



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Evolution
and
the Possibility of
Moral Knowledge

Silvan Wittwer

DOCTOR OF PHILOSOPHY
UNIVERSITY OF EDINBURGH
2018

MEINEN ELTERN
UND
MEINEM BRUDER

DECLARATION

I declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where states otherwise by reference or acknowledgment, the work presented is entirely my own.

August 30, 2018, Edinburgh

(Signature)

(Date & Place)

TABLE OF CONTENTS

Table of Contents	
Acknowledgements	
Abstract	
Lay Summary	
Chapter 1: Introduction	1
1. Introduction	1
2. Robust Moral Realism	2
3. Evolutionary Debunking Arguments	3
4. A Lacuna – And How to Fill It	6
5. Chapter Previews	7
Chapter 2: Explaining Reliability	13
1. Introduction	13
2. Evolutionary Debunking and Explaining Moral Reliability	14
3. Trivializing Moral Reliability by Modalizing It	18
4. Moral Reliability, Modal Stability and Epistemic Luck	20
5. Revisiting Evolutionary Debunking	27
6. Evolutionary Moral Scepticism at Last	31
7. Conclusion	36
Chapter 3: Higher-Order Evidence of Error	37
1. Introduction	37
2. Evolutionary Debunking and Higher-Order Evidence	39

3. The Threat of Self-Defeat	47
4. Avoiding Self-Defeat: A Strategy	49
5. The Total Evidence View	50
6. Evolution and Moral Unreliability	51
7. Evolution and Moral Disagreement	54
8. Conclusion	59
Chapter 4: Explanatory Dispensability	61
1. Introduction	61
2. Evolutionary Debunking and Explanatory Dispensability	62
3. The Nature of Evolutionary Explanations	68
4. Companions in Guilt	72
5. Deliberative Indispensability	81
6. Conclusion	89
Chapter 5: Evolution and the Possibility of Robust Moral Knowledge	91
1. Introduction	91
2. A Sketch of Moral Intuitionism	93
3. Moral Intuitionism and Epistemic Luck	97
4. Moral Intuitionism and the Best Evolutionary Explanation Argument	105
5. Conclusion	129
Bibliography	133

ACKNOWLEDGEMENTS

Writing a PhD thesis is no walk in the park. Rather, it is an often adventurous, always challenging, and ultimately satisfying hike. Any hike, though, is made infinitely better by proper hiking companions. And I have had the privilege of sharing my journey with truly wonderful ones.

First and foremost, I would like to thank my supervisors. Over the years, Matthew Chrisman has been the source of countless stimulating philosophical conversations, tremendously valuable feedback on my research, savvy career advice and entertaining insights into football. For all of that, I am truly grateful. I am also indebted to Duncan Pritchard, for supporting my ambition to become a metaethicist-cum-epistemologist, and to Geoff Sayre-McCord, for giving me the opportunity to visit Chapel Hill, teaching me how to teach as well as forcing me to always see the bigger philosophical picture.

In addition to my supervisors, many others have contributed towards my intellectual development over the last couple of years. A rotating cast of ‘work in progress buddies’, including Lucy Campbell, Joshua Blanchard, Marius Baumann, James Brown, Francois Jaquet and Felix Pinkert, helped me realize the value of daily writing – and that even (or: especially) bad ideas and rough drafts are worth talking about. Others shaped my project with kind words, fruitful questions, scathing objections, and helpful literature suggestions. Among them were Alfred Archer, Natalie Ashton, Cameron Boulton, Chris Cowie, Guy Fletcher, Giada Fratantonio, Camil Golub, Allan Hazlett, Sebastian Koehler, Elinor Mason, Joshua May, Maggie O’Brien, David Plunkett, Joey Pollock, Mike Ridge, Debbie Roberts, Stefan Riedener, Neil Sinclair, Patrick Todd, Lani Watson, Lee Whittington and Alan Wilson. Finally, my project benefitted from the feedback received at various talk venues, including the Universities of Budapest (CEU), Edinburgh, Nottingham, Sheffield and Tübingen. All those audiences deserve my thanks.

My time in Edinburgh would not have been as rewarding without all those who made it a second home to me. For their continuous encouragement, support and friendship, I am indebted to Liv Coombes, Ana Cretu, Giada Fratantonio, Blathnaid Greene & Matthew Steele, Iliana Lytra & George Tentokallis, Maggie O’Brien, Anna Ortin, Nina Poth, Rick Sendelbeck, Kegan Shaw, Ben Sworn, Lani Watson, Di Yang and especially Jonathan Macdonald and Francesca Bratton. (Apologies to anyone I may have forgotten!) Special thanks are reserved for Nick Rebol, who went way beyond the call of duty to help me with last minute formatting, and for Katie Keltie, who took care of all the administrative

business behind the scenes and offered some uplifting words whenever they were needed. Further, visiting Chapel Hill was a memorable and enriching experience, not least because of the following people: Amy Berg, Chris & Joanna & Annabelle Blake-Turner, Allende Cornego & Iain Stitt, Nilanjan Das, Chris Howard, Barry Maguire, Caroline & Philipp Lustenberger, Lindsay Brainard & Keshav Singh, Steve Swartzler, Caspar Wilson, and Alex Worsnip. Finally, I count myself blessed to be surrounded at almost all times with a group of very good, very old friends. Some of those make me miss home: Raphael Ochsenbein, Luca Liechti, Luise Baumgartner and especially Julian Reidy and Lukas Keller. Others never fail to provide me with a home away from home: Jönne, Jael & Johann Kriener, Felix Pinkert and Daniel Paffenholz.

But it wasn't just friends who helped me navigate academic life successfully. Even though it might be pretentious or overly dramatic to state, I probably wouldn't have finished writing my PhD without listening to post-rock. So, thanks to Godspeed You! Black Emperor, Mogwai and Russian Circles, amongst many others, for providing the right kind of soundtrack for many long, increasingly frantic creative efforts. Moreover, I am extremely grateful to two other staples of my Edinburgh life: Barbara for imparting the importance of self-care – and the good people at Pleasance for reminding me that, sometimes, the best way of thinking something through is not to think about it for a while and lift some weights instead.

I am also very grateful to various institutions whose financial support made my doctoral studies possible in the first place: the University of Edinburgh for their Principal's Career Development Scholarship, the Swiss Study Foundation, and especially the Swiss National Science Foundation as well as the Janggen-Pöhn Stiftung for giving me the freedom to focus on writing up during my final year.

Finally, I would like to express my deep gratitude to and love for my family. My parents have been unwavering in their emotional support, starting way before their wide-eyed teenage son impulsively decided to leave Switzerland to study Analytic philosophy at St. Andrews. And my brother has always had the most wonderful ability to keep me connected to my roots, ever so gently reassuring me that, all things told, I am and will always be a 'Bärner Gieu' at heart, wherever I happen to lay my head. This PhD thesis I dedicate to them.

ABSTRACT

This PhD thesis provides an extended evaluation of *evolutionary debunking arguments* in meta-ethics. Such arguments attempt to show that evolutionary theory, together with a commitment to robust moral objectivity, lead to moral scepticism: the implausible view that we lack moral knowledge or that our moral beliefs are never justified (e.g. Joyce 2006, Street 2005, Kahane 2011). To establish that, these arguments rely on certain *epistemic principles*. But most of the epistemic principles appealed to in the literature on evolutionary debunking arguments are imprecise, confused or simply implausible. My PhD aims to rectify that.

Informed by debates in cutting-edge contemporary epistemology, Chapter 1 distinguishes three general, independently motivated principles that, combined with evolution, seem to render knowledge of robustly objective moral facts problematic. These epistemic principles state that (i.) our getting facts often right in a given domain requires explanation – and if we cannot provide one, our beliefs about that domain are unjustified; (ii.) higher-order evidence of error undermines justification; and (iii.) for our beliefs to be justified, our having them must be best explained by the facts they are about. Chapters 2-4 develop and critically assess evolutionary debunking arguments based on those principles, showing that only the one inspired by (iii.) succeeds.

Chapter 2 investigates the argument that evolution makes explaining why we get moral facts often right impossible. I argue that Justin Clarke-Doane's recent response (2014, 2015, 2016, 2017) works, yet neglects an issue about epistemic luck that spells trouble for robust moral objectivity. Chapter 3 discusses the argument that evolution provides higher-order evidence of error regarding belief in robustly objective moral facts. I show that such an argument falls prey to Katia Vavova's (2014) self-defeat objection, even if evolutionary debunkers tweak their background view on the epistemic significance of higher-order evidence. Chapter 4 develops the argument that evolution, rather than robustly objective moral facts, best explains why we hold our moral beliefs. I offer a systematic, comprehensive defence of that argument against Andreas Mogensen's (2015) charge of explanatory levels confusion, Terrence Cuneo's (2007) companion in guilt strategy, and David Enoch's (2012, 2016) appeal to deliberative indispensability.

Chapter 5 brings everything together. It investigates whether robust moral objectivity survives the worry about epistemic luck raised in Chapter 2 and the explanatory challenge developed in Chapter 4. Making progress, however, requires a better idea of how we form

true, justified beliefs about and acquire knowledge of robustly objective moral facts. Since it offers the most popular and best-developed epistemology of robustly objective morality, my inquiry in Chapter 5 focuses on contemporary moral intuitionism: the view that moral intuitions can be the source of basic moral knowledge. I argue that its success is mixed. While moral intuitionism has the conceptual tools to tackle the problem of epistemic luck from Chapter 2, it cannot insulate knowledge of robustly objective moral facts against the sceptical worry raised by the evolutionary debunking argument developed in Chapter 4. Thus, evolutionary theory, together with a commitment to robust moral objectivity, *does* lead to a form of unacceptable moral scepticism.

