Q then my belief in P becomes unjustified even if I did not infer my belief in P from my belief in Q. Applied to our moral beliefs, this implies my beliefs in moral facts that I take to be non-basic become unjustified if I take their existence to depend on basic moral facts and I withhold belief in basic moral facts, and that this is true *regardless* of whether or not my beliefs in moral facts I take to be non-basic are inferred from my beliefs in moral facts I take to be basic.

An analogy with mathematics can establish a similar point. Say we have a number of non-axiomatic mathematical beliefs. Call them non-basic mathematical beliefs. Say we then infer, on the basis of these beliefs, beliefs in certain mathematical axioms, and we take the truth of these axioms to explain the truth of our non-basic mathematical beliefs. In fact, we believe that if these axioms were not true, then our non-basic mathematical beliefs could not be true either.

Surely, if we were given decisive reason to withhold belief in these mathematical axioms, we could no longer justifiably hold our non-basic mathematical beliefs. How could we continue to justifiably believe in these non-basic truths, whilst also believing that these mathematical truths depend on the axioms and withholding belief in these axioms?

This concludes my second argument for the conditional claim that, if our beliefs in moral facts that we take to be basic have been defeated, then beliefs in moral facts that we take to be non-basic have been defeated as well. In the next section, I will combine this argument with that of the previous section, and argue that this ultimately implies that minimalist responses to debunking arguments fail.

5.8: Minimalist Replies to Debunking Arguments Fail:

In section 5.5 I argued that our realistically construed basic moral beliefs cannot be tf-connected. The combination of this argument and evolutionary debunking arguments imply that our beliefs in moral facts that we take to be basic are unjustified (assuming that evolutionary debunking arguments succeed in establishing those beliefs are not e-connected). Next, I provided two arguments for the conditional claim that if our beliefs in moral facts that we take to be basic are unjustified, then so are our beliefs in moral facts that we take to be non-basic. Together, these two arguments imply that our beliefs in moral facts that we take to be basic and in those facts we take to be non-basic are unjustified.

This implies we are no longer justified in believing in moral facts that we take to be *either* basic or non-basic. If we believe that our moral belief in P must be either about a basic or a non-basic moral fact, then we must view it as being on the one hand neither e-connected or tf-connected and so defeated, or on the other defeated because our beliefs in facts that we take to be basic are defeated. Thus, if we view a given belief of ours as being about a moral fact that is either basic or non-basic, our belief in that fact is defeated.

Finally, we can point out that my distinction between basic and non-basic facts is exhaustive; a moral fact must fall in either category. Either a moral fact is not explained by some other moral fact (in which case it is basic) or it is explained by some other moral fact (in which case it is non-basic). No moral fact could be neither explained nor unexplained. This implies that all of our moral beliefs are about moral facts that are either basic or non-basic¹⁰⁹. Given my arguments in the previous section, this implies that the minimalist should take all of their moral beliefs to be either about basic moral facts (and defeated) or about non-basic moral facts (and defeated)¹¹⁰. This

¹⁰⁹ Of course, there is a third option that they are not about any moral facts because they are false. But, obviously, this option cannot be endorsed by the minimalist.

¹¹⁰ My argument implies that realistically construed beliefs that we take to be either basic or non-basic are defeated, and I have argued that minimalists ought to view all their moral beliefs as either basic or non-basic. There may be a worry that the realist who has not been exposed to this argument and remains unreflective about the metaphysical status of their moral beliefs will remain unaffected by this argument. But, as I argued in chapter one, evolutionary debunking arguments are most compelling when they are framed in terms of what the realist accepts or have reason to accept; I am not trying to present

implies that *all* of the minimalist's moral beliefs have been defeated by evolutionary debunking arguments which have been combined with the arguments of this chapter.

I refer to EDAs which are expanded to include the arguments presented in this chapter as “*EDAs”. Given (IEC), the minimalist cannot legitimately rely on substantive moral claims when responding to *EDAs. The minimalist accepts that evolutionary debunking arguments establish that our moral beliefs are not e-connected, and I have shown that, with this assumption in place, additional arguments can be employed to show that all our realistically construed moral beliefs are defeated. Even if we grant that those beliefs were justified prior to these arguments, this justification is defeated as a result of such arguments. Minimalist replies therefore fail because they rely on substantive moral beliefs that have been defeated.

I have suggested in previous chapters that the minimalist is allowed to rely on their moral beliefs until the claim that is supposed to defeat those beliefs has been established. Given what I argued in chapter 4, the claim that defeats their beliefs is that they fail (IEC). Thus, minimalists are allowed to employ their moral beliefs if, by doing so, they can show that those beliefs are either e-connected or tf-connected, because they would then be relying on their moral beliefs to show that the claim that is supposed to defeat those beliefs is not true.

But if what I have argued is correct, then *even if* (while relying on moral claims) the third-factor replier can show that some of their moral beliefs are tf-connected, they cannot show this for their basic moral beliefs. Their basic moral beliefs are incapable of such vindication and so, according to (IEC), they are defeated. Furthermore, once beliefs in basic moral facts are off the table, the minimalist cannot resist my arguments that their non-basic moral beliefs are therefore defeated even whilst relying on their non-basic moral beliefs. So my arguments do not assume that the minimalist's beliefs have been defeated in the process of arguing that they have been defeated; the

an argument that will be effective against all moral realists independently of what they have been given reason to believe. As such, we should not be worried about the possibility of an unexposed realist who is not aware of my arguments and does not have any opinion about the metaphysical status of their beliefs.

arguments that are supposed to defeat those beliefs work even if we grant the minimalist *prima facie* justification for those beliefs.

There is a related concern that, so far, I have only shown that the minimalist *ought* to withhold belief that their basic moral beliefs are e-connected and tf-connected. From this it follows that their non-basic moral beliefs are defeated as well. But (IEC) does not say that our belief is defeated if we *ought* to withhold belief that it is e-connected and tf-connected; rather, it stipulates that our belief is only defeated if we actually withhold belief that it is e-connected or tf-connected. So there remains the theoretical possibility that a minimalist could accept that they ought to withhold belief that their basic moral beliefs are e-connected and that they are tf-connected, but steadfastly refuse to adopt this attitude. By doing so they avoid the attitude that defeats their basic moral beliefs and my argument cannot go through.

I doubt that this is a particularly attractive response to my argument. The reason is that, even if this could allow the minimalist to escape the problem that I have been pressing (that they problematically rely on defeated moral beliefs when making their argument) their position would become criticisable for another reason. They would be criticisable for accepting that they ought to withhold belief that their basic moral beliefs are e-connected and tf-connected but then failing to adopt this attitude. So the minimalist would have just re-located the problem with their position elsewhere. In order to iron out this new problem, they would presumably have to adopt the attitude that I have argued defeats all their moral beliefs, and my original argument goes through.

Of course, this is not to say that the minimalist *has* to accept the force of the arguments that I have presented. They might, for example, deny the truth of (IEC), or reject the structure of normative reality outlined in section 1. Though I think that the arguments in favour of each of these positions is strong, I invite the minimalist to engage with these arguments in order to try to establish where I have gone wrong. My point is simply that the minimalist cannot accept the legitimacy of such arguments and then escape defeat merely by refusing to adopt the attitude that they acknowledge they ought to accept.

If what I have argued for in this chapter is correct, then EDA's do not by themselves defeat the realist's moral beliefs. In order to defeat the realist's moral beliefs, EDAs have to be bolstered with the additional arguments presented in this chapter, thus morphing them into an *EDA. Minimalist replies, strictly speaking, object to EDAs rather than *EDAs, and so in a sense they succeed against the debunking arguments to which they were responding. One might object, then, that I have not shown that minimalist replies fail, because they were intended as responses to EDA's. I have only shown that they fail as responses to *EDAs.

But this is a hollow victory for the minimalist. Even if we concede that minimalist replies were successful as responses to EDAs, if what I have argued is correct, there is a genuinely valid argument against which minimalist replies are unsuccessful. All the debunker needs to do is adopt this new argument to neutralise any threat posed by minimalist replies. Even if I have not, strictly speaking, shown that minimalist replies fail, I have at least removed the significance of their success by presenting an alternative, equally valid argument against which they cannot be effectively deployed.

I will end this chapter by considering two more potential objections from the minimalist. They both involve ways in which the realist might be able to respond to the lack of an explanatory connection between our basic moral beliefs and the facts that make them true. The first is to accept that two correlate by chance, while stomaching the unease of this position for the sake of realism's other benefits. The second is to cite the modal security of our basic moral beliefs. I will argue that neither of these options are tenable, and so neither present a way for the realist to cope with the explanatory disconnect between our basic moral beliefs and moral truth.

5.9: Just a Small Miracle:

One possible response to the arguments I have just presented is to accept that our beliefs in basic moral facts cannot be explanatorily related to the facts that make them true, but then to take them to be true by chance, as a kind of small miracle. Though positing this kind of unexplained coincidence might make us feel slightly uneasy, the benefits of realism elsewhere might more than make up for this small drawback of having to posit one unexplained correlation in our meta-ethical theory.

Having presented his third-factor response to debunking arguments, Enoch considers the following objection:

“On the story just told, had evolution "aimed" at something else, something that is not of value or perhaps is even of negative value, our normative beliefs would have been systematically mistaken. So doesn't it follow that the story just told, far from showing how the realist can avoid commitment to miraculous correlations, *relies* on a miracle?” (2010: 433)

What exactly is the miraculous correlation to which Enoch is referring? It is the correlation between the aim of evolution (i.e, survival) and what is evolution-independently morally good (also survival). Enoch states that for this correlation *not* to have occurred, either basic moral truths would have to have been different, or the aim of evolution would have to have been different (2010: 433).

This implies that, according to Enoch, the “aim” of evolution corresponds to a basic, metaphysically necessary moral truth. Given this correlation, we can tell a story about why our beliefs in non-basic moral facts and the non-basic moral facts themselves are systematically correlated: they are both explained by the same natural fact (hence, they are tf-connected).

But this story says nothing about why there is a correlation between the aim of evolution and basic moral truth. The account presupposes, but does not explain, this correlation. As such, the account relies on the “small miracle” that the aim of evolution is good. But by relying on this correlation between just two facts, we can explain the systematic correlation between our beliefs in non-basic moral facts and the non-basic moral facts themselves.

Referring to this explanatory gap, Enoch writes:

“Let me not give the impression that this suggested way of coping with the epistemological challenge is ideal. Indeed, because of the (perhaps) remaining small miracle, perhaps the realist loses some plausibility points here. But not, it seems to me, too many. Given the rules of the explanatory game - and given the other, independent, advantages of realism (which I don't discuss in this paper) - this conclusion may be enough for the realist to have satisfactorily addressed the epistemological challenge.” (2010: 435).

This response indicates how Enoch might respond to my claim that there can be no explanatory connection between basic moral facts and our belief in those facts. Enoch might be inclined to answer that, yes, according to his position, there is no such explanatory connection between the two, and as such it is difficult to see how we could explain the fact that the two correlate with each other. But, once we have posited this unexplained correlation, Enoch can explain how our other beliefs (in non-basic moral facts) might be systematically, non-accidentally reliable. And though positing this initial unexplained correlation might be difficult for us to stomach, this one drawback of the realist position is unlikely to render it less plausible than competing meta-ethical positions all-things-considered. The realist might therefore suggest that they can live with lack of an explanatory connection between their basic moral beliefs and moral truth, and that this (admittedly negative) feature of their account lacks any major significance when comparing meta-ethical theories.

Wielenberg also notes that, on his account, there is an unexplained coincidence between the moral principles that according to which evolution encourages us to reason and necessary moral truths (2014: 174). Wielenberg's response is to argue that theistic meta-ethical accounts will *also* involve unexplained coincidences between necessary truths. Thus, Wielenberg concludes, "there is a plausible case to be made that the unexplained necessary coincidences worry has equal force against my view and a theistic approach to meta-ethics, and hence there is no advantage for the theistic approach here." (2014: 175). This implies that Wielenberg might also be prepared to countenance an unexplained coincidence between basic moral truths and our beliefs about them, with the hope that realism's *other* features might make up for this small defect when it is compared with other meta-ethical positions all-things-considered.

In this thesis, I have made two positive arguments that are relevant to assessing the viability of this response. The first, argued for over the course of chapters three and four, is that if we withhold belief that our belief is e-connected and we withhold belief that it is tf-connected, then our belief in P is defeated. In other words, (IEC) is true. The second, argued for in sections 5.6 and 5.7 of this chapter, is that if the realist's beliefs in basic moral facts are defeated, then this ultimately implies that *all* the realist's moral beliefs are defeated.

With these arguments in place, we can see the true price of accepting that our basic moral beliefs can only be true as a result of a “small miracle”. Accepting that the realist’s belief in basic moral facts can only be true as a result of a chance alignment (rather than any explanatory relationship) implies, given my epistemological arguments in favour of (IEC), that none of those beliefs can be justified. And my arguments in the latter part of this chapter show that, once these beliefs in basic moral facts are unjustified, then so become the realist’s beliefs in non-basic moral facts as well. Thus, the arguments that I have defended in this thesis imply that the consequence of accepting a mere chance alignment between our basic moral beliefs and moral facts is that none of our moral beliefs can be justified.

Accepting that beliefs in basic moral facts can only be true as the result of a chance alignment therefore drastically alters the appeal of moral realism when it is weighed, all things considered, against its meta-ethical competitors. We do not only have to countenance the slight unease of just one unexplained correlation between our moral psychology and moral reality, as some of the quoted passages imply; we also have to countenance the devastating epistemological consequences of this attitude. When these consequences become clear, the chances that realism will be picked as the most all-round attractive meta-ethical theory becomes much lower, because the flaws with this theory that are far more extreme than initially recognised. I therefore reject the possibility that the realist can posit a correlation between our basic moral beliefs and basic moral truths as a result of a “small miracle”, with the hope of making up for this defect elsewhere.

5.10: Going Modal:

Some minimalists try to justify their beliefs in basic moral facts by noting the modal security of those beliefs. Enoch hints at this response in the following paragraph:

“Fundamental normative truths are presumably necessary in a fairly strong sense, or at the very least so we are entitled to assume in the context of critically evaluating the epistemological challenge to Robust Realism. So the main way in which the evolutionary “aim” (which is actually of value) could have failed to be of value is if evolution had a very different “aim”. But it’s not clear what to make of this suggestion: For surely, it’s not contingent that evolution has something to do, for instance, with survival and reproductive success rather than their opposites.” (2010: 433)

The thinking here, familiar from chapter two, is that, owing to their necessity, there are no close possible worlds in which basic moral truths might have been different. Furthermore, there are no close possible worlds in which evolution could have encouraged us to have different basic moral beliefs, because there are no close possible worlds in which evolution aims at something other than survival. As such, there is no sense in which it is a “miracle” that our basic moral beliefs align with the moral facts, and so there is nothing troubling about the lack of an explanatory connection between the two (2010: 433).

Wielenberg defends this argument more explicitly. First, Wielenberg notes that

“because the basic ethical facts are necessary truths, if there is any luck in the correspondence between our psychological dispositions and moral reality, it must lie entirely on the psychological side of the equation. Where there is no contingency, there is no luck. (2014, pg 167).

Whether or not there is any luck involved in the accuracy of our basic moral beliefs therefore depends on whether we are lucky to have those beliefs. Then, Wielenberg suggests, we are not lucky to have those beliefs, because evolution could not have produced beings with the psychology complexity required to have basic moral beliefs without encouraging those beings to have roughly the same basic moral beliefs that we have (2014: 170).

In other words, it is no accident that evolution encouraged us to have the basic moral beliefs that we have. Even if there is no explanatory connection between our basic moral beliefs and the moral facts, the laws of nature make it likely that beings with the capability to form any moral beliefs will have the moral beliefs that we have. So there is no contingency on the psychological side of the correlation between our basic moral beliefs and the moral facts. And, as we have already established, neither is there any on the side of basic moral facts. So there is no significant sense in which our basic moral beliefs are luckily correct, and there is no reason to think that the absence of an explanatory connection between the two should threaten our justification for these beliefs.

As I argued in chapter three, accepting that our belief in P is explanatorily unrelated to P defeats that belief regardless of the implications this has for the modal security of that belief. As such, even if the realist can successfully establish that there are no close

possible worlds in which the basic moral truths are different, and no close possible worlds in which the content of our beliefs about basic moral truths are different, they do not thereby establish that those beliefs are justified. My argument presented above, that such beliefs cannot be explanatorily related to each other, remains unaffected by such considerations. As I argued in chapter 3, this argument alone is enough to establish that such beliefs are defeated. The project of vindicating our beliefs in basic moral facts by showing that they are modally secure is therefore fundamentally misguided.

5.11: Conclusion:

My goal in this chapter has been to show how, contrary to appearances, (IEC) can ground an objection to minimalist replies to debunking arguments. We can supplement EDA's with further argumentation that establishes the minimalist should withhold belief that their basic moral beliefs are e-connected and tf-connected. According to (IEC) this defeats those beliefs, from which it follows that all the minimalist's moral beliefs are defeated. These beliefs therefore cannot be legitimately relied upon in response to *EDAs.

There are various ways that the realist might resist my argument and justify their reliance on substantive moral claims in response to debunking arguments. They might, for example, deny the debunker's genealogy of our moral beliefs. This could potentially allow for those beliefs being e-connected or tf-connected, which would allow them to pass (IEC). Or they could accept the debunker's proposed genealogy, but argue that this genealogy does not imply that those beliefs are not e-connected (Mogenson, 2016, Hanson, 2017).

Importantly, though, pursuing either of these strategies would presumably have to involve engaging a scientific debate about the empirical claims on which the debunker relies, and would therefore involve a very different kind of argument to the one usual presented by the minimalist. Minimalist responses are characterised by their granting the debunker's genealogy of our moral beliefs and their continued reliance on substantive moral claims in spite of this concession. The possibility of resisting my expanded EDA by rejecting the debunker's genealogy of our beliefs therefore signals the abandonment of the minimalist's position, rather than a means of defending it.

The important point is that, if the arguments presented in chapter are good, then the realist cannot respond to debunking arguments by accepting that their moral beliefs are not e-connected and then relying on their moral beliefs. For reasons independent of evolutionary debunking arguments, our beliefs in realistically construed basic moral facts cannot be tf-connected, so if the realist gives up the possibility that they are e-connected they thereby give up the possibility that they might be appropriately connected to the moral facts. And, once we give up these basic moral beliefs, our other moral beliefs should be quick to follow. So the realist cannot respond to debunking arguments by accepting that their moral beliefs are not e-connected and then relying on those moral beliefs, and, as these are the moves that characterise minimalist replies to debunking arguments, minimalist replies to debunking arguments fail.

I ended by considering two possible ways in which the minimalist might deal with the claim that our basic moral beliefs are explanatorily disconnected from the facts that make them true. In the next chapter will consider one more potential way in which the minimalist might argue our basic moral beliefs might be justified despite everything I have said so far. This involves the possibility that our basic moral beliefs might be justified in virtue of being inferred from our non-basic moral beliefs. Diagnosing why this response fails sheds more light on the kind of attitude we can have towards our belief such that the belief in question is no longer justified.

Chapter 6: Inferring From Our Non-Basic Beliefs:

6.1: Introduction:

In the previous chapter I argued that evolutionary debunking arguments can be expanded to imply that our basic moral beliefs are neither e-connected nor tf-connected under the assumption of moral realism. Given (IEC), this implies that, assuming moral realism, our beliefs in moral facts that we take to be basic are unjustified. I then argued that this implies that all our moral beliefs are unjustified under the assumption of moral realism. Minimalist replies to evolutionary debunking arguments therefore fail.

In this chapter I want to consider the following potential response from the minimalist. It is consistent with the debunker's evolutionary genealogy of our moral beliefs that our non-basic moral beliefs can be tf-connected. This implies that those non-basic moral beliefs can remain undefeated according to (IEC). But, if these non-basic moral beliefs can form the basis of our basic moral beliefs, then it seems we can justifiably infer these basic moral beliefs from our (tf-connected and therefore justified) non-basic moral beliefs. So maybe our basic moral beliefs can be justified because they can be inferred from our non-basic moral beliefs.

I begin by outlining this response in more detail. I then argue that, if our basic moral beliefs are inferred on the basis of non-basic moral beliefs, then (according to the structure of normative reality presented in the last chapter) our basic moral beliefs share an alternative explanatory connection with the moral facts that is distinct from either an e-connection or a tf-connection: they are a-connected. If those basic moral beliefs can thereby remain justified, this represents a counter-example to (IEC). The success of this response therefore threatens the legitimacy of (IEC), as well as my argument against minimalist replies to debunking arguments.

In section 6.3, I argue that the realist response under consideration fails. The problem, I argue, is that taking our non-basic beliefs to be tf-connected presupposes the very basic beliefs that we hope to infer from those non-basic beliefs, and so this attempted

vindication of our moral beliefs is problematically circular. This is true, I will argue, *regardless* of whatever basis we initially had for our non-basic moral beliefs. In section 6.4 I consider what my arguments imply about the conditions under which our beliefs are defeated, and I end by considering a potential response from the realist.

But, before we expand on this potential response to the arguments I presented in the previous chapter, it is important to keep in mind who those arguments were aimed at and where this potential response might come from. The arguments of the previous chapter were aimed at minimalist responses to debunking arguments that grant the explanatory disconnect implied by debunking arguments but argue that our moral beliefs can nonetheless remain justified. The response I am now considering is an objection to this argument, and hopes describe a way in which minimalist responses to debunking arguments might succeed. As a result, this response is an attempt to show how our moral beliefs might remain justified *even if* we have accepted the explanatory disconnect implied by debunking arguments, as this is one of the defining claims of the minimalist replies to debunking arguments that this response is hoping to vindicate. As I did in the last chapter, then, I will assume this imagined minimalist-defender grants that their moral beliefs are not explained by the moral facts. Unless they can show our moral beliefs to be justified even granting this assumption, they cannot show that minimalist response to debunking arguments succeed. I will now go into more detail about what this response from the minimalist might look like.