LAY SUMMARY

Is morality a matter of mere taste? Or is it, as many would claim, objective? Those of us who lean towards the latter believe that it is a matter of moral fact that, say, killing is wrong. What is more, we take ourselves to *know* that moral fact. However, some philosophers and biologists have recently suggested that these commitments are incompatible with evolutionary theory. Even if there are objective moral facts, they argue, we did not evolve to find them out. Instead, we evolved to survive. Given that evolution doesn't favour our finding out about moral facts, how could we ever know any? How could we know, for example, that killing is wrong? So, evolution seems to undermine our putative knowledge of objective morality.

My PhD thesis evaluates this evolutionary critique in three steps. First, it clarifies the epistemic principles (about knowledge, justification, evidence, etc.) on which this critique rests. Since those principles are often imprecise or implausible, I propose three more compelling ones. For instance, my first epistemic principle states that our getting moral facts mostly right requires explanation – which evolutionary theory renders impossible. Second, my thesis develops and critically assesses the evolutionary critique of objective moral knowledge in light of those more compelling epistemic principles. As I argue, evolutionary theory does threaten objective moral knowledge. For our moral beliefs to be justified, our having them would have to be best explained by objective moral facts. But that is not the case, for evolutionary facts best explain our moral beliefs. Third, the thesis investigates how defenders of objective moral knowledge could respond to this evolutionary challenge. Making progress requires a better idea of how we acquire such knowledge. Since it offers the most popular and best developed account of objective moral knowledge, my inquiry focuses on contemporary moral intuitionism: the view that moral intuitions can be the source of basic moral knowledge. However, I argue that moral intuitionism lacks the conceptual resources to tackle the evolutionary challenge. Thus, the evolutionary critique has been vindicated: evolutionary theory does undermine our knowledge of objective morality. We cannot know the objective moral fact that, say, killing is wrong.

CHAPTER 1: INTRODUCTION

1. Introduction

My PhD thesis provides an extended evaluation of *evolutionary debunking arguments* in meta-ethics. Such arguments attempt to show that evolutionary theory, together with a commitment to robust moral objectivity, leads to moral scepticism: the implausible view that we lack moral knowledge or that our moral beliefs are never justified. The idea is that, even if there are robustly objective moral facts, we did not evolve to find them out. Rather, we evolved to survive. And while our survival may require certain moral beliefs, it does not require or even encourage ones that would constitute knowledge of robustly objective moral facts. In that manner, awareness of the evolutionary origins of our capacity to form moral beliefs is supposed to lead to scepticism about robustly objective morality.

To get from robust moral objectivity and evolution to moral scepticism, these arguments rely on certain *epistemic principles*. But most of the epistemic principles appealed to in the literature on evolutionary debunking arguments are imprecise, confused or simply implausible. As a result, most of these arguments aren't epistemologically sound and therefore fail. My PhD thesis aims to reconsider the evolutionary threat to knowledge of robustly objective moral facts in light of more precise and plausible epistemic principles. To achieve that, I proceed in three steps. Chapter 1 distinguishes three epistemic principles that, combined with evolution, do indeed seem to render knowledge of objective moral facts problematic. These principles are informed by cutting-edge debates in contemporary epistemology, from the nature of reliability to the epistemic significance of higher-order evidence and the relevance of explanation to justification. Then, Chapters 2 to 4 develop new and improved evolutionary debunking arguments based on those principles. Finally, Chapter 5 brings everything together: it investigates whether there is any way for a commitment to robust moral objectivity to survive in the face of these arguments. In the end, I think that these new and improved evolutionary debunking arguments don't imperil the mere possibility of robustly moral knowledge. However, they undermine any claim to know what robust morality is *substantively* like. So, evolutionary theory, together with a commitment to robust moral objectivity, *does* lead to an implausible form of moral scepticism.

This chapter – as just mentioned – takes the first step towards a comprehensive assessment of evolutionary debunking arguments in meta-ethics. Its aim is to set the stage, clarify the theoretical background, introduce the state of the literature and motivate my emphasis on

three specific epistemic principles. Its structure is as follows. I first define robust moral realism, the view that conceives of moral facts as robustly objective (§2). Then, I sketch the basic evolutionary challenge to robust moral realism (§3), before identifying a lacuna, namely the lack of compelling underlying epistemic principles, and proposing in generic terms how to fill it (§4). I conclude with a preview of the remainder of my PhD thesis (§5).

2. Robust Moral Realism

Robust moral realism is the meta-ethical view that there are mind-independent moral facts which cannot be reduced to natural facts, yet are *in principle* knowable.¹ For taxonomical purposes, it proves helpful to think of any meta-ethical view as addressing four key aspects of moral thought or discourse: its psychology, semantics, metaphysics and epistemology.² Robust moral realism is no different. First, the view is an instance of *cognitivism* in moral psychology: the doctrine that moral judgments are beliefs as opposed to desires or other conative states. Second, robust moral realism subscribes to *truth-conditional* moral semantics: moral concepts get their meaning by purporting to refer to robust moral properties. Third, the view is metaphysically committed to the existence of *robust moral facts*. What are robust moral facts? To begin with, such facts are objective or mind-independent: they hold independently of whether we form beliefs about them. Further, those facts are ‘robust’: they resist reduction to or identification with facts as described by the natural and social sciences. By implication, they are also typically assumed to lack causal powers. Fourth and finally, robust moral realism is a form of *moral non-scepticism*: it holds that we can have knowledge of robust moral facts, at least under epistemically favorable conditions.³ Together, I take those four core commitments to characterize robust moral realism and thus provide the theoretical background for my investigation.

Before proceeding, two points about the scope of my inquiry are in order. First, for the purposes of my argument, I take the psychology, semantics and metaphysics of robust moral realism for granted. As far as I am concerned, moral judgments are indeed beliefs whose constituent concepts at least sometimes refer to robust moral properties. However,

¹ Prominent contemporary defenders of the view include Cuneo 2007, Enoch 2012, Fitzpatrick 2008, Huemer 2005, Oddie 2005, Parfit 2013, Shafer-Landau 2003, and Wedgwood 2007. I prefer ‘robust’ to ‘non-naturalist’ due to the difficulties involved in defining the latter. See Ridge 2018.

² See Chrisman 2016: Introduction for more on that.

³ For Chapters 2 to 4, where I assess different strands of evolutionary debunking arguments, we don’t need to assume anything more specific about how exactly we form true, justified beliefs about and know moral facts. In other words, my assessment is largely neutral with respect to the correct first-order moral epistemology. That changes in Chapter 5, however, where I examine whether moral intuitionism has the conceptual tools to tackle any residual epistemic worries raised (or inspired by) evolutionary debunking arguments.

I want to question the epistemological core commitment of robust moral realism, namely moral non-scepticism. More specifically, my PhD thesis asks: assuming (possibly for *reductio*) the existence of robust moral facts, evolutionary explanations of how we form moral beliefs and certain independently plausible epistemic principles, can we ever know or even justifiably believe in robust moral facts?

Second, my inquiry will be limited to robust moral realism as the target of evolutionary debunking arguments. Admittedly, that goes against some recent developments in the relevant literature. Some have argued that the evolutionary challenge (or something inspired by it) also arises for other meta-ethical views, including moral naturalism and quasi-realism.⁴ But there is a good methodological reason for my focus on robust moral realism. After all, I intend to figure out whether evolutionary debunking arguments raise an epistemic problem *at all*. To do that, it seems promising to start with a meta-ethical view that is as metaphysically extravagant as robust moral realism, thus offering the evolutionary debunker a comparably easy target.⁵ It is only after (if at all) having successfully isolated the epistemic worry inspired by evolutionary theory that we should concern ourselves with the vulnerability other meta-ethical views. So, my investigation examines evolutionary debunking exclusively against the backdrop of robust moral realism.⁶

3. Evolutionary Debunking Arguments

Robust moral realism has long been criticized on epistemological grounds.⁷ However, an empirically informed epistemological challenge has recently garnered significant philosophical attention.⁸ *Evolutionary debunking arguments* attempt to show that knowledge

⁴ See, for example, Barkhausen 2016 and Golub 2017, respectively.

⁵ This echoes Bedke (2014), claiming that robust moral realism is ‘a nice test case for seeing that whether there is an epistemic problem at all’ (2014: 6).

⁶ I recognize that evolutionary debunking arguments are often developed not just to criticize robust moral realism, but to motivate alternative meta-ethical views such as moral constructivism (e.g. Street 2006, 2008) or moral error theory (e.g. Joyce 2006). However, my thesis focuses on the complex task of establishing the viability of these arguments against robust moral realism. What exactly follows for the broader meta-ethical landscape lies beyond its scope.

⁷ See, for example, Mackie 1977 on the epistemological queerness of robust moral facts.

⁸ The two most influential contributions to date are Joyce (2006) and Street (2006). Since their defense of evolutionary debunking arguments, the meta-ethical literature on the merit of these arguments has exploded. (According to Google Scholar, more than 15100 articles with keywords ‘evolutionary debunking’ were published between 2005 and 2018.) To name just a few: Bedke 2008, Berker 2014, Brosnan 2011, Copp 2008, Enoch 2011: Ch. 7, Fitzpatrick 2015, Joyce 2008 & 2014, Kahane 2011, Parfit 2013, Schafer 2010, Shafer-Landau 2012, Skarsaune 2011, Street 2008 and Wielenberg 2010. There also have been several recent contributions to the very renowned *Oxford Studies in Metaethics*, including Bedke (2014), Sinclair (2018) and Vavova (2014). For more literature, see Chapters 2 to 4.

of robust moral facts stands in tension with evolutionary theory. These arguments share the following general structure:⁹

- (1) There are robust moral facts.
- (2) Evolutionary theory explains how we form beliefs, including moral beliefs.
- (3) If evolutionary theory explains how we form moral beliefs, we cannot justifiably believe or know robust moral facts.
- (4) Therefore, we cannot justifiably believe or know robust moral facts.