6.2: Leveraging Justification From Our Non-Basic Beliefs:

My distinction between basic and non-basic moral facts is a metaphysical one; basic moral facts are those that do not metaphysically depend on any other moral facts, and non-basic facts are those that do metaphysically depend on other moral facts. But, as I noted in the previous chapter, beliefs about basic moral facts need not be “basic” in the sense that they are the basis from which for all other moral beliefs are inferred. It is perfectly possible that we justifiably infer beliefs in moral facts that we take to be metaphysically basic from beliefs in facts that we take to be metaphysically non-basic.

And this is sometimes what happens. We might have a number of moral beliefs about highly specific actions and situations, and on the basis of such beliefs infer more and more general beliefs, until we arrive at what we take to be the most general facts about

what makes actions right or wrong. These general facts are those that we take to be metaphysically basic; they do not depend for their existence on any other moral facts in combination with some natural fact¹¹¹.

There seems to be no good reason for privileging more general moral beliefs as the only legitimate starting points of moral reasoning. Sarah McGrath (2019) presents the following argument for this point: prior to any philosophical investigation, people have a number of highly specific moral beliefs that qualify as knowledge (or, at least, as reasonable belief). And, when engaging in moral enquiry, we ought to employ any relevant knowledge (or reasonable beliefs) that we possess. Because our judgements regarding highly specific moral claims clearly *are* relevant to moral enquiry, we ought to utilise our highly specific moral beliefs when engaging in moral enquiry (McGrath, 2019: 51-2). Unless we are to implausibly deny that our everyday judgements about specific moral claims can qualify as knowledge, there is no reason to think that we cannot start with a number of beliefs in moral facts that we take to be non-basic, and then infer beliefs in more general, basic moral facts¹¹².

Let us assume that the moral claim that is posited by the third-factor replier is a basic moral fact. That is, assume that Enoch takes the goodness of survival to be basic¹¹³. It is possible that the realist's belief in this basic moral fact is inferred on the basis of their beliefs in non-basic facts. We might believe that we ought to look after our offspring, and that we ought not murder, and on the basis of these beliefs come to believe more generally that survival is morally good.

What's more, as the third-factor replier has correctly pointed out (and I have conceded in the previous chapter) traditional evolutionary debunking arguments do not imply that our moral beliefs cannot be tf-connected; they just show that our moral beliefs are not e-connected. If what I have argued previously is correct, then we have good reason

¹¹¹ I do not mean to imply that reasoning never occurs in the opposite direction; our moral general beliefs about basic moral facts will often inform our more specific beliefs about which actions are right and wrong. But it is at least possible that a given belief in a basic moral fact is inferred on the basis of beliefs in moral facts that one takes to be non-basic.

¹¹² For the opposing view, that we should privilege our judgements about general moral principles in moral enquiry, see Singer (1974, 2005).

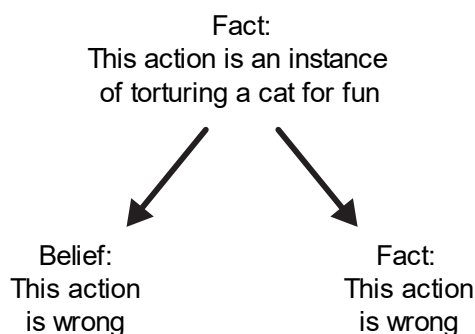
¹¹³ We can argue for the same conclusion even if the survival of goodness is not itself basic, and is itself ultimately explained by a separate basic moral fact. The argument is just slightly more complex and, since my comments apply equally to both, I will focus on the simplest case in which the goodness of survival is itself the basic moral fact.

to think that our basic moral beliefs cannot be tf-connected, and this (in combination with evolutionary debunking arguments) implies that those beliefs are defeated. Once we have established that these basic beliefs are defeated, this implies that our non-basic beliefs are defeated as well. But it seems that, *until* our basic beliefs have been defeated, there is no need to withhold belief that our non-basic beliefs are tf-connected, and our non-basic beliefs can therefore pass (IEC).

But, if prior to our basic beliefs being defeated our non-basic moral beliefs are undefeated, why can we not retain justification for our basic moral beliefs by inferring them from these justified non-basic beliefs? Surely inferring a belief in P from other justified beliefs is one way to justify that belief in P, and so why can we not leverage the justification we have for our non-basic beliefs that we take to be tf-connected in order to justifiably infer (and believe in) basic moral facts? If we can thereby retain justification for our basic moral beliefs, the argument against our non-basic moral beliefs (which crucially relied on the claim that our basic moral beliefs are defeated) does not go through, and the realist can retain justification for all her moral beliefs. The minimalist could therefore unproblematically rely on such beliefs when responding to the evolutionary debunker.

Though no minimalist I am aware of actually makes this argument, some of them seem primed to do so. For example, according to Wielenberg's account of moral psychology, our "verdictive" moral beliefs (i.e. "beliefs about the rightness or wrongness of particular acts performed by particular agents in particular circumstances" (2014: 108)) are formed by our unconsciously categorising the scene as having a number of natural properties, and some of these unconscious classifications trigger the verdictive moral belief. If we turn a corner and see a number of hoodlums abusing a cat for fun, our unconscious will categorise the scene as having the property of "torturing a cat for fun". This unconscious classification will, in turn, cause us to have the conscious belief that what the hoodlums are doing is morally wrong. When these verdictive beliefs are formed reliably, they are tf-connected because the natural properties that prompt our

Figure 17:



belief that the action in question is wrong do in fact explain that the action is wrong, as in figure 17.

The more general moral bridging principles that regulate these unconscious inferences from a scene having some natural property to the scene having some moral property are “hidden”, in the sense that they are often not consciously believed or utilised by the observer. Wielenberg notes “The judger may lack direct conscious access to the relevant general moral principles altogether” (2014: 98).

But Wielenberg goes on:

“However, it is important to realize that this leaves open the possibility that the judger can *infer* that her judgments conform to the m-possessed principle¹¹⁴ (*hidden* does not mean *unknowable*). That is not merely a theoretical possibility; there is empirical evidence that it actually sometimes occurs... despite the common use of terms like “rationalization” or “confabulation” to characterize people’s efforts to identify the principles to which their particular moral judgments conform, it is a mistake to suppose that such efforts are always unsuccessful. Sometimes such efforts or “rationalizations” succeed in uncovering the hidden moral principles that actually describe the patterns of our particular moral judgments.” (2014: 100)

Thus, according to Wielenberg, when our conscious verdictive moral beliefs are produced reliably, they are tf-connected. What’s more, it is possible for us to infer from moral beliefs about specific actions more general moral beliefs in the principles to which our unconscious inferences adhere. It therefore seems that Wielenberg’s account of moral psychology is amenable to the response that I have presented in this section: why can’t our beliefs in basic moral facts be inferred from tf-connected (and thereby justified) beliefs in non-basic moral facts? And if they can be, why is this inference from a number of justified non-basic beliefs not enough for the basic beliefs to themselves be justified?

6.3: (IEC) and Attenuated Explanatory Links:

The first thing to note about this reply is that, if it succeeds, it entails that (IEC) is false. This is because I have argued, in the previous chapter, that our beliefs in basic

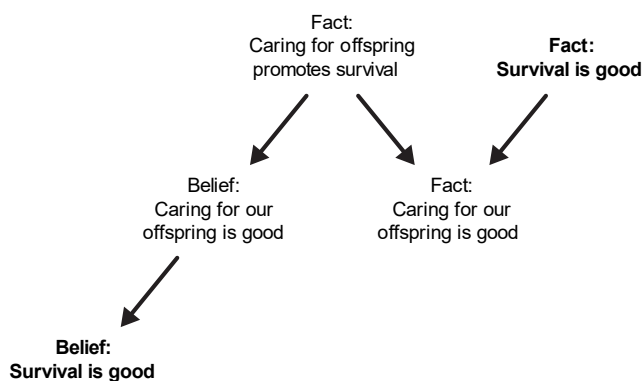
¹¹⁴ Here Wielenberg is referring to the more basic bridging principle that regulates our verdictive beliefs.

moral facts cannot be e-connected or tf-connected. This follows from the debunker's genealogy of our beliefs (which minimalists grant for the sake of argument) in combination with the supposition that basic moral truths are unexplained. The current suggestion is that our beliefs in basic moral facts can still be justified because they can be inferred from our tf-connected non-basic beliefs. But this suggestion does not imply that our basic moral beliefs can in fact be e-connected or tf-connected; it does not challenge either the debunker's genealogy of our moral beliefs (which implies our beliefs in basic moral facts cannot be e-connected) or my claim that realistically construed basic moral facts are unexplained (which implies that they cannot be tf-connected). Rather, it just suggests that the debunker's basic moral beliefs might still be justified even though, presumably, the reasons for thinking they cannot be e-connected or tf-connected still stand. As such, if this response is successful, it implies that (contrary to what I argued in chapter 4) our belief can still be justified even if we withhold belief that it is both e-connected and tf-connected, and (IEC) is false.

In fact, this response draws our attention to an alternative explanatory connection, other than e-connection or tf-connection, that might hold between our belief in P and P. Say our belief that we ought to care for our offspring is non-basic, and from this belief we infer our belief in the basic moral fact that survival is good. Under the assumption that, if our belief in P is based on Q then Q explains our belief, then this implies our non-basic belief in the goodness of caring for our offspring explains our basic belief that survival is good. Furthermore, if our non-basic belief that caring for our offspring is good is explained by the fact that caring for our offspring promotes survival, and our basic belief that survival is good is true, then the explanatory structure represented in figure 18 holds for our beliefs.

According to the explanatory structure outlined in figure 18, our belief in the goodness of survival does in fact share some kind of explanatory relationship with the fact that makes it true. This explanatory connection is more complex than either an e-connection or a tf-connection. If our belief in P is e-connected then P explains our

Figure 18:



belief in P, and our belief in P is tf-connected if it is explained by some fact that explains our belief in P. According to the explanatory structure outlined in figure 18, our belief that survival is good is explained by some fact that explains a fact which is explained by the fact that makes it true. We can refer to this kind of connection as an “attenuated” explanatory relationship¹¹⁵. When our belief in P shares this attenuated relationship with the facts that make it true, it is “a-connected”.

Now, (IEC) says:

(IEC): If (i) S withholds belief that her belief that P is explained by the fact that P and (ii) S withholds belief that there is some fact that explains both her belief that P and the fact that P, then S’s belief that P is unjustified.

So according to (IEC), taking our belief to be a-connected does not allow the minimalist’s basic moral beliefs to remain justified. As I have pointed out, the explanatory link between the belief in the goodness of survival and the fact that survival is good outlined in figure 18 is neither an e-connection nor a tf-connection. The possibility of this kind of explanatory connection therefore does not allow the realist to retain justification for these basic moral beliefs, given (IEC).

This implies that if the realist response under consideration succeeds, then it presents a counter-example to (IEC). It shows that our basic moral beliefs that we take to be a-connected, but not e-connected or tf-connected, can remain justified. We should then reject (IEC) in favour of a principle that allows beliefs to remain justified if we do not withhold belief in this more attenuated explanatory connection to the relevant fact, just as we rejected (EC) in favour of a principle that allows beliefs to remain justified if we take them to be tf-connected.

In the remainder of this chapter I will argue that this realist response fails, and that our basic moral beliefs cannot be justified in virtue of being inferred from non-basic beliefs that we take to be tf-connected, even if this does imply that those beliefs are a-

¹¹⁵ True, this attenuated relationship is built out of more simple e-connections and tf-connections. But this does not imply that a-connections do not constitute a distinct class of explanatory connections, any more than the fact that tf-connections are built out of e-connections implies tf-connections are not distinct from e-connections. Furthermore, if a-connections were not distinct from e-connections or tf-connections then my argument that our basic moral beliefs cannot be e-connected or tf-connected would also rule out the possibility that they are a-connected. That it does not rule out this possibility is evidence that the a-connections are genuinely distinct from e-connections or tf-connections, despite being comprised of them.

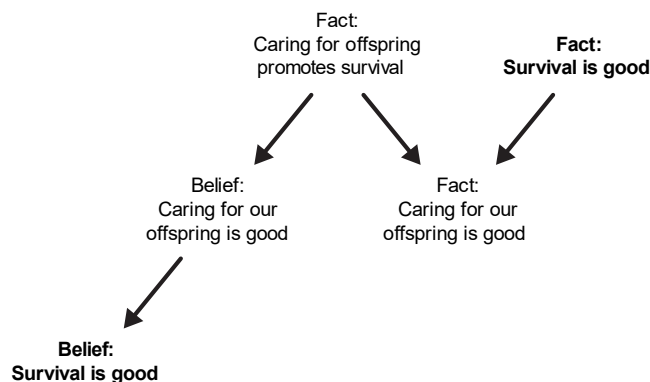
connected. This vindicates my expanded evolutionary debunking argument against the potential response from the minimalist, and neutralises the threat it poses to (IEC) as a legitimate constraint on belief.

6.4: Why This Response Fails:

In this section I will argue that realist response under consideration fails because it implies that the truth of our basic moral beliefs are presupposed in the process of justifying our non-basic moral beliefs, and our basic moral beliefs are then inferred from these very non-basic moral beliefs. This implies that justification for both our non-basic and our basic moral beliefs becomes problematically circular, and so this strategy does not succeed in vindicating our realistically construed moral beliefs.

According to the realist response Figure 18:

under consideration, the realist accepts the debunker's evolutionary genealogy of our moral beliefs which implies they are not e-connected, but argues our non-basic moral beliefs are nonetheless justified because they



are tf-connected. I argued in the previous chapter that we cannot take our non-basic belief to be tf-connected if we withhold belief in the relevant bridging principle in virtue of which it is tf-connected. That is, we cannot take our belief that caring for our offspring is good to be tf-connected via the fact that looking after our offspring promotes survival if we withhold belief that survival is good.

Let's say the realist does not withhold belief in this bridging principle, and she takes her non-basic belief to be tf-connected in virtue of this bridging principle. The suggestion is that these non-basic moral beliefs can therefore be justified, and they can form the basis of our belief in the basic moral fact that survival is good¹¹⁶. Hence the explanatory structure outlined in figure 18.

¹¹⁶ Below, I will consider the possibility that we could infer basic moral beliefs *other* than the specific bridging principle that is required for the non-basic beliefs to be tf-connected.

I think that this response fails because the realist presupposes the truth of the relevant bridging principle (that survival is good) in the process of justifying her non-basic belief (that looking after our offspring is good), and her belief in this bridging principle is based on this very non-basic belief.

The realist's non-basic belief (that caring for offspring is good) is only justified, given the debunker's evolutionary genealogy of our moral beliefs, because she takes there to be some third-factor (caring for our offspring promotes survival) that explains both her belief and the fact that makes it true. But she can only take her non-basic belief to be tf-connected in this way if she presupposes the basic moral fact that survival is good; if survival is not good, then the fact that caring for our offspring promotes survival does not explain the fact that caring for our offspring is good, and her non-basic belief that caring for our offspring is good is not tf-connected. As such, the realist's justification for her non-basic belief (that caring for our offspring is good) is dependent on her belief in the relevant bridging principle that survival is good. If she does not believe in this bridging principle, she ought to withhold belief that her belief that she ought to look after our offspring is tf-connected, and, because she grants that this belief is not e-connected, this belief ceases to be justified.

But, according to the suggestion under consideration, the realist's belief in this very bridging principle (that survival is good) is only justified on the basis of her belief that looking after our offspring is good. This is how the imagined minimalist proposes to retain justification for her basic moral beliefs even after I have established, in the previous chapter, that those beliefs cannot be e-connected or tf-connected. As such, her justification for her non-basic belief depends on her belief in the relevant bridging principle, and her justification for her belief in the relevant bridging principle depends on her non-basic belief. This implies that her reasoning for accepting these beliefs is problematically circular, and that neither belief can be vindicated as a result of this process.

We can see this problem most clearly if we imagine that the realist forms her non-basic and basic beliefs by explicitly reasoning in a way that reflects the explanatory structure outlined in figure 18. That is, say she believes that some act promotes survival, infers from this belief that the act is morally good and, on the basis of her

beliefs that this act both promotes survival and is morally good, she infers that actions that promote survival are morally good.

The problem with this line of reasoning is clear: the inference from the belief that the action promotes survival to the belief that the action is good *presupposes* that actions that promote survival are morally good. She only has reason to think that the action in question is morally good under the assumption that promoting survival is morally good, and so she cannot then base her belief in the goodness of survival on her belief that this action is morally good. If she does, the reasoning for both her belief that a particular action is good and her belief that survival is good become problematically circular and neither belief can be justified.

Unfortunately for the realist, I think that this strategy fails *regardless* of whatever basis the realist initially had for her non-basic beliefs. In order to see why, it is important to remind ourselves that we are assuming that the realist accepts (in virtue of the debunker's proposed genealogy of our moral beliefs) that our moral beliefs are not e-connected. This is the concession that characterises minimalist responses to debunking arguments, and we are currently considering the possibility that the realist can argue that, despite the fact that none of their moral beliefs are e-connected, their non-basic beliefs can be tf-connected and can therefore remain justified. These tf-connected beliefs can then form the basis of their basic moral beliefs, and justification for these basic moral beliefs has been salvaged. Assume, for now, that the realist does not take there to be any *other* tf-connection between their non-basic moral beliefs and the non-basic moral facts, and takes them to be only tf-connected via that fact that survival is good, as in figure 18.¹¹⁷

With this in mind, imagine that the realist initially formed her non-basic moral beliefs on the basis of her intuitions that the actions in question were right or wrong. She believes that she ought to look after her offspring on the basis of her intuitive sense that this is the case. We are now supposing that she accepts that none of her non-basic beliefs are e-connected. This implies that none of her intuitions, on which she bases

¹¹⁷ Below, I consider the possibility that there is some alternative tf-connection between our non-basic beliefs and the non-basic facts.

on her non-basic beliefs, are explained by the non-basic facts¹¹⁸. If they were, then this would imply that her beliefs were explained by the moral facts because the moral facts would explain her intuitions, which would explain her beliefs. Assuming explanatory transitivity holds, this would imply that her moral beliefs are explained by the moral facts. As such, when she accepts that her non-basic moral beliefs are *not* e-connected, she can no longer take her intuitions (on which she bases her non-basic beliefs) to be explained by the relevant facts. In short, the minimalist cannot take her beliefs to be e-connected to the facts via her intuitions, because we are stipulating that she accepts that her non-basic beliefs are *not* e-connected.

The same reasoning applies however we imagine that the realist initially formed her non-basic beliefs. Her non-basic beliefs have some basis: X. She then accepts that her non-basic moral beliefs are not e-connected. She ought also accept that the basis of her belief, X, is not explained by the non-basic moral facts. As such, she ought to accept that her non-basic beliefs are not e-connected via this basis.

In general, assuming that the realist takes the tf-connection expressed in figure 18 to be the only explanatory connection between her non-basic beliefs and the non-basic facts, the realist cannot take the basis of her belief to imply any *other* explanatory connection between her non-basic beliefs and the non-basic facts. This would imply that there was in fact an alternative explanatory connection between her non-basic moral belief and the non-basic moral fact, and we are stipulating that the realist does not take there to be any such alternative connection.

But if, regardless of the basis of her beliefs, there cannot be any other explanatory connection between her non-basic belief and the non-basic facts than that outlined in figure 18, then, regardless of the basis of her beliefs, justification for her non-basic beliefs will depend on her believing in this explanatory connection. Now that she has accepted the debunker's genealogy of her beliefs, her taking this explanatory connection to hold is the only thing stopping her from accepting that her beliefs have

¹¹⁸ This assumes that, if our belief in P is based on some fact Q, then Q explains our belief. This assumption is most obviously consistent with causal accounts of the basing relation, according to which our belief in P is based on Q only if Q causes our belief in the appropriate way (for example, see Longino, 1978; Audi 1986; Moser 1989 and Ye, 2020a. But even doxastic theories of the basing relation, according to which there are cases in which our belief in P can be based on Q even though Q does not explain our belief, are consistent with the claim that such cases are unusual (for example, see Lehrer (1971)) and that the basis of our belief typically explains our belief.

no explanatory connection with the fact that makes them true. And, as I have argued, her belief in this explanatory connection crucially depends on her belief in the relevant bridging principle. As such, regardless of the basis of her belief, once she accepts that her non-basic belief is only tf-connected via the fact that caring for our offspring promotes survival, justification for her belief becomes dependent on her belief in the relevant bridging principle in virtue of which this tf-connection can hold. And, if we further stipulate that her belief in this bridging principle is based on this tf-connected non-basic belief (as is consistent with the minimalist response under consideration) then justification for her beliefs becomes problematically circular.

In order to see the problem with this strategy more clearly, consider the following example:

Ian's height:

Chris meets Ian and forms the following two beliefs

- (i) Ian is tall, and
- (ii) Ian is selfish.

He forms (i) because he can see Ian is tall. He forms (ii) because Chris has a strong intuition that Ian is selfish. On the basis of this intuition, he forms the belief that Ian is selfish.

Chris then meets various other people who he takes to be both (i) tall and (ii) selfish. Chris forms the belief that each person is tall on the basis of observation, and the belief that they are selfish on the basis of a strong intuition that they are selfish. Once he has encountered a sufficient number of supposedly selfish tall people, Chris forms the belief that

- (iii) being tall predisposes one towards selfishness.

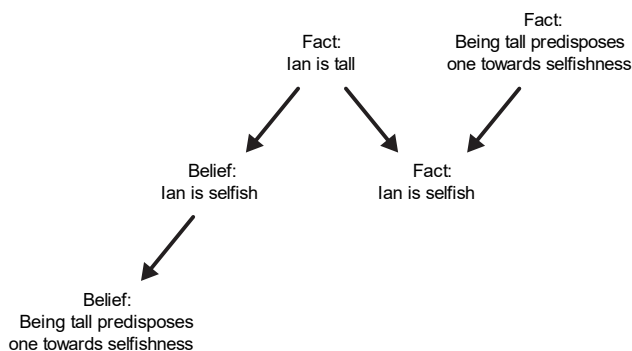
Having noted the persistent correlation between height and selfishness, Chris reasons that being tall must increase the likelihood that one is selfish. This belief in (iii) is inferred on the basis of a number of Chris's beliefs about individuals that take the form "X is tall" in addition to a number of corresponding beliefs about individuals of the form "X is selfish". For any one person, Chris's belief that they are tall and Chris's belief that they are selfish both partially (in addition to corresponding beliefs about a number of other individuals) explain his belief in (iii).