The first premise simply reiterates the metaphysical commitments of robust moral realism: that there are moral facts, that those facts are objective, and that they are robust. As mentioned above, I shall take that for granted. So, the first premise does no more than articulating the dialectical background.

Meanwhile, the second premise appeals to the evolutionary explanation of moral beliefs. Very roughly, such an explanation states that we believe what we believe, morally speaking, in virtue of natural selection.¹⁰ For instance, we believe that incest is morally wrong or that we have special moral obligations to family members because thinking with those moral concepts proved *adaptive* or helped our ancestors survive.¹¹ Now, the evolutionary explanation of our moral beliefs differs significantly from that of our perceptual beliefs about, say, mid-sized objects. It is overwhelmingly plausible that evolution has selected for their *truth*. After all, having true perceptual beliefs would have enhanced the fitness of our ancestors, helping them locate food and detect environmental

⁹ My template follows Kahane (2011) – with two major differences. First, he takes the first premise to say that robust moral realism is correct. Given my definition of robust moral realism as epistemologically optimistic, that would make above template a *reductio ad absurdum*. That, though, seems too strong to me: evolutionary debunking arguments don't show that robust moral realism is absurd. Rather, they show, if successful, that robust moral realism faces a significant epistemological challenge. Second, Kahane spells out the third premise in terms of off-track evolutionary influence. Tracking, however, suggests Nozick's (1981) sensitivity, a modal condition on knowledge. Even though I discuss a modal interpretation of premise 3 in Chapter 2, I don't think all evolutionary debunking arguments rely on modality. Therefore, my template opts for a more open formulation of premise (3).

¹⁰ This is obviously loose talk: it is implausible that evolution explains *particular* moral beliefs. Rather, it explains our dispositions to apply moral concepts (e.g. Joyce 2001, 2006) or basic evaluative dispositions that shape our moral beliefs (e.g. Street 2006, 2008, 2011). However, I follow the literature on evolutionary debunking in assuming that this complication doesn't matter for my argument above.

¹¹ I will have more to say about the nature of such evolutionary explanations in Chapter 4. However, for my purposes, their exact details don't really matter for the most part – as long as such explanations don't refer to robust moral facts. By side-stepping the empirical details and focusing instead on the epistemic significance of evolutionary explanations, I follow most of the literature on evolutionary debunking. (See, for example, Berker 2014.) But for more on evolutionary moral psychology, see Fraser 2014, Joyce 2006: Chs. 1-4, and Mogensen 2014: Ch. 1 & 2.

threats. So, the evolutionary explanation of perceptual beliefs refers to perceptual facts.¹² However, importantly, the evolutionary explanation of *moral* belief doesn't avert to robust *moral* facts. After all, evolution hasn't selected for *true* moral beliefs, but for those that are most fitness-enhancing. As Richard Joyce (2006) writes: '[w]hether we assume that the concepts *right* and *wrong* succeed in denoting properties in the world, or whether we think that they suffer from a referential failure that puts them on a par with the concepts *witch* and *ghost*, the plausibility of the hypothesis concerning how moral judgments evolved remains unaffected' (183). Similarly, Sharon Street (2008) writes that '...the best explanation of why we tend to value our survival is not that it's independently true that our survival is valuable ... but rather, much more simply, that creatures who valued their survival tended to do what promoted it, and therefore left more descendants' (209). So, as the second premise states, evolutionary forces, not robust moral facts, explain why we hold certain moral beliefs.

The third premise articulates an epistemic principle that ties the evolutionary explanation of moral belief to the defeat of moral justification or knowledge.¹³ More specifically, it says that if evolutionary theory explains how we form moral beliefs, then those beliefs cannot constitute knowledge or even count as justified. But why not? There are various ways of spelling that out. To illustrate, here are the two most influential recent formulations of the epistemic principle.

First, Joyce (2006: 179-84) compares becoming aware of the evolutionary origins of our moral beliefs to finding out that we have been slipped belief pills: pills that made us form the true belief that, say, Napoleon lost at Waterloo. While this realization doesn't show that our belief about Napoleon's campaign is false, it seems to defeat its justification or '...undermine your faith in your belief that Napoleon lost Waterloo' (179). And once we become aware of their evolutionary origins, moral beliefs are no different: we should '...cultivate agnosticism regarding all positive beliefs involving these [moral] concepts until we find some solid evidence either for or against them' (181). In both cases, therefore, finding out about the origins of your beliefs defeats their justification.

Second, Street (2006, 2008) argues that evolutionary explanations of moral belief raise a puzzle for the robust moral realist. In particular, they must specify the relationship between our evolved moral attitudes and the robust moral truth. But tackling that task lands them in a 'Darwinian dilemma'. Either, robust moral realists could deny that the

¹² The same evolutionary vindication might extend to basic scientific, arithmetic and logical beliefs, given that reliable formation of such beliefs plausibly proved fitness-enhancing as well. See Joyce 2006: Ch. 6 for more.

¹³ We usually distinguish between rebutting and undermining defeaters. Putative evolutionary defeaters are most plausibly thought of as the latter, not the former. (See, for instance, Mogensen 2015.) So, in what follows, I shall mean that by 'defeater', dropping the qualifier.

relationship between evolved moral beliefs and robust moral truths is anything other than random. That, though, is highly problematic: it forces the robust moral realist to acknowledge that ‘our normative judgments are in all likelihood hopelessly off track’ (208) – or means that “an astonishing coincidence took place [...], that as a matter of sheer luck, evolutionary pressures affected our evaluative attitudes in such a way that they just happened to land on or near the true normative views among all the conceptually possible ones’ (ibid.). Since both implications seem implausible, this amounts to the first horn of the dilemma. Alternatively, the robust moral realist could posit a positive relationship between our evolved moral beliefs and robust moral truths: that it was somehow adaptive for our ancestors to grasp or track robust moral truths. But that is scientifically untenable: as stated above, evolutionary explanations of moral belief do not avert to robust moral facts. Evolution hasn’t selected for *true* moral beliefs, but for those that are most fitness-enhancing. This constitutes the second horn of the dilemma. Caught between a rock and a hard place, robust moral realists seem forced to conclude that we cannot justifiably believe or know robust moral facts.

Together, premises (1) – (3) imply *moral scepticism*: we cannot justifiably believe or know robust moral facts. So, knowledge of robust moral facts seems incompatible with evolutionary theory. But, of course, moral scepticism is unacceptable to robust moral realists. As we saw in §2, their view is committed to a form of moral non-scepticism, according to which we can have robust moral knowledge in epistemically favorable circumstances. So, evolutionary debunking arguments give us reason to reject robust moral realism – or, at least, to deem it significantly less attractive than initially thought.

4. A Lacuna – And How to Fill It

Despite its initial plausibility, the evolutionary debunking argument above isn’t sound. The problem lies with premise (3), the epistemic principle tasked with licensing the inference from the existence of robust moral facts and evolutionary explanation to moral scepticism. As it turns out, most attempts at substantiating that principle found in the literature are imprecise, confused or simply implausible. To begin to see that, I want to revisit and closely examine the two influential recent proposals introduced above.

Recall that Joyce (2006) compares becoming aware of the evolutionary origins of our moral beliefs to finding out that we have been slipped pills making us form the true belief that Napoleon lost at Waterloo. In both cases, he argues, what we find out defeats the beliefs’ justification. But what seems to defeat the pill-induced beliefs is our finding out about their *groundlessness*. Our evolved moral beliefs, though, do have grounds.¹⁴ For instance, take my belief that pain is *pro tanto* bad. That belief has a multitude of grounds:

¹⁴ See White 2010: 575f.

that pain hurts, that most people don't like it, that it seems intuitively bad, that denying the truth of its contents strikes most as absurd, and so forth. So, Joyce seems to appeal to an implausible epistemic principle.

Similarly, the first horn of Street's (2006, 2008) Darwinian dilemma alleges that it is conceptually *possible*, given evolution, that our moral beliefs are adaptive, yet massively *false* – and that we have no good independent reason to rule that out. So, robust moral realists should become moral sceptics. However, her argument relies on an epistemic principle familiar from arguments for scepticism about the external world: if you have no good independent reason to think that your perceptual beliefs are true, you cannot rationally maintain them.¹⁵ Those arguments, though, develop their threat without relying on evolution (or any other empirical assumption) and they don't just apply to moral or normative beliefs. So, the epistemic principle that Street appeals to seems imprecise, confused and ill-suited for evolutionary debunking.

Evolutionary debunkers are thus in desperate need of more compelling ways of spelling out premise (3) of their argument. That is the *lacuna*. To fill that lacuna, I propose that we examine three clear, precise and more plausible epistemic principles which might work for evolutionary debunking arguments. Together with evolutionary theory, these principles indeed pose a threat to knowledge of robust moral facts. The principles I have in mind are informed by ongoing debates in contemporary epistemology about whether moral reliability requires explanation (which evolution complicates), whether higher-order evidence of error (which evolution arguably provides) undermines justification and whether robust moral facts earn their epistemological keep by being cited in the explanation of moral beliefs (which evolutionary explanations threaten). Evolutionary debunkers *should* rely on something like one of these epistemic principles as premise (3) of their argument – and may sometimes *implicitly* do. For instance, the intuitive force of Joyce's (2006) belief-pill analogy above plausibly comes down to receiving higher-order evidence of error that undermines justification. It is just that he fails to make that explicit. Street (2006, 2008) can be charitably recast as questioning whether robust moral realists can explain why they are morally reliable. But what would evolutionary debunking arguments based on these three principles look like? The final section introduces them in more detail and provides guidance to Chapters 2 to 5.

5. Chapter Previews

To conclude Chapter 1, I shall briefly elaborate on each epistemic principle – and, by doing so, preview the remainder of my PhD thesis.

¹⁵ See Elga ms, Vavova 2015:15 and White 2010: 604.

The first epistemic principle states that reliability about a given domain requires explanation – and that, if we fail to provide one, our beliefs about that domain are unjustified. When combined with the thought that evolution complicates such an explanation for belief in robust moral facts, this principle threatens to undermine the justification of moral beliefs, robustly construed. Chapter 2 closely examines that form of the evolutionary debunking argument.

To see in brief how this evolutionary debunking argument works, assume that most of our moral beliefs are true or that we are mostly morally *reliable*: we get things mostly right about morality. What explains our success? On reflection, it is difficult to see how robust moral realists could explain such success.¹⁶ To start with, robust moral facts – by definition – don't *cause* our moral beliefs. So, we can neither directly perceive (see, hear, smell, etc.) them and nor rely on empirical investigation of the world to find out about them. Further, robust moral facts cannot *depend* on our moral beliefs. After all, they are objective: they hold independently of what we think of moral matters. So, they cannot be a function or construction of our moral beliefs. Finally, *evolution* cannot explain our moral reliability either: evolution favors *adaptive*, not true, moral beliefs. But if we cannot explain why we get robust moral facts mostly right, why think that we do? It looks like we are no longer *justified* in holding our moral beliefs. Moral scepticism follows.

But what exactly does 'explaining reliability' mean? Or, put differently, what form would such an explanation take? Justin Clarke-Doane (2015, 2016, 2017, forthcoming) has argued that a *good* explanation of moral (mathematical, logical, etc.) reliability would show that our moral (mathematical, logical, etc.) beliefs are *stable* in a certain way. More specifically, our moral beliefs would have to be both *sensitive* to the robust moral facts (such that if those facts changed, so would our beliefs) and *safe* from error (such that they could not easily be false).¹⁷ But, Clarke-Doane continues, our moral beliefs already are. After all, they are sensitive to the robust moral facts by default: since such facts hold necessarily, they could not possibly change. So, any true moral belief must be sensitive as well. Similarly, our moral beliefs are safe from error: they are true (as we assume) and could not easily have evolved differently. So, our moral beliefs could not easily be false. In short, according to Clarke-Doane, once we realize what a good explanation of moral reliability looks like, it also dawns on us that we already have one. As a result, any evolutionary debunking argument based on the idea that evolution makes explaining moral reliability very hard does not work.

Clarke-Doane's attempt at securing the sensitivity and safety of our robust moral beliefs raises various questions. For instance, does it really succeed? Is the task of showing that

¹⁶ Consequently, this line of argument is often called the *reliability challenge*. For more, see Chapter 2.

¹⁷ For more technical definitions of these modal conditions on knowledge, see Chapter 2.

our moral beliefs are sensitive and safe really all there is to explaining moral reliability? Or are there residual epistemic worries in the vicinity? And does this task plausibly underlie evolutionary debunking arguments? So, further discussion seems required, which Chapter 2 of my PhD thesis provides. In it, I argue that Clarke-Doane's response succeeds against evolutionary debunking, but neglects an issue about epistemic luck that spells trouble for robust moral realism.

The second epistemic principle says that higher-order evidence of error undermines justification. When combined with the idea that evolution provides us with such higher-order evidence regarding belief in putative robustly objective moral facts, this principle quickly leads to moral scepticism. Chapter 3 investigates whether such an evolutionary debunking argument is viable.

What is higher-order evidence of error? Higher-order evidence of error is evidence that something is wrong with one's first-order evidence or cognitive functioning. Suppose that you are a competent meteorologist, reliably predicting the weather for your local TV station, based on data from the national weather service. You predict that there will be light rain in your area tomorrow afternoon. But then you learn that a fellow meteorologist from another local TV station disagrees, despite being equally reliable and using the same data as you. Here, finding out about that disagreement gives you higher-order evidence of error: it tells you that there is something wrong with your first-order evidence, namely the data from the national weather service, or how you assessed it.¹⁸ Now, what does this evidence do to your belief that it is going to rain tomorrow afternoon? Many believe that such higher-order evidence of error defeats the *justification* of your belief.¹⁹ After all, your fellow meteorologist is in the same epistemic situation as you – and you are therefore both equally likely to be mistaken. As a result, you should be significantly less confident in your own prediction of the tomorrow afternoon's local weather.

But why think that evolution provides us with higher-order evidence of error with respect to our moral beliefs? To see why, it helps to consider a more general phenomenon first: sometimes, thinking about the causes of our beliefs (or, their etiology) can provide us with higher-order evidence of error.²⁰ Suppose that you grow up in a liberal community, coming to hold corresponding political beliefs. Once you attend university, though, you

¹⁸ For more on *peer disagreement*, see Christensen 2007, Elga 2007, Feldman 2006, or Kelly 2005.

¹⁹ This view is often called *Conciliationism*. For Conciliationism regarding peer disagreement, see, for instance, Christensen 2007, Elga 2007 or Feldman 2006. For Conciliationism regarding higher-order evidence more generally, see Christensen 2010 and Lasonen-Aarnio 2014. Conciliationism is by no means universally accepted. For criticism, see Kelly 2005, 2010 or Lackey 2008. I will have more to say about Conciliationism in Chapter 3.

²⁰ Often, these appeals are either fallacious or innocuous (see Dworkin 1996). But there are genuinely problematic cases covered by an emerging literature, starting with Cohen's (2000: 16-19) and culminating (so far) in White 2010 and Vavova 2018. For more, see Chapter 3.

realize that you could easily have grown up in a conservative community instead. In that case, you would hold politically conservative beliefs now. So, you hold liberal beliefs *just because* you grew up in a liberal community. That insight seems unsettling, should give you pause and forces you to re-examine your first-order evidence in support of your political convictions. Here, then, thinking about the causes of your beliefs gives you higher-order evidence of error.²¹ And many have thought that this evidence defeats the justification of your politically liberal beliefs. Importantly, our moral beliefs are like your political beliefs in that respect: we hold them *just because* we evolved in certain ways. Had we evolved in different ways, we would hold different moral beliefs. For instance, we might have evolved to believe that we have a moral obligation to kill our firstborns. Again, that insight seems unsettling, forcing us to re-examine our first-order moral evidence. So, thinking about the evolutionary origins of our moral beliefs gives us higher-order evidence of error.²² And that evidence defeats the justification of our moral beliefs. As a result, moral sceptics looms.

This evolutionary debunking argument looks attractive, drawing on interesting observations from recent debates about the nature of higher-order evidence and the epistemic significance of etiology. Still, it has not received much attention. To date, only Bogardus (2016) and Mogensen (2016) have put forth an argument along those lines. So, again, further discussion seems required, which I offer in Chapter 3. I argue that such an argument ultimately cannot avoid Katia Vavova's (2014) self-defeat objection, even if the evolutionary debunker adopts an attractive alternative background view on the epistemic significance of higher-order evidence of error.

The third and final epistemic principle I consider states that, for our beliefs to be justified, our having them must be best explained by the facts they are about. When combined with the idea that evolution, not robust moral facts, provides the best explanation of our moral beliefs, this principle can quickly lead to moral scepticism: our moral beliefs are unjustified. Chapter 4 develops and critically assesses an evolutionary debunking argument along those lines.

The principle articulates a plausible *explanatory* constraint on justified belief. For instance, consider your *perceptual* belief that there is a coffee mug on the desk in front of you. If it turned out that the fact that there is indeed a coffee mug on the desk in front of you did *not* best explain why you hold the corresponding belief, your belief would – intuitively – not be justified. The same seems to hold true for our *scientific* beliefs. Suppose you form

²¹ There is a nascent debate about what kind of higher-order evidence etiology provides. According to some, it is peer disagreement, according to others, fallibility. See Chapter 3 for more.

²² Again, contributors to the literature disagree about whether evolutionary theory provides evidence of possible moral peer disagreement or evidence of moral unreliability for robust moral realists. We shall engage with that issue in depth in Chapter 3.

the belief that a proton is in the cloud chamber, based on your observation of a vapor trail (which you know to reliably indicate that). Here, your belief is best explained by the fact that there is indeed a proton in the cloud chamber. So, it seems plausible that, for our beliefs to be justified, they must best explained by the facts they are about. But what about our (robustly) moral beliefs? Here, matters get slightly more complicated. Suppose you believe that killing your own children would be a serious moral wrong (at least *ceteris paribus*). What best explains that belief? To start with, it seems implausible that your belief is directly *caused* by the robust moral fact that killing your own children is seriously wrong. After all, such facts don't have causal powers. So, you cannot *perceive* moral wrongness. But neither is it plausible that you rely on empirical investigation of the world to figure it out. Instead, what best explains your moral belief might be *evolution*: killing your own children would not be adaptive, for it ends their survival prospects. So, more generally, facts about evolution, not robust moral facts, seem to best explain why we hold our moral beliefs. That, though, once again quickly leads to moral scepticism, given that, for our beliefs to be justified, the facts they are about must figure in best explanation of why we have them.²³

Surprisingly, despite its apparent simplicity, few have examined this evolutionary debunking argument in depth.²⁴ So, the discussion I provide in Chapter 4 is pressing. More specifically, I offer a systematic, comprehensive defense of an argument along those lines, showing that neither Andreas Mogensen's (2015, 2016) recent observations about the nature of evolutionary explanation, nor Terrence Cuneo's (2007) companion in guilt strategy nor David Enoch's (2012, 2016) appeal to deliberative indispensability should make us doubt its soundness.

By the end of Chapter 4, I will have completed the second step of my investigation. More specifically, I will have developed and critically assessed three new versions of the evolutionary debunking argument, based on clear and independently plausible epistemic principles. My discussion will have shown that anybody committed to robust moral objectivity faces two distinct epistemic worries. The first, articulated in Chapter 2, is loosely inspired by evolutionary debunking and centers on epistemic luck. Meanwhile, the second, as discussed in Chapter 4, is directly raised by evolutionary considerations, focusing on whether robust moral facts figure in the best explanation of why we hold our moral beliefs.