Chris then realizes that all of his initial beliefs that tall people are selfish are not explained by the fact that they are selfish. They are explained by a range of factors, included among which is the fact that they are tall. Thus, the fact that each of these people are tall explains both his belief that they are selfish and then also partially explains his subsequent belief that (iii) being tall predisposes one towards selfishness.

But, Chris reasons, if being tall explains selfishness, then each of his beliefs of the form “x is selfish” is explained by some fact that also explains the truth of that belief: that they are tall. Thus, each of his initial beliefs that X is selfish can be tf-connected and therefore justified, and they seem to imply the general principle that tall people are selfish, and so this inferred belief can also remain justified.

Figure 19 represents the explanatory structure that Chris accepts for a particular set of beliefs about a selfish tall person: Ian. Chris accepts that his belief that (ii) Ian is selfish is explained by the fact that Ian is tall, and he takes Ian’s being tall to explain Ian’s selfishness. He then takes Ian’s selfishness to *also* be explained by the fact that being tall predisposes one towards selfishness. Because his belief in (iii) is explained by his belief in (ii), this implies that his belief in (iii) is a-connected to the fact that makes it true. Given Chris’s pre-existing beliefs in (ii) and (iii), his realisation about the cause of his belief that each tall person is selfish does nothing to count against this a-connection, and in fact supports it.

Figure 19:



I think that all of Chris’s beliefs of the form “X is selfish” become unjustified when he accepts that his belief in Ian’s selfishness is not explained by Ian’s selfishness and is instead explained (among other things) by Ian’s height. I also think that Chris’s attempt to retain justification for his belief, by citing a tf-connection between his belief in Ian’s selfishness and Ian’s selfishness, fails. This this is the case even though Chris did not initially form his belief in Ian’s selfishness by reasoning explicitly from Ian’s height to his selfishness, and his belief was instead formed on the basis of an intuition.

Let's imagine how we would actually feel if someone were to try to retain justification for their beliefs this way. We are speaking with Chris, who believes that being tall predisposes one towards being selfish. We then find out that he has this view because he has met various tall people who he feels, intuitively, are selfish, and on the basis of these observations he has inferred his belief in the general rule that being tall makes one selfish. We then manage to convince him that he only ever believed that any one of these people were selfish, not because they were *actually* selfish, but because they were tall. Chris, we ask, does this not undermine that your initial beliefs that each one of these people are selfish, and in turn undermine the belief you inferred on the basis of these beliefs that that being tall predisposes one towards being selfish?

And now Chris responds as follows: "Aha! My belief that these people are selfish is explained by their being tall, but I already believe that being tall makes people selfish. This implies that I only believe that they are selfish because of some fact that ensures that they are selfish, and my belief can be tf-connected and therefore justified. And, because I am justified in each of these beliefs, I can also remain justified in my inferred belief in the fact that being tall predisposes one towards being selfish. So you have given me no reason to give up any of my beliefs!"

The problem with this line of reasoning should be clear; having realised that his belief in Ian's selfishness is not explained by Ian's selfishness, Chris hopes to retain justification for this belief by arguing that it is tf-connected to the fact that makes it true. However, the claim that this belief is tf-connected presupposes the fact that being tall predisposes one towards selfishness. As such, his belief that Ian is selfish can only be justified under the assumption that being tall predisposes one towards selfishness. But Chris only believes that being tall predisposes one towards selfishness on the basis of this very belief in Ian's selfishness. As such, Chris cannot rely on the claim that being tall predisposes one towards selfishness in the process of justifying his belief that Ian is selfish.

This seems to be a problem with Chris's proposed response *even though* Chris initially formed his belief in Ian's selfishness on the basis of an intuition, rather than on the basis of a conscious line of inference from the fact that Ian is tall to the fact that Ian is selfish. Once he accepts that his belief that Ian is selfish is just tf-connected via the fact that Ian is tall, then any bases for his belief cannot offer an alternative explanatory

connection between his belief in P and P, and so taking that belief to be explanatorily connected to the relevant fact (and therefore justified) will depend on this tf-connection regardless of the initial basis of that belief.

I think that the case of *Ian's height* bears important similarities to the realist response under consideration. The realist has some unspecified basis for her belief that various actions that promote survival are good (just as Ian had a strong intuition that various people who are tall are selfish). From these beliefs, the realist infers the more general belief that survival is good (just as Ian inferred the more general belief being tall predisposes one towards selfishness). Finally, the realist is told (and accepts) that none of her beliefs are e-connected, which implies that her non-basic beliefs in specific actions being morally good are not actually explained by the fact that the actions in question are morally good. This, in turn, implies that whatever basis she initially had for her non-basic beliefs are not explained by the non-basic facts.

The suggestion under consideration is that the realist maintain justification for their non-basic moral beliefs as follows: “well, my beliefs that these actions are good are explained by the fact that they promote survival, but, because survival is good, I only have these beliefs because of some fact that ensures its truth, and those beliefs are therefore tf-connected. As such, I can continue to justifiably believe that each of these actions are good, and they can therefore continue to form a basis for my basic belief that survival is good.”

This reasoning is flawed in the same way that the corresponding proposed response from Chris is flawed. The problem arises because the realist takes her beliefs to be only be justified in virtue of being tf-connected in such a way that presupposes the relevant bridging principle, and as long as this bridging principle is inferred on the basis of these very non-basic beliefs, then her reasoning will be problematically circular. She will therefore succeed justifying neither her non-basic nor her basic moral beliefs. But these features that generate the problem are just those features that characterise the realist response under consideration, according to which tf-connected (and therefore justified) non-basic moral beliefs can form the basis of basic moral beliefs.

I therefore conclude that the minimalist response under consideration fails. Because this response fails, it does not present a counter-example to (IEC). We are not

presented with an instance of a belief that one takes to be a-connected (and not e-connected or tf-connected) but which can remain justified. As such, this minimalist response gives us no reason to doubt either my expanded evolutionary debunking argument against moral realism nor the legitimacy of (IEC) as a constraint on belief.

6.5: An Independence Clause:

My critique of this minimalist response implies that, if we accept some explanation of our belief in P and we then have no independent reason for thinking that our belief in P might be explanatorily related to P, our belief in P is unjustified. This was the case with our non-basic beliefs in the response under consideration; the minimalist accepted that her non-basic beliefs can be given an evolutionary explanation (according to which those beliefs are not e-connected), and her only reason for thinking that her non-basic beliefs were tf-connected were basic moral beliefs that were inferred from those very non-basic beliefs. This rendered justification for those non-basic beliefs unacceptable circular, and so these beliefs were not justified.

This implies that we could add some sort of independence clause to (IEC), of the sort that Lutz includes for his own preferred explanatory constraint (2018). According to Lutz:

“Explaining Away Defeats (EAD): New evidence, D, defeats the support that E provides for S’s belief that P if: D is evidence in favor of a complete explanation, A, of E, such that S may not infer P from A and S’s independent background information.” (2018: 1110)

According to (EAD), then, new evidence (D) defeats the support that evidence (E) provides for our belief in P if (D) supports a complete explanation of (E) and we cannot infer the truth of our belief from that explanation in combination with our *independent* background information. Background information is independent when it is distinct from our evidence (E) and our belief in P and is not justified on the basis on either our evidence (E) or our belief in P. Once we have reason to accept a particular explanation of our evidence (E), if we can only infer P from that explanation of our evidence by relying on dependent background information, then our belief in P is defeated.

I think that a lesson from this chapter is that a similar clause should be added to (IEC), modelled on Lutz's independence clause of (EAD). When (IEC) has been so amended, we would get something like the following explanatory constraint on belief:

(IEC*): If (i) S withholds belief that her belief that P is explained by the fact that P and (ii) S withholds belief that there is some fact that explains both her belief that P and the fact that P,

or if S accepts an explanation of her belief in P and S cannot infer that her belief in P is e-connected or tf-connected from this explanation and her independent background information,

then S's belief that P is not justified.

According to (IEC*), our belief is defeated if we accept some explanation of that belief and we cannot infer that our belief in P is e-connected or tf-connected from that explanation in combination with our independent background information. If P is featured in our accepted explanation of P, then this explanation of our belief entails that our belief is e-connected because it entails that our belief in P is explained by P, and we can therefore infer that our belief is e-connected from this explanation of our belief. If one of our background beliefs includes the belief that some fact in our accepted explanation of our belief is explained by or explains P, then we can infer that our belief in P is e-connected or tf-connected from this explanation of our belief in combination with this background belief. Importantly, though, this background belief has to be *independent* of our belief in P in order for this possibility to allow our beliefs to remain justified. (IEC*) therefore specifies a further way in which our belief in P might be defeated. It is still true, as claimed by (IEC), that our belief in P is defeated if we withhold belief that P is e-connected or tf-connected, but it is also the case that our belief in P is defeated if we accept some explanation of our belief such that we can only justifiably take our belief in P to be e-connected or tf-connected by relying on beliefs that are dependent on P.¹¹⁹

¹¹⁹ This clause does *not* imply that, for any belief to be justified, we must have some independent reason for believing that our belief in P is explanatorily related to P (Lutz, 2018: 1114). The clause is consistent with our having no independent reason for thinking our belief in P is explanatorily related to P prior to accepting any explanation of that belief. But once we accept some explanation of our belief, we must have some reason for thinking that our belief in P is explanatorily related to P that is not dependent on P.

Applied to the realist response under consideration, the realist accepts that their moral beliefs can be given an evolutionary explanation and that this explanation implies that realistically construed moral facts do not explain moral beliefs. Furthermore, they only have reason to think that their moral beliefs are tf-connected, given this explanation, if they assume some bridging principle that implies a tf-connection between their moral beliefs and one of the natural facts that feature in this explanation. The realist response under consideration involved the claim that their belief in this bridging principle can be inferred from these non-basic beliefs, and this implies that their belief in the bridging principle is dependent on their non-basic beliefs. When they can only rely on *independent* beliefs, then, they have no way of inferring from their accepted explanation of their beliefs that those beliefs are e-connected or tf-connected. As such, the minimalist's moral beliefs fail (IEC*) and are therefore unjustified.

Lutz's own debunking argument involves the claim that *only* our beliefs in natural facts are independent relative to evolutionary debunking arguments against non-naturalism. Because we cannot infer moral beliefs from the evolutionary explanation of those beliefs in combination with our beliefs in natural facts, our moral beliefs (given non-naturalism) are defeated according to (EAD) (2018: 1119-1120).

Now that I have suggested my own independence clause for (IEC*), we might wonder whether I can rule out minimalist replies to debunking arguments using similar reasoning. That is, I could first argue that only natural facts are independent relative to the debunking challenge. Next, I could point out that the minimalist accepts that the complete explanation of their beliefs will not involve the fact that P. Given my suggested independence clause, then, in order for the minimalist's beliefs to be justified they must have some independent reason for thinking that those beliefs are tf-connected. And, as our moral beliefs can only be tf-connected given the truth of *moral* bridging principles, when relying on only independent beliefs in natural facts the minimalist cannot infer that their moral beliefs are tf-connected from their accepted explanation of their moral beliefs. Minimalist replies therefore fail because they cannot infer that their moral beliefs are e-connected or tf-connected from their accepted explanation of those beliefs plus their independent background beliefs.