Chapter 5 investigates whether robust moral realism can deal with both the worry about epistemic luck raised in Chapter 2 and the explanatory challenge developed in Chapter 4. However, making progress on that issue requires a better idea of how we form true,

²³ Of course, this is an evolutionary version of Harman's (1977) challenge.

²⁴ Woods 2016 is the notable exception. More on him follows in Chapter 4.

justified beliefs about and acquire knowledge of robust moral facts. Since it is the most popular and best-developed moral epistemology available to robust moral realism, my inquiry in Chapter 5 focuses on contemporary moral intuitionism, the view that moral intuitions can be the source of basic moral knowledge. I argue that its success is mixed. While moral intuitionism has the conceptual tools to tackle the problem of epistemic luck developed in Chapter 2, it cannot insulate robust beliefs about *substantive* moral matters against the worry raised by best evolutionary explanation argument in Chapter 4. As a result, robust moral realism is saddled with unacceptable moral scepticism: we cannot justifiably believe substantive facts about morality, robustly construed. While evolution doesn't imperil the mere possibility of robustly moral knowledge, it undermines any claim to know what robust morality is substantively like.

Admittedly, the discussion of my PhD thesis will be most directly relevant to those invested in the debate about evolutionary debunking specifically or the viability of robust moral realism more generally. Still, the dialectic of Chapters 2 to 5 should also, more broadly, be of interest to meta-ethicists or epistemologists who don't follow the evolutionary debunking debate or who think that robust moral realism is a non-starter for other reasons. How so?

As we have seen, my discussion will revolve around three general, independently plausible epistemic principles, having to do with the nature of reliability to the epistemic significance of higher-order evidence and the relevance of explanation to justification. Importantly, those principles are central to any foundational debate in the epistemology of the normative and, more generally, the *apriori*. For example, it would be fruitful to ask whether reliability about logic requires explanation – or what role prudential facts exactly play in the formation of beliefs about prudential value. Further, my investigation doesn't just appeal to those principles in order to gauge the viability of different evolutionary debunking arguments. Rather, by examining how those principles work (or don't work) in the context of evolutionary debunking, we also generate insights into how to formulate them most compellingly – and how to apply them to other domains. Put differently, through *applying* these general epistemic principles extensively to the evolutionary debunking of robust moral belief, we find ways of *refining* them. For instance, if my line of argument in Chapter 2 is correct, 'explaining reliability' is best understood as a requirement on normative or apriori knowledge, not justified normative or apriori belief. In that manner, my discussion should be of interest to anyone working on the epistemological foundations of the normative and, more generally, the *apriori*.

CHAPTER 2: EXPLAINING RELIABILITY

1. Introduction

Evolutionary debunking arguments attempt to show that the evolutionary theory, when combined with a commitment to robust moral objectivity, leads to moral scepticism: the view that we lack moral knowledge or that our moral beliefs are never justified. But which epistemic principle licenses that inference? Recently, a promising suggestion has garnered significant philosophical attention: the fact that we are reliable in a given domain requires explanation – and if we cannot provide one, our beliefs about that domain are unjustified.¹ When combined with the idea that reflection on evolution undermines such an explanation in the moral case, this principle implies moral scepticism.

In response, Justin Clarke-Doane (2015, 2016, 2017, forthcoming) has argued that evolutionary considerations do not undermine an explanation of reliability in the moral case. According to him, a good explanation of moral reliability requires that our robust moral beliefs be *modally stable*. More specifically, our moral beliefs would have to be both *sensitive* to the robust moral facts (such that if those facts changed, so would our beliefs) and *safe* from error (such that they could not easily be false). But, Clarke-Doane continues, our moral beliefs already are: they are trivially sensitive, given that robust moral facts hold necessarily, and safe, given their actual truth and the fact that they could not easily have evolved differently. In short, once we realize what a good explanation of moral reliability looks like, it also dawns on us that we already have one.

In this chapter, my aim is to argue that Clarke-Doane's line is mistaken: as moral instances of Gettier (1963) cases illustrate, robust moral beliefs can be epistemically lucky – and therefore insensitive and unsafe. That insight, in turn, allows me to recast the principle that reliability requires explanation. Its explanatory demand isn't trivial. Rather, it forces robust moral realists to offer an account of why their moral beliefs, even if these are necessarily true and justified, aren't epistemically lucky. And if they fail to do so, they cannot claim to know robust moral facts. However, as I shall argue, reconceiving of the principle in that manner has significant implications for evolutionary debunking. Since the principle doesn't plausibly combine with evolutionary considerations to yield moral scepticism, it cannot serve as the key premise of an evolutionary debunking argument. So,

¹ For some relevant references, see footnote 2 below. 'Robust moral belief' is a shorthand for 'moral belief about (or purporting to be about) robust moral facts'. 'Robust' doesn't imply anything about its epistemic credentials.

my response to Clarke-Doane highlights a worry about epistemic luck widely overlooked by robust moral realists. But, more importantly, my discussion shows that evolutionary debunking arguments, conceived as challenges to explain the reliability of moral belief, fail. As a result, we need a different epistemic principle for evolution to debunk robust moral knowledge.

Here is my plan. In §2, I introduce the principle that reliability requires explanation and explain its putative role in evolutionary debunking arguments. In §3, I present Clarke-Doane's criticism in more detail. In §4, I argue – *contra* Clarke-Doane – that robust moral beliefs aren't already modally stable and proceed to spell out the implications for our understanding of the principle that reliability requires explanation. In §5, I show that the reconceived principle cannot plausibly serve as the key epistemic premise of the evolutionary debunking argument developed in §2. In §6, I defend my line of argument against two recent alternative formulations of the principle by Kieran Setiya (2013) and Matt Bedke (2014), respectively.

2. Evolutionary Debunking and Explaining Moral Reliability

Evolutionary debunking arguments attempt to show that knowledge of robust moral facts stands in tension with evolutionary theory. To succeed, these arguments must rely on epistemic principles that license the inference from the commitment to the robust objectivity of moral facts and the evolutionary explanation of moral belief to moral scepticism. However, as we saw in Chapter 1, identifying such principles proves harder than perhaps initially thought.

In response, many have recently made the following promising suggestion: reliability requires explanation – and if we fail to provide one in the moral case, our moral beliefs are unjustified. When combined with the thought that evolution renders such an explanation impossible, this principle would force the robust moral realists to embrace moral scepticism. Put more formally, the resulting argument reads as follows:²

² Both proponents and critics of evolutionary debunking seem to advance an argument along those lines, either explicitly (e.g. Bedke 2014, Clarke-Doane 2012 & 2013 & 2015, Enoch 2010 & 2012: Ch. 7, Schechter 2010, Setiya 2013: Ch. 2-4) or implicitly (e.g. Berker 2014, Copp 2008, Schafer 2010, Shafer-Landau 2012, Street 2006 & 2008). However, there is disagreement about the exact role played by evolutionary explanations in the argument. See below for more on that.

- (1) If we fail to explain why we are reliable in a given domain, we cannot justifiably believe facts about that domain.
- (2) Evolutionary explanations of moral belief make it impossible to explain why we are reliable with respect to morality, robustly construed.
- (3) Therefore, we cannot justifiably believe robust moral facts.³

In the remainder of this section, I shall briefly explain what these premises mean, why they might be true – and how they jointly seem to entail the conclusion.

The first premise articulates the principle at the heart of the so-called *reliability challenge* to robust forms of realism about the *apriori*.⁴ To understand how that challenge arises in the moral case, assume first that most of our moral beliefs are true or that we are mostly morally *reliable*: we get robust moral facts mostly right. What explains that? On reflection, that is difficult to see. To start with, robust moral facts – by definition – don't *cause* our moral beliefs. So, we cannot directly perceive (see, hear, smell, etc.) them. And since they – by definition – aren't reducible to or identical with scientific facts, we cannot rely on scientific theorizing to find out about them either. Further, robust moral facts cannot *depend* on our moral beliefs. After all, they are objective: they hold independently of what we think of moral matters. So, they cannot be a function of our moral beliefs. But if we cannot (even *in principle*) explain why we get robust moral facts mostly right, why think that we do? It looks like we are no longer *justified* in holding our robust moral beliefs. Moral scepticism seems to follow.

Three aspects of the epistemic principle that makes up the first premise are worth elaborating on. As already mentioned, the principle is *broad* in scope: it applies generally to robust forms of realism about the *apriori*. After all, explaining why one's beliefs are mostly true seems puzzling for any domain populated by facts with certain putative metaphysical characteristics. More specifically, explaining reliability seems problematic when the facts we attempt to form beliefs about lack causal powers and hold independently of what we think about them. In addition to robust moral facts, robust mathematical, logical, modal and perhaps even epistemic facts fit that bill. And the literature bears that out: the contemporary philosophical discussion of the principle starts

³ For the sake of argument, the evolutionary debunker allows the robust moral realist to assume that robust moral facts exist. Further, we may assume that evolutionary theory indeed explains how we form moral beliefs.

⁴ See references in fn. 2 for the reliability challenge to robust *moral* realism. In addition, the principle also raises a challenge for robust *mathematical* (e.g. Clarke-Doane 2012 & 2013 & 2015 & 2016 & 2017 & forthcoming) and *logical* (e.g. Schechter 2010 & 2013) realism.

with the Field-Benaceraff challenge to robust *mathematical* realism – and much of the most recent discussion, including Justin Clarke-Doane’s (2012, 2014, 2015, 2016, 2017, forthcoming) contribution presented in §3 below, still primarily engages with that. So, the principle isn’t restricted to the robust moral domain, but seems perfectly general.

Second, the principle is *forceful*, for it doesn’t depend on an idiosyncratic epistemological background picture. Therefore, it cannot be easily dismissed. For example, it does not presuppose the *causal* theory of justification or knowledge.⁵ Rather, it only appeals to a plausibly necessary condition on *justification*: for a belief to be justified, it cannot be impossible to explain why it might be true or reliably formed. Similarly, the challenge does not assume the truth of *externalism* about justification, even though it is cashed out in terms of reliability.⁶ Of course, doing so would beg the question against many forms of robust moral realism. After all, many robust moral realists subscribe to epistemic internalism: they take justification to be a matter of what is reflectively accessible to an epistemic subject, not of whether her cognitive faculties are reliable.⁷ However, should the explanatory demand above indeed remain unmet, even epistemic internalists need to worry. For if we cannot explain why we get robust facts about the *apriori* mostly right, why think that we do at all? In other words, failing to explain reliability provides *evidence* that our belief-forming faculties about a given domain are unreliable. And while epistemic internalists can allow for unreliably formed, yet justified beliefs, they cannot plausibly argue that beliefs *knowingly* formed in an unreliable fashion are still justified.⁸

Third and finally, the principle is *distinctive*. Most notably, it is not just a thinly veiled re-statement of a more general or radically sceptical challenge for perceptual belief (about the external world). First, it is comparably easy to explain why our perceptual beliefs are reliably formed: they are *caused* by perceptual facts. So, in the case of perceptual beliefs, the antecedent of the conditional encoding the principle turns out to be false. Further, the principle does not invoke the possibility of massive, yet undetected error.⁹ Instead, the

⁵ See Clarke-Doane 2017. In that respect, Field’s (1989) formulation of the challenge to robust mathematical realism improves on Benaceraff’s (1973).

⁶ This point is widely accepted in the literature on evolutionary debunking and reliability challenges. See, for instance, Joyce 2006: 179-182 or Enoch 2010: 423f., respectively. For more on the distinction between epistemic internalism and externalism, see Pappas 2017.

⁷ See Chapter 5 for further discussion of a prominent form of internalism in moral epistemology, namely *moral intuitionism*.

⁸ Put differently, while reliability may not be necessary for justification, evidence of unreliability is surely sufficient for defeating justification, even for epistemic internalists.

⁹ In addition to avoiding such sceptical hypotheses, the principle also doesn’t appeal to closure or evidential under-determination, as standard formulations of arguments for radical scepticism do. See, for example, Prichard 2005a.

principle grants the approximate truth and defeasible justification of beliefs about robust *apriori* facts.¹⁰ Then, it insists that, if the robust realist about the *apriori* cannot explain why her beliefs are formed reliably, that undermines the defeasible justification of those beliefs. So, the principle is *moderately*, not radically, sceptical – and therefore articulates a distinctive epistemic worry for robust realism about the *apriori*.

The second premise states that evolutionary explanations of moral belief make it impossible to explain why we are reliable with respect to morality, robustly construed. How so? As we saw in Chapter 1, evolutionary processes favor *adaptive*, not robustly true, moral beliefs. So, evolution cannot explain why we would get robust moral facts mostly right. (In that respect, moral beliefs differ significantly from perceptual beliefs: having *true* perceptual beliefs would have enhanced the fitness of our ancestors. Therefore, evolution could explain why we get perceptual facts mostly right.) That, however, deepens the mystery of moral reliability. So far, we have seen that moral beliefs can neither be directly caused by nor themselves fix robust moral facts. But it gets worse: neither did our moral beliefs evolve to be about robust moral facts.¹¹ By implication, it is even more difficult – perhaps impossible – to see why we would ever be reliable in our moral beliefs.

Admittedly, some contributors to the literature doubt that the appeal to evolutionary theory in the second premise adds much to the challenge of explaining moral reliability. According to them, the evolutionary story serves a merely heuristic or illustrative purpose: it supplies some details as to why moral beliefs aren't caused by robust moral facts. But, in principle, any other causal explanation that is indifferent to robust moral truth (e.g. cultural background, upbringing, neurochemistry) would do the trick. Therefore, the evolutionary debunking argument above, based on the principle that reliability requires explanation, is simply an instance of the more general challenge to moral reliability. As David Enoch (2010: 426) writes: '...[f]or robust [moral] realists, though, it is quite safe to assume that Street's dilemma is merely a particular instance of the general epistemological challenge'.¹² There is nothing distinctively evolutionary about it.

But I think that this assessment is too hasty: on closer inspection, the evolutionary explanation of moral belief does indeed significantly complicate the explanation of moral reliability. After all, not all causal, putatively truth-indifferent explanations are created equal. Rather, some have the *potential* to serve as or, at least, facilitate an explanation of

¹⁰ See Clarke-Doane 2017: 21 or Enoch 2010: 420.

¹¹ Of course, some robust moral realists question that assumption, developing so-called third-factor explanations of moral reliability. Chapter 5 will examine those in depth.

¹² Bedke 2014 and Klenk 2017 echo that sentiment.

reliability. And evolutionary explanations are amongst them: plausibly, an evolutionary story could be told about why we get robust logical or mathematical facts mostly right.¹³ But, importantly, we couldn't do the same with respect to our reliability about robust moral facts. So, the *unavailability* of an evolutionary explanation of moral reliability adds something to the challenge, making it even more difficult to meet. In that sense, the appeal to evolutionary theory in the second premise does play a distinctive role in motivating moral scepticism.

Jointly, premises (1) to (2) seem to entail (3): that we cannot justifiably believe robust moral facts. So, justified belief about robust moral facts seems incompatible with evolutionary theory. But, of course, that kind of moral scepticism is unacceptable to robust moral realists. As we saw in Chapter 1, their view is committed to the claim that we can justifiably believe robust moral facts under epistemically favorable conditions. Therefore, robust moral realism is in trouble.

But how compelling is the argument above really? In what follows, I shall first present a recent criticism of its key principle that reliability requires explanation, which evolution problematizes in the moral case (§3), before developing my conception of the principle in response (§4) and examining its suitability for evolutionary debunking (§5).

3. Trivializing Moral Reliability by Modalizing It

In a recent series of papers, Justin Clarke-Doane (2015, 2016, 2017, forthcoming) has argued that evolutionary considerations do not undermine an explanation of reliability in the moral case.¹⁴ According to him, once we truly understand what a good explanation of moral reliability looks like, it also dawns on us that we already have one. As a result, any evolutionary debunking argument based on the idea that evolution makes explaining moral reliability impossible does not work. In this section, my aim is to present Clarke-Doane's argument in more detail, thus setting the stage for my critical assessment of it in the remainder of the chapter.

Clarke-Doane's argument proceeds in two distinct steps. First, he points out that it is *unclear* what exactly 'explaining reliability' means, even though the evolutionary debunker demands it and says that evolution makes such explanations impossible in the moral case.

¹³ See Schechter 2013 for logic and Kitcher 2005 for mathematics.

¹⁴ In an earlier paper, Clarke-Doane (2012) defends a sensitivity-based evolutionary debunking argument both against robust moral and mathematical realism. So, the series of more recent papers discussed in this chapter prove a significant departure from his earlier work.

To fill that lacuna, Clarke-Doane offers a distinctively *modal* interpretation of what it means to ‘explain reliability’.¹⁵ In more detail, he suggests that a *good* explanation of moral reliability would show that our moral beliefs are *stable* in a certain way. More precisely, our moral beliefs would have to be both *sensitive* to the robust moral facts and *safe* from moral error. Both sensitivity and safety are technical *modal* notions: a subject’s S belief that p is sensitive if and only if, were p false, S would not believe that p.¹⁶ And a subject’s S belief that p counts as safe if and only if, were S to believe p (in close-by possible worlds), p would be true.¹⁷ So, a good explanation of moral reliability would demonstrate that our moral beliefs change in accordance with the robust moral facts and couldn’t easily be false. Conversely, if we lacked such an explanation, we would have reason to believe that our robust moral beliefs were insensitive and/or unsafe. And that, in turn, might defeat their justification.¹⁸

Second, Clarke-Doane argues that our moral beliefs are *already* modally stable – and we thus have a good explanation of why we get the robust moral facts mostly right. To begin with, our moral beliefs are *trivially* sensitive to the robust moral facts. Recall that a belief that p is sensitive just in case, were p false, we would not believe p. But if p is a basic, metaphysically necessary moral truth (e.g. killing sentient beings just for fun is morally wrong), there is simply no metaphysically possible world in which p is false (e.g. it is permissible to kill sentient beings for fun). And since we may assume that our moral belief that p is true, it must therefore also be sensitive. In short, our moral beliefs are sensitive to the robust moral facts by default: since such facts hold necessarily, they could not possibly change. So, any true moral belief must be sensitive as well.

Similarly, our moral beliefs are safe from error, given their truth and evolutionary history. To see why, recall that a belief that p is *safe* just in case, were we to believe that p (in close-by possible worlds), p would be true. But, as mentioned in §2, we may assume that our moral beliefs are actually true. Moreover, we have a perfectly plausible *evolutionary explanation* why we end up holding the moral beliefs that we do – and could not have ended up with radically different explanatorily basic moral beliefs. (For example, given our evolutionary past, we could not easily have ended up believing that pleasure was *pro tanto* bad, that suicide was morally required or that parents don’t have moral obligations to look after our children.) Putting those two claims together, we would hold that same

¹⁵ He takes his proposal to be perfectly general, covering the demand to explain mathematical and logical reliability as well. However, for ease of exposition, I shall focus on moral reliability in the main text.

¹⁶ For sensitivity as a condition on knowledge, see Nozick 1981.

¹⁷ For safety as a condition on knowledge, see Pritchard 2005b, Sosa 1999, and Williamson 2000.

¹⁸ I will have to say more about which necessary condition on defeat Clarke-Doane takes to underlie the reliability challenge below.

true moral beliefs in close-by possible worlds – which is tantamount to saying that our robust moral beliefs are safe.¹⁹ So, according to Clarke-Doane, our moral beliefs are both sensitive to the robust moral facts and safe from error. In other words, they are modally stable. But that is all a good explanation of moral reliability requires. So, our moral reliability has been explained – and our robust moral beliefs therefore remain justified.