I think this line of reasoning only rules out minimalist replies to debunking arguments under the assumption that the presupposed bridging principle is inferred on the basis

of those non-basic beliefs that are then taken to be tf-connected (as with the realist response that I have been considering in this chapter). In that case, the realist has no *independent* reason for thinking that their beliefs are tf-connected, because the bridging principle that would imply their non-basic beliefs are tf-connected depend on those very beliefs.

But it seems open to the third-factor replier to insist that their beliefs in the relevant bridging principle are *not* based on these beliefs. Above I argued that there is no reason to think that our non-basic moral beliefs must be inferred from our basic moral beliefs, but it also seems that there is no reason to think that our basic beliefs *must* be inferred from our non-basic beliefs. And, if this is true, then it is possible for the realist to insist that their beliefs in basic moral facts (that function as the requisite bridging principles) are not based on, and are therefore independent of, their non-basic moral beliefs. Once this possibility is allowed, the minimalist's belief in the relevant bridging principle is among those independent beliefs from which she can infer that her non-basic beliefs are tf-connected given her accepted explanation of those beliefs. And once the realist is allowed to rely on this belief, it seems she can infer that her moral beliefs are tf-connected given the explanation of her moral beliefs that she accepts, and her beliefs no longer fail the independence clause of (IEC*).

I therefore think that the independence clause of (IEC*) only implies that minimalist replies fail if we assume, as we have done in this chapter, that our belief in the relevant bridging principle is inferred on the basis beliefs that we take to be tf-connected in virtue of this principle. Without this assumption, there is no reason to think that the realist cannot infer that her moral beliefs are tf-connected given the truth of the debunker's evolutionary explanation in combination with her independent background beliefs. In order to see why the minimalist replies fail even if we do not assume that the realist's basic beliefs are dependent on their non-basic beliefs in this way, we need the arguments of the previous chapter. The independence clause of (IEC*) therefore does not render those arguments superfluous.

6.6: An Objection:

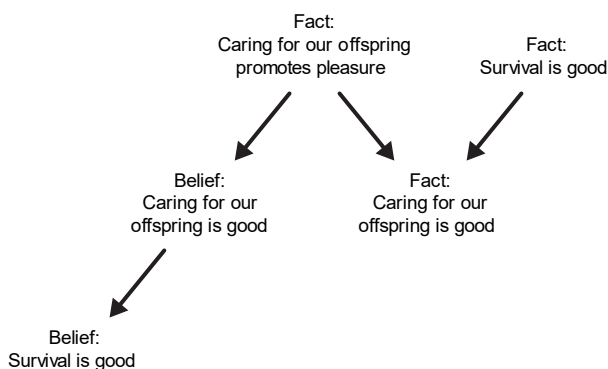
I think that the most pressing objection to my argument in this chapter is as follows: I have assumed that, according to the minimalist response under consideration, the basic

moral belief that gets justified on the basis of our tf-connected, non-basic moral beliefs is in the very bridging principle that is presupposed by our taking our non-basic beliefs to be tf-connected. Because the inferred basic belief is the same as the relevant bridging principle, the realist cannot take her belief to be tf-connected in this way without presupposing a belief that is inferred on the basis of these non-basic beliefs, and this reply fails.

But is it not possible that we could take our non-basic beliefs to be tf-connected in some way other than that which presupposes the inferred basic moral belief? Why must the inferred basic belief and the required bridging principle be the same? And, if they can be different, then we can take our belief to be tf-connected in such a way that does not presuppose the basic belief that we infer from those non-basic beliefs, and our justification for both our non-basic and our basic beliefs is not problematically circular in the way that I have been suggesting.

This is an interesting suggestion, Figure 20:

but I think that it has a number of problems. In order to see this, imagine a concrete case in which the inferred basic belief is distinct from the bridging principle that is required for the non-basic belief to be tf-connected. Imagine that we



take our belief that caring for our offspring is good to be explained by the natural fact that caring for our children promotes pleasure. In order for our non-basic belief to be tf-connected in this way we do not need to assume that survival is good. Instead, this tf-connection presupposes an alternative bridging principle: that promoting pleasure is good. We can therefore take our non-basic beliefs to be tf-connected in this way without presupposing that survival is good. Our belief in the basic moral fact that survival is good can then be based on our belief that caring for our offspring is morally good, and justification for this basic belief is not problematically circular. Figure 20 represents the explanatory structure that we accept for our beliefs according to this suggestion.

The first problem with this suggestion is that it is unclear whether, once we take our non-basic beliefs in the goodness of survival to be tf-connected in this way, we can continue to justifiably base our belief in the goodness of survival on this non-basic belief. We take the fact that caring for our offspring is good to be explained by the fact that it promotes pleasure. According to structure of normative reality outlined in the previous chapter, this implies that the basic moral fact that explains this belief is the fact that promoting pleasure is good (hence why we need to presuppose this fact in order to believe that our belief is tf-connected *via* the fact that looking after our children promotes pleasure). We therefore ought to take the goodness of looking after our offspring to be explained by the fact that looking after our offspring promotes pleasure, in addition to the fact that promoting pleasure is good. But if we take these facts to explain the fact that looking after our offspring is morally good, why should we take this fact to be a good indication of the fact that *survival* is morally good? It appears we can fully account for this non-basic moral fact without needing to posit the goodness of survival, and we therefore cannot justifiably infer from our belief that looking after our offspring is good that survival is good.

It therefore seems there is a connection between the natural fact that we take to be responsible for a given non-basic moral fact and the more basic belief that we can justifiably infer from our belief in that non-basic moral fact. If we take our non-basic moral belief that looking after our offspring is good to be tf-connected *via* the fact that promoting pleasure is good, then it seems that we can only justifiably infer from this belief that promoting pleasure is morally good. This implies that the bridging principle that we have to presuppose in order to take our non-basic belief to be tf-connected must be identical to the basic belief that we infer from those non-basic moral beliefs in order for that inference to be justified.

The second problem with this proposal that the presupposed bridging principle might be distinct from the inferred basic moral fact is that it leaves us with no account of how the presupposed bridging principle that promoting pleasure is good gets justified. We cannot withhold belief in this fact if we take our belief that caring for our offspring is good to be tf-connected in the way outlined in figure 20. But what justification do we have for believing in this bridging principle?

This is not a problem when the presupposed bridging principle is identical with the inferred basic moral belief, as I was assuming previously. In this case we have a ready account of how the presupposed bridging principle gets justified: it is justified on the basis of those non-basic beliefs that we take to be tf-connected¹²⁰. But when the inferred basic moral belief is distinct from the presupposed bridging principle, the presupposed bridging principle cannot be justified on the basis of those non-basic beliefs because we are stipulating that it is not inferred on the basis of those beliefs. But then we have a question: how *can* it be justified?

It cannot be justified on the basis of our non-basic belief that looking after our offspring is good, because then the presupposed bridging principle is identical to the inferred basic moral belief and the reasoning for this belief becomes problematically circular in the way that I have been arguing. It cannot be justified on the basis of other non-basic beliefs that we take to be tf-connected, because either those beliefs being tf-connected presupposes that very belief and the justification for that belief is circular, or the non-basic belief presupposes an alternative bridging principle, and the problem starts over again. As a result, then, this proposal presupposes a bridging principle and it is unclear how belief in this crucial bridging principle gets justified.

I therefore think that the minimalist response under consideration cannot succeed if the non-basic beliefs are tf-connected in a way that presupposes a bridging principle that is distinct from the basic moral fact that is inferred on the basis of those beliefs. Unfortunately for the minimalist, I have already argued that this response cannot succeed if the presupposed bridging principle and the inferred basic moral fact are the same. The upshot is that the minimalist response under consideration fails, and it cannot establish that our realistically construed basic moral beliefs can be justified.

6.7: Conclusion:

In this chapter I have considered and argued against a potential way in which the minimalist might hope to retain justification for their basic moral beliefs even once they have accepted that those beliefs can be neither e-connected nor tf-connected and therefore fail (IEC). The suggestion was that they could be justified by being inferred

¹²⁰ Though, for the reasons I have been arguing, I think that this account of how the relevant bridging principle is unsatisfactory.

from our non-basic beliefs which we *can* take to be tf-connected, and which might therefore form a good basis for those basic moral beliefs.

The fundamental problem with this suggestion, I have argued, is that we can only take our non-basic beliefs to be tf-connected by presupposing the truth of the very basic belief that we hope to infer from those non-basic beliefs. This renders the strategy unacceptably circular, and it therefore fails to establish that our realistically construed basic moral beliefs can remain justified in spite of the debunker's genealogy of those beliefs. As such, this possibility does not pose a threat to either my expanded evolutionary debunking argument, or (IEC) as a legitimate constraint on belief. And, with this objection dealt with, I conclude my defence of this expanded evolutionary debunking argument, and thereby conclude my assessment of why minimalist responses to evolutionary debunking arguments fail.

Conclusion:

If the arguments of this thesis are successful, then our realistically construed moral beliefs are defeated under the assumption that evolutionary debunking arguments successfully establish they are not e-connected. This implies that minimalist responses to these arguments fail because, having accepted that evolutionary debunking arguments establish this point, their moral beliefs are unjustified, and they cannot rely on any substantive moral claims. As all such responses *do* crucially rely on substantive moral claims, those responses all fail.

The argument for this position was a long one. Having outlined evolutionary debunking arguments and minimalist responses in chapters one and two, I first had to argue against those commentators who felt that explanatory considerations cannot defeat our beliefs if they do not also threaten the modal security of those beliefs. I did this by considering examples in which an agent makes an explanatory concession without having reason to doubt the modal security of that belief. Because the belief in question is undermined, this implies that explanatory concessions can undermine independently of what they imply about the modal security of that belief. The modal security response, as outlined in chapter two, is therefore fundamentally misguided.

Next, in chapter four, I defended a specific explanatory constraint on belief: (IEC). I formulated (IEC) by considering the flaws of other explanatory constraints and by showing that (IEC) can avoid them. If the arguments in this chapter are correct, then they imply that our belief is defeated if we withhold belief that it is e-connected *and* we withhold belief that it is tf-connected, and this contrasts with a number of other constraints that have been suggested in the literature

In the final two chapters I have formulated and defended an expanded evolutionary debunking arguments which implies that our basic moral beliefs fail (IEC) under the assumption of moral realism. If (IEC) is valid, then, according to the structure of normative reality accepted by non-naturalists, our beliefs in basic moral facts cannot be tf-connected. As minimalists accept that these beliefs are not e-connected, this

implies that we should withhold all such beliefs given (IEC). And there are good reasons, I argued, for thinking that if our beliefs in moral facts that we take to be basic are defeated then so are our beliefs in facts we take to be non-basic.