What does Clarke-Doane take his argument to show more generally? It *trivializes* the explanatory demand raised by the principle that, if we fail to explain why we are reliable in a given domain, we cannot justifiably believe facts about that domain. More precisely, for the principle to have any undermining force, he thinks that it would need to turn on the following necessary condition on defeat: ‘If information, E, undermines all of our beliefs of a kind, D, then it does so by giving us reason to doubt that our D-beliefs are both sensitive and safe’ (2016: 31). But this condition – which we might call *Modal Defeat*²⁰ – is never met for robust moral beliefs, for we can show that they are both sensitive and safe. We thus don’t have reason to doubt their modal stability. So, reliability is easy to explain in the moral case, in spite of evolutionary considerations. By implication, any evolutionary debunking argument based on the idea that evolution makes explaining moral reliability impossible must fail.

Clarke-Doane’s attack on the principle that reliability requires explanation is thought-provoking, but I don’t think it succeeds. While I am on board with his distinctively modal interpretation of what it means to ‘explain reliability’, I think we should resist the second step of his argument, namely that our robust moral beliefs are *already* modally stable. To make visible why our robust moral beliefs are more modally precarious than Clarke-Doane admits, I argue in §4 below that robust moral beliefs can be epistemically lucky.

4. Moral Reliability, Modal Stability and Epistemic Luck

In this section, I argue – *contra* Clarke-Doane – that robust moral beliefs aren’t already modally stable, despite being necessarily true, defeasibly justified and evolved (§4.1). Then, I spell out the implications for our understanding of the principle that reliability

¹⁹ Admittedly, I doubt that all our basic – let alone most non-basic – moral beliefs can plausibly be explained in evolutionary terms. However, I shall grant Clarke-Doane that assumption for the sake of argument. My criticism follows in §4 and §5 below.

²⁰ Clarke-Doane (2015, 2016, 2017, forthcoming) calls this condition ‘Modal Security’. But that strikes me as misleading: the condition concerns defeat, not justification, knowledge, ‘modal security’ – or any other positive epistemic status for that matter. And that distinction is important, if often missed (see, for instance, Mogensen ms: Ch. 4). For criticism of ‘Modal Security’, see Woods 2016.

requires explanation (§4.2). I conclude by defending my line of argument against a potential worry (§4.3).

4.1. Epistemic Luck and Modal Instability

As we saw in §3 above, Clarke-Doane argues that our moral beliefs are *already* modally stable: they are both sensitive to the robust moral facts and safe from error. Thus, we already are in possession of a good explanation of why we get the robust moral facts mostly right. However, Clarke-Doane's line of argument strikes me as fishy. But why? Here's a hunch: it seems problematic that the modal status of the truths we form beliefs about or their evolutionary history guarantees our infallibility. After all, reliability is a property of belief-forming processes, not (true) propositions or belief contents – or, for that matter, evolutionary processes. So, the combination of metaphysics and evolutionary history should not fix epistemology in that manner.²¹

But why exactly not? One way to corroborate this hunch and start developing a response would be to show that robust moral beliefs can be *epistemically deficient*, despite being necessarily true, defeasibly justified and evolved. For if they could be epistemically deficient, evolutionary history and the metaphysics of robust moral truth would no longer fix our moral epistemology. But in what sense are such robust moral beliefs epistemically deficient? Despite being necessarily true, defeasibly justified and evolved, robust moral beliefs can be *epistemically lucky*, as illustrated by moral Gettier (1963) cases. Epistemically lucky beliefs, however, are both insensitive and unsafe – and therefore modally unstable.²² So, Clarke-Doane's argument is unsound: our robust moral beliefs aren't already modally stable. But if they don't already exhibit modal stability, their reliability hasn't been explained yet.

Why think that robust moral beliefs can be *epistemically lucky*? We should think that moral beliefs can be epistemically lucky if we can, as mentioned above, find moral Gettier (1963) cases. In such cases, our moral beliefs are true and justified, yet lucky – and therefore not knowledge. So, are there moral Gettier cases? In my view, there are, even though they are rather hard to find.²³ To begin, let's examine a *non-moral* Gettier case that illustrates more

²¹ See also Besson 2009: 5 for a similar point regarding metaphysics.

²² See Pritchard 2005b: 133-41 and Orozco 2011 for more on kinds of epistemic luck. Most of them (e.g. capacity, evidential and content luck) don't undermine knowledge and/or modal stability. The same holds for the lucky use of a reliable method, as I shall argue in §6.1. below. However, *veritic luck*, i.e. that it is a matter of luck that the agent's belief is true, plausibly precludes knowledge, as illustrated by Gettier-style cases. Hence it is the focus of this section.

²³ There are – in my estimation – two main reasons for this. First, many consider *moral testimony* to be epistemically problematic. That rules out testimony-style moral Gettier cases modeled after Besson's (2009)

generally that even beliefs about necessary, namely mathematical, facts can be epistemically lucky. Consider:

Broken Calculator: you rely on your very trustworthy pocket calculator to form a belief about the product of '11 x 12'. You type it in and your calculator gives you the correct answer, namely '132'. But, unbeknownst to you, your calculator has recently broken and generates random results.²⁴

This case seems to show that we can form true, justified, yet epistemically lucky, mathematical beliefs, even though mathematical facts hold necessarily. After all, your mathematical belief is true and necessarily so: 11 x 12 equates 132. Further, your belief is justified: up until now, using your calculator has been a highly reliable method of mathematical belief-formation. However, your mathematical belief is also *epistemically lucky*, given that you are using an unknowingly broken calculator. Had you typed in another set of multiplying factors or used the calculator at another time, you might have gotten a different, incorrect answer. So, you were lucky that you got it right – and thus, intuitively, you don't know that 11 x 12 equates 132.

The **Broken Calculator** scenario above provides us with a model for Gettier cases about necessarily true beliefs: combine the bad luck of a usually highly reliable, yet broken, method of belief-formation with the good luck of correct belief. So, can we come up with a structurally similar case for robust moral beliefs? I think we can. However, before we can construct a moral analogue to **Broken Calculator**, we need to fix our method of moral belief-formation. For the time being, let's just stipulate that a basic form of *moral intuitionism* is true: intuiting certain true moral propositions is sufficient for moral justification and knowledge.²⁵ Thus, we get:

Brain Lesion: suppose that you highly reliably intuit (and thus know) various moral truths. However, curiously enough, you have never thought about whether honesty is a virtue or not. One night, without realizing, you suffer a

Gettier cases for logic. Second, since moral facts hold as a matter of metaphysical necessity (and we quantify over metaphysically possible worlds), we struggle to construct Goldman's (1976) Fake-Barn-style Gettier cases. After all, there are no metaphysically possible worlds in which, say, torturing children is morally permissible (you are surrounded by fake barns) even though it seems impermissible (you seem to look at a real barn) and actually is (you are looking at the only real barn in miles). As a result, we have to settle for Gettier-style cases for moral belief modelled after broken calculator scenarios, as I do below.

²⁴ See Huemer 2005: 123, Miscevic 2007: 49, 56, Pritchard 2012: 10, Roland & Cogburn 2011: 4.

²⁵ For a comprehensive characterization and in-depth critical assessment of moral intuitionism, see Chapter 5.

brain lesion, which has the following effect: you find intuitively compelling *any* moral proposition you entertain. Luckily for you, though, the first moral proposition you entertain the next morning is that honesty is a virtue, which is necessarily true.²⁶

This case, as **Broken Calculator** above, seems to show that we can form true, justified, yet epistemically lucky, *moral* beliefs. And that is so even though robust *moral* facts hold as a matter of metaphysical necessity. After all, your moral belief is true and necessarily so: honesty is indeed a virtue. Further, your belief is justified: up until now, intuiting moral truths has been a highly reliable method of moral belief-formation, often resulting in robust moral knowledge. However, your moral belief is also *epistemically lucky*, given that you – without realizing – suffer from a brain lesion that makes you find intuitively compelling any moral proposition you entertain. So, if you had considered a moral falsehood (e.g. it is morally required to kill one's first born), you would have formed a corresponding false moral belief. So, you were lucky that you got it right – and thus you don't know that honesty is a virtue.

What does epistemic luck have to do with modal stability, though? In both **Broken Calculator** and **Brain Lesion**, you don't know due to epistemic luck. That lack of knowledge, however, can plausibly be explained by the lack of modal stability: your beliefs are neither sensitive to the robust mathematical or moral facts nor safe from error. To begin with, your method of belief-formation in both scenarios isn't appropriately sensitive to the facts. Had you typed in a different multiplication (e.g. $11 \times 11 = 132$) or considered another moral proposition (e.g. it is morally required that one kill one's firstborn), you would have still believed them, despite their falsity. Put more formally, it is not the case that, if p were false, you would not believe that p . So, it looks as if robust mathematical and moral beliefs could be *insensitive*.²⁷ Further, both cases also seem to show that necessarily true belief and evolution are not sufficient for safety. After all, we could easily

²⁶ A word on the choice of candidate moral proposition: nothing hinges on it. According to Huemer (2005: 102), there is a moral intuition with the content to that effect, where intuition is an apriori, basic (i.e. non-inferential) source of knowledge about necessary moral truths. However, my case can be run with any moral proposition that the moral intuitionist wants to classify as intuitive, apriori and necessary (e.g. suffering is bad). For more on moral intuitionism, see Chapter 5.

²⁷ Importantly, neither your mathematical belief in **Broken Calculator** nor your moral belief in **Brain Lesion** is vacuously or trivially sensitive (or safe), just because its content is a metaphysically necessary truth. We can see that by plausibly relativizing sensitivity (and safety) to the basis or method of belief formation (cf. Pritchard 2012: 10). Interestingly, Clarke-Doane seems to think that, in this context, method-relativization only works for safety, yet not sensitivity. This strikes me as implausible, as I discuss further in §4.3 below.

add an evolutionary story about why we hold core mathematical and moral beliefs (such as that $11 \times 12 = 132$ or that honesty is a virtue). If we did so, however, both the method in **Broken Calculator** and **Brain Lesion** would remain *unsafe*: based on it, in close-by possible worlds, you could easily form false mathematical or moral beliefs such as that 11×12 does not equate 132 or that being insensitive to others is virtuous.