In chapter six, I considered a potential response from the minimalist according to which justification for our basic moral beliefs can be secured if they are inferred from our non-basic moral beliefs. I argued that this response fails because it implies that justification for our non-basic moral beliefs and our basic beliefs becomes problematically circular. These arguments imply that we can amend (IEC) to include a further condition, according to which our beliefs are defeated if we accept some explanation of those beliefs and we cannot infer that our beliefs are e-connected or tf-connected from that explanation in combination with our independent background information.

If the arguments of this thesis are successful, they imply that minimalist responses to debunking arguments all fail. These responses are characterised by their acceptance that none of their moral beliefs are explained by the facts that make them true and, having made this concession, they then rely on substantive moral claims without argument. But I have argued that, once they accept that their moral beliefs are not e-connected then this ultimately implies that *all* our moral beliefs have been defeated, and any *prima facie* justification the minimalist had for assuming that their moral beliefs are true has been defeated.

One way of stopping this process is to resist the debunker's claims that our moral beliefs are not explained by the facts that make them true. I see two main ways in which they might do this. Firstly, they might object to the claim that our moral beliefs can be given an evolutionary explanation at all. Perhaps the evidence is, at this point, too thin to support this claim. Secondly, the realist might argue that our moral beliefs *can* be given an evolutionary explanation, but not one that implies our moral beliefs are not explained by the moral facts. One way of doing this is to argue that evolution selected for true moral beliefs, but another (in my view more plausible) strategy is to argue that the evolutionary explanation is silent on the proximate causes of our moral

beliefs, and it is therefore consistent with the possibility that we are in touch with moral facts even if evolution did not select for true moral beliefs¹²¹.

This is where I see the most potential for further research on the debunking side of things. There is always room for further empirical research into the evolutionary origins of our moral beliefs, and how we should interpret this evidence when it comes to the proximate causes of our moral beliefs requires more attention. Without fully addressing these issues, evolutionary debunking arguments (both in general and when expanded in the way I have suggested) are only conditional; *if* our moral beliefs can be given an evolutionary explanation, and *if* this explanation implies that our moral beliefs are not explained by the moral facts, then our moral beliefs cannot be justified under the assumption of moral realism.

We also cannot ignore the possibility that the realist might accept that their basic moral beliefs are not explained by the moral facts, but argue against my claim that their basic moral beliefs cannot be tf-connected. As this point is straightforwardly entailed by the supposition that basic moral facts are unexplained, the realist pursuing this strategy would need to deny the existence of unexplained moral facts. As I have tried to indicate, I find it hard to see how the realist could do this consistently, but it is worth noting that this is a potential avenue for the realist who hopes to accept the debunker's claim that our moral beliefs are not explained by the moral facts whilst retaining justification for their basic moral beliefs. The realist, of course, might also deny (IEC), and thereby deny that the impossibility of our basic moral beliefs being e-connected or tf-connected implies that those beliefs are unjustified.

On the epistemological side of things, I think more needs to be said about different kinds of explanatory relationships that our belief might have towards the facts that make it true. I have shown why some examples of beliefs one takes to be a-connected cannot be justified, but I think that *any* belief we take to be a-connected is bound to face similar problems. The structure of an a-connection implies we can never take our belief to have this relationship with the fact that makes it true in such a way that does not depend on our belief in P itself. This will render our justification for that belief unacceptably circular and it will fail the independence clause of (IEC). Establishing

¹²¹ For more on this suggestion, see Andreas Mogensen (2016) and Louise Hanson (2017)

this point can shed further light on why e-connections and tf-connections play such an important role in epistemic justification.

Relatedly, in this thesis, I have been concerned with cases in which explanatory transitivity holds. In these cases, if P explains some fact Q and Q explains our belief, then P explains our belief and our belief is e-connected. But interesting questions arise when we consider cases in which transitivity breaks down. If we take our belief in P to be connected to P by a chain of e-connections, but not e-connected because transitivity breaks down somewhere along the way (implying that P does not explain our belief in P), then can our belief still be justified? If not, why not? I think pursuing this question can further our understanding of why certain attitudes towards the explanatory history of our belief are inconsistent with that belief being justified.

In this thesis I have defended an epistemological principle that stipulates the kind of attitude we can have towards the explanatory history of our belief such that our belief is unjustified. I have *not* attempted to provide a complete analysis of when a belief of ours gets defeated (IEC presents a sufficient, not a necessary, condition for defeat), and I have certainly not attempted to provide an analysis of epistemic justification in general.

But some of the arguments in this thesis clearly relate to an increasingly popular explanationist analysis of justification (Ted Poston, 2014, McCain, 2014, Lutz, 2020a & 2020b)¹²². Investigating the plausibility of this stronger form of explanationism, and investigating its relationship with my far weaker claim that explanatory considerations are sufficient for defeat, is another area that I think requires further attention.

¹²² See also Bogardus and Perrin (*forthcoming*) who argue for an explanationist account of knowledge.

Abbreviations:

EDA: Evolutionary debunking argument.

IEC: Inclusive Explanatory Constraint.

Bibliography:

- Audi, Robert. 'Belief, Reason, and Inference'. *Philosophical Topics* 14, no. 1 (1986): 27–65.
- Bedke, Matthew S. 'Intuitive Non-Naturalism Meets Cosmic Coincidence'. *Pacific Philosophical Quarterly* 90, no. 2 (2009): 188–209.
- . 'No Coincidence?' *Oxford Studies in Metaethics* 9 (2014): 102–25.
- Benacerraf, Paul. 'Mathematical Truth'. *Journal of Philosophy* 70, no. 19 (1973): 661–79.
- Bergmann, Michael. 'Defeaters and Higher-Level Requirements'. *Philosophical Quarterly* 55, no. 220 (2005): 419–36.
- Berker, Selim. 'Does Evolutionary Psychology Show That Normativity Is Mind-Dependent?' In *Moral Psychology and Human Agency: Philosophical Essays on the Science of Ethics*, edited by Justin D'Arms and Daniel Jacobson, 215–52. Oxford University Press, 2014.
- Blackburn, Simon. *Essays in Quasi-Realism*. Vol. 1. Oxford University Press, 1993.
- Bogardus, Tomas. 'Only All Naturalists Should Worry About Only One Evolutionary Debunking Argument'. *Ethics* 126, no. 3 (2016): 636–61.
- Bogardus, Tomas, and Will Perrin. 'Knowledge Is Believing Something Because It's True'. *Episteme*, forthcoming, 1–19.
- Brosnan, Kevin. 'Do the Evolutionary Origins of Our Moral Beliefs Undermine Moral Knowledge?' *Biology and Philosophy* 26, no. 1 (2011): 51–64.
- Byerly, T. Ryan. 'Explanationism and Justified Beliefs about the Future'. *Erkenntnis* 78, no. 1 (2013): 229–43.
- Byerly, T. Ryan, and Kraig Martin. 'Problems for Explanationism on Both Sides'. *Erkenntnis* 80, no. 4 (2015): 773–91.
- Campbell, Richmond. 'What Is Moral Judgment?' *Journal of Philosophy* 104, no. 7 (2007): 321–49.
- Christensen, David. 'Higher Order Evidence'. *Philosophy and Phenomenological Research* 81, no. 1 (2010): 185–215.
- Clarke-Doane, Justin. 'Morality and Mathematics: The Evolutionary Challenge'. *Ethics* 122, no. 2 (2012): 313–40.
- . 'Justification and Explanation in Mathematics and Morality'. *Oxford Studies in Metaethics* 10 (2015).

- Clarke-Doane, Justin, and Dan Baras. 'Modal Security'. *Philosophy and Phenomenological Research* 102, no. 1 (2021): 162–83.
- Climenhaga, Nevin. 'Evidence and Inductive Inference'. In *The Routledge Handbook of the Philosophy of Evidence*, edited by Maria Lasonen-Aarnio and Clayton Littlejohn. Routledge, 2021.
- Cohen, Stewart. 'Justification and Truth'. *Philosophical Studies* 46, no. 3 (1984): 279–95.
- Copp, David. 'Darwinian Skepticism about Moral Realism'. *Philosophical Issues* 18, no. 1 (2008): 186–206.
- . 'How to Avoid Begging the Question against Evolutionary Debunking Arguments'. *Ratio* 32, no. 4 (2019): 231–45.
- Crisp, Roger. *Reasons and the Good*. Vol. 57. Clarendon Press, 2006.
- Cuneo, Terence, and Russ Shafer-Landau. 'The Moral Fixed Points: New Directions for Moral Nonnaturalism'. *Philosophical Studies* 171, no. 3 (2014): 399–443.
- DiPaolo, Joshua. 'Higher-Order Defeat Is Object-Independent'. *Pacific Philosophical Quarterly* 99, no. 2 (2018): 248–69.
- Dreier, James. 'Meta-ethics and the Problem of Creeping Minimalism'. *Philosophical Perspectives* 18, no. 1 (2004): 23–44.
- Dworkin, Ronald. 'Objectivity and Truth: You'd Better Believe It'. *Philosophy and Public Affairs* 25, no. 2 (1996): 87–139.
- . *Justice for Hedgehogs*. Belknap Press of Harvard University Press, 2011.
- Dyke, Michelle M. 'Bad Bootstrapping: The Problem with Third-Factor Replies to the Darwinian Dilemma for Moral Realism'. *Philosophical Studies* 177, no. 8 (2020): 2115–28.
- Enoch, David. 'The Epistemological Challenge to Metanormative Realism: How Best to Understand It, and How to Cope with It'. *Philosophical Studies* 148, no. 3 (2010): 413–38.
- . *Taking Morality Seriously: A Defense of Robust Realism*. Oxford University Press UK, 2011.
- Faraci, David. 'Groundwork for an Explanationist Account of Epistemic Coincidence'. *Philosophers' Imprint* 19 (2019).
- Fine, Kit. 'Varieties of Necessity'. In *Conceivability and Possibility*, edited by Tamar Szabo Gendler and John Hawthorne, 253–81. Oxford Up, 2002.
- FitzPatrick, William Joseph. 'Robust Ethical Realism, Non-Naturalism, and Normativity'. *Oxford Studies in Metaethics* 3 (2008): 159–205.

- Foot, Philippa. 'Moral Arguments'. *Mind* 67, no. 268 (1958): 502–13.
- Fraser, Ben. 'Evolutionary Debunking Arguments and the Reliability of Moral Cognition'. *Philosophical Studies* 168, no. 2 (2014): 457–73.
- Gibbard, Allan. *Thinking How to Live*. Vol. 116. Harvard University Press, 2003.
- Goldman, Alvin I. 'A Causal Theory of Knowing'. *Journal of Philosophy* 64, no. 12 (1967): 357–72.
- . 'What Is Justified Belief?' In *Justification and Knowledge*, edited by George Pappas, 1–25. D. Reidel, 1979.
- Greene, Joshua. 'The Secret Joke of Kant's Soul'. In *Moral Psychology, Vol. 3*, edited by W. Sinnott-Armstrong. MIT Press, 2007.
- Guigon, Ghislain. 'A Universe of Explanations'. In *Oxford Studies in Metaphysics*, edited by Karen Bennett and Dean W. Zimmerman, 345–75. Oxford University Press, 2015.
- Hanson, Louise. 'The Real Problem with Evolutionary Debunking Arguments'. *Philosophical Quarterly* 67, no. 268 (2017): 508–33.
- Hattiangadi, Anandi. 'Moral Supervenience'. *Canadian Journal of Philosophy* 48, no. 3–4 (2018): 592–615.
- Hausman, Daniel M. 'Why Don't Effects Explain Their Causes?' *Synthese* 94, no. 2 (1993): 227–44.
- Huemer, Michael. *Ethical Intuitionism*. Vol. 78. Palgrave Macmillan, 2005.
- Jenkins, Carrie. 'Is Metaphysical Dependence Irreflexive?' *The Monist* 94, no. 2 (2011): 267–76.
- Joyce, Richard. *The Myth of Morality*. Cambridge University Press, 2001.
- . *The Evolution of Morality*. Bradford, 2005.
- . *Essays in Moral Skepticism*. Oxford University Press UK, 2016.
- Kahane, Guy. 'Evolutionary Debunking Arguments'. *Noûs* 45, no. 1 (2011): 103–25.
- Kim, Jaegwon. 'Explanatory Realism, Causal Realism, and Explanatory Exclusion'. *Midwest Studies in Philosophy* 12, no. 1 (1988): 225–39.
- Kitcher, Philip. 'Psychological Altruism, Evolutionary Origins, and Moral Rules'. *Philosophical Studies* 89, no. 2–3 (1998): 283–316.
- . 'Biology and Ethics'. In *The Oxford Handbook of Ethical Theory*, edited by David Copp. Oxford University Press, 2005.
- Klenk, Michael. 'Evolutionary Ethics'. In *Introduction to Philosophy: Ethics*, edited by George W. Matthews, 76–90, 2019a.

- . ‘Objectivist Conditions for Defeat and Evolutionary Debunking Arguments’. *Ratio* 32, no. 4 (2019b): 246–59.
- . ‘Third Factor Explanations and Disagreement in Metaethics’. *Synthese* 197, no. 1 (2020): 427–46.
- Korman, Daniel Z. ‘Debunking Arguments’. *Philosophy Compass* 14, no. 12 (2019a).
- . ‘Debunking Arguments in Metaethics and Metaphysics’. In *Metaphysics and Cognitive Science*, edited by Alvin Goldman and Brian McLaughlin, 337–63. Oxford University Press, 2019b.
- Korman, Daniel Z., and Dustin Locke. ‘Against Minimalist Responses to Moral Debunking Arguments’. *Oxford Studies in Metaethics* 15 (2020): 309–32.
- Lange, Marc. *Because Without Cause: Non-Causal Explanations in Science and Mathematics*. Oxford University Press USA, 2016.
- Lehrer, Keith. ‘How Reasons Give Us Knowledge, or the Case of the Gypsy Lawyer’. *Journal of Philosophy* 68, no. 10 (1971): 311–13.
- . ‘Ultimate Preference and Explanation’. *Grazer Philosophische Studien* 97, no. 4 (2020): 600–615.
- Lehrer, Keith, and Stewart Cohen. ‘Justification, Truth, and Coherence’. *Synthese* 55, no. 2 (1983): 191–207.
- Levy, Arnon, and Yair Levy. ‘Evolutionary Debunking Arguments Meet Evolutionary Science’. *Philosophy and Phenomenological Research* 100, no. 3 (2020): 491–509.
- Lewis, David. ‘Causal Explanation’. In *Philosophical Papers Vol. Ii*, edited by David Lewis, 2:214–40. Oxford University Press, 1986.
- Locke, Dustin. ‘Darwinian Normative Skepticism’. In *Challenges to Moral and Religious Belief: Disagreement and Evolution*, edited by Michael Bergmann and Patrick Kain. Oxford University Press, 2014.
- Longino, Helen E. ‘Inferring’. *Philosophy Research Archives* 4 (1978): 17–26.
- Lutz, Matt. ‘What Makes Evolution a Defeater?’ *Erkenntnis* 83, no. 6 (2018): 1105–26.
- . ‘Background Beliefs and Plausibility Thresholds: Defending Explanationist Evidentialism’. *Synthese* 197, no. 6 (2020a): 2631–47.
- . ‘Explanationism Provides the Best Explanation of the Epistemic Significance of Peer Disagreement’. *Philosophical Studies* 177, no. 7 (2020b): 1811–28.
- McCain, Kevin. *Evidentialism and Epistemic Justification*. Routledge, 2014.
- . ‘Explanationism: Defended on All Sides’. *Logos and Episteme* 6, no. 3 (2015): 333–49.

- . ‘Undaunted Explanationism’. *Logos and Episteme* 8, no. 1 (2017): 117–27.
- McGrath, Sarah. *Moral Knowledge*. Oxford University Press, 2019.
- Mogensen, Andreas L. ‘Do Evolutionary Debunking Arguments Rest on a Mistake about Evolutionary Explanations?’ *Philosophical Studies* 173, no. 7 (2016): 1799–1817.
- Morton, Justin. ‘A New Evolutionary Debunking Argument Against Moral Realism’. *Journal of the American Philosophical Association* 2, no. 2 (2016): 233–53.
- Moser, Paul K. *Knowledge and Evidence*. Cambridge University Press, 1989.
- Nagel, Thomas. *The View From Nowhere*. Vol. 37. Oxford University Press, 1986.
- Olson, Jonas. ‘What Can Debunking Do for Us (Sceptics and Nihilists)?’ *Ratio* 32, no. 4 (2019): 290–99.
- Parfit, Derek. *On What Matters: Two-Volume Set*. Oxford University Press, 2011.
- Paul, Audi. ‘Explanation and Explication’. In *Palgrave Handbook of Philosophical Methods*, edited by Chris Daly. Palgrave Macmillan, 2015.
- Paul, L. A., and Ned Hall. *Causation: A User’s Guide*. Oxford University Press UK, 2013.
- Pollock, John L. *Knowledge and Justification*. Vol. 26. Princeton University Press, 1974.
- . ‘Reliability and Justified Belief’. *Canadian Journal of Philosophy* 14, no. 1 (1984): 103–14.
- . ‘Defeasible Reasoning’. *Cognitive Science* 11, no. 4 (1987): 481–518.
- Pollock, John, and Joe Cruz. *Contemporary Theories of Knowledge, 2nd Edition*. Rowman & Littlefield, 1999.
- Rosen, Gideon. Metaphysical Dependence: Grounding and Reduction. In *Modality: Metaphysics, Logic, and Epistemology*, edited by Bob Hale and Aviv Hoffmann, 109–135. Oxford University Press, 2010.
- . ‘Metaphysical Relations in Metaethics’. In *The Routledge Handbook of Metaethics*, edited by Tristram McPherson and David Plunkett, 151–69. Routledge, 2017.
- . ‘What is Normative Necessity?’. In *Metaphysics, Meaning, and Modality: Themes from Kit Fine*, edited by Mircea Dumitru, 205–233. Oxford University Press, 2020.
- Ruben, David-Hillel. *Explaining Explanation*. Routledge, 1990.
- Ruse, Michael. *Taking Darwin Seriously: A Naturalistic Approach to Philosophy*. Vol. 3. Prometheus Books, 1986.
- Schafer, Karl. ‘Evolution and Normative Scepticism’. *Australasian Journal of Philosophy* 88, no. 3 (2010): 471–88.

- Schaffer, Jonathan. 'Grounding in the Image of Causation'. *Philosophical Studies* 173, no. 1 (2016): 49–100.
- Shafer-Landau, Russ. 'Evolutionary Debunking, Moral Realism and Moral Knowledge'. *Journal of Ethics and Social Philosophy* 7, no. 1 (2012): 1–38.
- . *Moral Realism: A Defence*. Oxford University Press, 2003.
- Singer, Peter. 'Sidgwick and Reflective Equilibrium'. *The Monist* 58, no. 3 (1974): 490–517.
- . 'Ethics and Intuitions'. *The Journal of Ethics* 9, no. 3–4 (2005): 331–52.
- Sinnott-Armstrong, Walter. *Moral Skepticisms*. Oxford University Press, 2006.
- Skarsaune, Knut Olav. 'Darwin and Moral Realism: Survival of the Iffiest'. *Philosophical Studies* 152, no. 2 (2011): 229–43.
- Skow, Bradford. 'Are There Non-Causal Explanations (of Particular Events)?' *British Journal for the Philosophy of Science* 65, no. 3 (2014): 445–67.
- Sosa, Ernest. 'Reliability and the a Priori'. In *Conceivability and Possibility*, edited by John Hawthorne and Tamar Gendler, 369–384. Oxford University Press, 2002.
- Strandberg, Caj. 'Particularism and Supervenience'. In *Oxford Studies in Metaethics*, edited by Russ Shafer-Landau. Oxford University Press, 2008.
- Street, Sharon. 'A Darwinian Dilemma for Realist Theories of Value'. *Philosophical Studies* 127, no. 1 (2006): 109–66.
- . 'Constructivism about Reasons'. *Oxford Studies in Metaethics* 3 (2008): 207–45.
- . 'Reply to Copp: Naturalism, Normativity, and the Varieties of Realism Worth Worrying About'. *Philosophical Issues* 18, no. 1 (2008): 207–28.
- . 'What Is Constructivism in Ethics and Metaethics?' *Philosophy Compass* 5, no. 5 (2010): 363–84.
- . 'Mind-Independence Without the Mystery: Why Quasi-Realists Can't Have It Both Ways'. In *Oxford Studies in Metaethics, Volume 6*, edited by Russ Shafer-Landau, 6:1–32. Oxford University Press, 2011.
- . 'Objectivity and Truth: You'd Better Rethink It'. *Oxford Studies in Metaethics* 11 (2016).
- Strevens, Michael. *Depth: An Account of Scientific Explanation*. Harvard University Press, 2008.
- Stroud, Barry. *Meaning, Understanding, and Practice: Philosophical Essays*. Oxford University Press, 2000.

- Swain, Marshall. ‘‘Justification and the Basis of Belief’’. In *Justification and Knowledge*, edited by George Pappas, 25–50. D. Reidel, 1979.
- Taylor, Elanor. ‘Against Explanatory Realism’. *Philosophical Studies* 175, no. 1 (2018): 197–219.
- Ted, Poston. *Reason and Explanation*. Palgrave Macmillan, 2014.
- Tersman, Folke. ‘Debunking and Disagreement’. *Noûs* 51, no. 4 (2017): 754–74.
- Vavova, Katia. ‘Debunking Evolutionary Debunking’. *Oxford Studies in Metaethics* 9 (2014): 76–101.
- Wielenberg, Erik J. ‘On the Evolutionary Debunking of Morality’. *Ethics* 120, no. 3 (2010): 441–64.
- . *Robust Ethics: The Metaphysics and Epistemology of Godless Normative Realism*. Oxford University Press, 2014.
- . ‘Ethics and Evolutionary Theory’. *Analysis* 76, no. 4 (2016): 502–15.
- Ye, Ru. ‘A Doxastic-Causal Theory of Epistemic Basing’. In *Well-Founded Belief: New Essays on the Epistemic Basing Relation*, edited by J. Adam Carter and Patrick Bondy, 15–33, 2020a.
- . ‘Higher-Order Defeat and Intellectual Responsibility’. *Synthese* 197, no. 12 (2020b): 5435–55.