In sum, Clarke-Doane is wrong: our moral beliefs are not already modally stable, despite being necessarily true, defeasibly justified and evolved. Rather, as illustrated by **Brain Lesion** above, they can be both insensitive to the robust moral facts and unsafe from error. As a result, we still lack a good explanation of why we get the robust moral facts mostly right.

4.2. Reviving the Principle That Reliability Requires Explanation

What does my line of response imply for our understanding of the principle that reliability requires explanation? To begin with, we remain committed to the distinctively modal interpretation of ‘explaining reliability’. According to that interpretation, a good explanation of moral reliability amounts to establishing that our robust moral beliefs are modally stable. However, my response developed above shows that – *pace* Clarke-Doane – the explanatory demand the principle articulates isn’t trivial. Rather, it forces robust moral realists to put forth an account for why our moral beliefs aren’t epistemically lucky. If they failed to do so, then they wouldn’t have an account of why our moral beliefs are modally stable – and thus in the market for moral knowledge. In other words, the challenge to explain moral reliability turns out to be an instance of the more general problem of epistemic luck.²⁸ But robust moral realists haven’t done anything yet to address that worry – and their moral beliefs are thus still at risk. So, the principle that reliability requires explanation is far from toothless: if we fail to explain moral reliability, we cannot *know* robust moral facts.

Reviving the principle that reliability requires explanation along the lines of epistemic luck is attractive. Remarkably, it manages to be both reasonably conservative and original. How so? First, my understanding preserves most of the distinctive features of reliability challenges, as explicated in §2. Reconceived as a puzzle about epistemic luck, the principle is still broad in scope: it applies to other forms of robust *apriori* realism. For instance, we can construct Gettier-style scenarios for robust mathematical beliefs, as illustrated by **Broken Calculator** above. Further, the principle doesn’t rely on an idiosyncratic

²⁸ See Setiya 2013: Chs. 3 & 4 for a somewhat similar claim. However, as we shall see in §6.1. below, Setiya relies on a different, non-modal notion of epistemic luck that isn’t illustrated by Gettier cases.

epistemological background picture. After all, it doesn't presuppose the causal theory of justification or knowledge (since **Brain Lesion** and similar cases don't appeal to any causal requirements) and it proves problematic for both epistemic internalists and externalists (for both camps struggle equally with finding an anti-luck condition on knowledge).²⁹ Finally, the principle doesn't give rise to a form of radical scepticism, for it allows that the beliefs in question be (necessarily) true and defeasibly justified. (However, as I shall argue in §5 below, it is questionable whether the principle, together with evolutionary considerations, leads to the *moderately* sceptical conclusion that we lack moral knowledge.) In sum, my conception of the principle that reliability requires explanation is reasonably conservative.³⁰

Second, my suggestion is also *original*: the issue of epistemic luck has been widely overlooked in contemporary moral epistemology. In fact, there is almost no discussion of epistemic luck and moral knowledge, whether in the context of explaining reliability or beyond.³¹ Further, there is very little discussion of epistemic luck in the context of *non-inferential a priori* knowledge more generally (e.g. logic, mathematics or modality).³² That is surprising, given the importance of epistemic luck to discussions in mainstream epistemology, especially scepticism and the analysis of knowledge.³³ So, my conception of the principle that reliability requires explanation doesn't just withstand Clarke-Doane's criticism and preserve features that many hold distinctive, but it also broaches an underexplored issue in the epistemology of the *apriori*.

4.3. A Worry About Counterpossibility

Before assessing whether my revived principle plausibly underlies the evolutionary debunking argument presented in §2 above, let me address a potential worry. Some might feel that my argument in §4.1 and §4.2 above ignores the elephant in the room: *the*

²⁹ For more on this so-called *Gettier problem*, see Ichikawa & Steup 2018.

³⁰ In addition to preserving much of what is commonly thought as distinctive about the reliability challenge, my conception of its underlying principle also neatly accommodates the nearly ubiquitous talk of 'coincidence' and 'luck' found in the relevant literature. See, for example, Bedke 2009 & 2014, Street 2006 & 2008, Schafer 2010 and Schechter 2010 & 2013.

³¹ Huemer (2005: Ch. 5) is the only one who – to my knowledge – gestures towards an epistemic luck conception of explaining reliability. More on him follows in Chapter 5. Meanwhile, Setiya (2013: Ch. 3 & 4) is the only one who discusses epistemically lucky moral belief in depth.

³² To my knowledge, exceptions include Huemer 2005: Ch. 5, Miscovic 2007, Roland & Cogburn 2011 and Besson 2009. Williamson 2013 and, in response, Cohen & Comesana 2013, discuss whether Gettier cases can be formalized via epistemic logic. However, importantly, their discussion features perceptual Gettier cases only.

³³ See, for instance, Greco 2003, Pritchard 2005b, Riggs 2007, Sosa 1999, Unger 1968 and Zagzebski 1994.

problem of counterpossibles.³⁴ According to David Lewis' (1973) standard semantics for counterfactuals, counterfactuals with necessarily false antecedents (e.g. 'If we can square the circle, pigs can fly') or necessarily true consequents (e.g. 'If Freddie Mercury has a moustache, triangles have three sides') are trivially true.³⁵ Jeffrey Roland & Jon Cogburn (2011) have argued that counterfactual formulations of sensitivity and safety face an instance of this problem, given necessary truths. After all, both 'if it were not true that p, S would not believe that p' and 'if S believes that p (in close-by possible worlds), p is true' are trivially satisfied because there is no world in which p is false and any close-by possible world in which S believes that p is also a world in which p is true. Now, why is that problematic for me? My argument against Clarke-Doane crucially depends on the idea that moral beliefs about necessary robust moral truths can be epistemically lucky – and thus *insensitive* and *unsafe*. But if Roland and Cogburn are correct, that is just plainly impossible, at least within the standard Lewisian semantics for counterfactuals.³⁶ So, my argument is in trouble.

In response, we can point out that a potential solution to the problem, at least for sensitivity and safety, is already implicit in **Brain Lesion** above. In particular, we can relativize these modal conditions to processes or methods of belief-formation. Doing so is compelling on independent grounds: many proponents of both sensitivity and safety already understand them as relativized to methods. Accordingly, a belief is *method-sensitive* if and only if were p false, S would not believe that p via method m.³⁷ Alternatively, a belief is *method-safe* if and only if S believes that p (in close-by possible worlds) via method m, p is true.³⁸ Importantly, robust moral beliefs in **Brain Lesion** fail both of these conditions, despite the metaphysical necessity of robust moral truth: due to the brain lesion as a method of belief-formation, one would believe that, say, courage is not a virtue despite its falsity and one would believe it in close-by possible worlds despite its falsity. So, it appears that the *problem of counterpossibles* can be addressed, at least for sensitivity and safety.³⁹

³⁴ See Brogaard & Salerno 2013.

³⁵ See also Stalnaker 1968.

³⁶ Clarke-Doane's claim that necessarily true robust moral beliefs are trivially sensitive is an expression of this problem. Curiously, he doesn't seem to consider that instance problematic. However, he oddly enough thinks that trivial satisfaction is a genuine worry for safety – that can be addressed by appeal to method-relativization I mention below (see Clarke-Doane 2017). Admittedly, I don't really know what to make of this inconsistency.

³⁷ See, for instance, Nozick 1981, Luper 1987 or Becker & Black 2012: 88.

³⁸ For method-safety, see Dunaway 2017, Pritchard 2012: 10 or Williamson 2000.

³⁹ If method-relativization should fail as a response, other options are available. First, we might – following Bedke (2014) – understand the relevant modality as conceptual or epistemic. Doing so, however, would need to avoid the overgeneralization worry developed for his proposal in §6.2 below. Second, we could understand reliability and epistemic luck *non-modally*. Alternatives include primitivism (e.g. Comesana 2005),

Where are we now? In this section, I have developed an argument to the effect that robust moral beliefs aren't already modally stable, despite being necessarily true, defeasibly justified and evolved. That, in turn, allowed me to revive the principle that reliability requires explanation along the lines of epistemic luck. The resulting principle proved attractive, simultaneously being original and preserving distinctive features commonly associated with reliability challenges. I concluded by defending my line of argument against a potential worry about counterpossibility. So far, so good. But can my understanding of the principle that reliability requires explanation plausibly serve as the key epistemic premise of the evolutionary debunking argument presented in §2?

5. Revisiting Evolutionary Debunking

In this section, I examine whether my conception of the principle that reliability requires explanation can serve as the key epistemic premise of the evolutionary debunking argument presented in §2. I argue that it cannot: on my conception, the principle doesn't plausibly combine with evolutionary considerations to yield moral scepticism. Still, it raises a separate, forceful worry about epistemic luck that robust moral realists must deal with.

Recall that we are interested in the soundness of the following evolutionary debunking argument:

- (4) If we fail to explain why we are reliable in a given domain, we cannot justifiably believe facts about that domain.
- (5) Evolutionary explanations of moral belief make it impossible to explain why we are reliable with respect to morality, robustly construed.
- (6) Therefore, we cannot justifiably believe robust moral facts.

So far, my aim has been to closely examine premise (2). In §3, I presented Clarke-Doane's criticism: that 'explaining reliability' should be construed modally and that our robust moral beliefs are already modally stable (in spite of evolutionary considerations). In §4, I argued against the latter claim: robust moral beliefs can be epistemically lucky and

conditional probability (e.g. Roush 2005), or virtue epistemology (e.g. Miscevic 2007). Any of these alternatives would allow me to argue – *contra* Clarke-Doane – that necessarily true and defeasibly justified robust moral beliefs can be epistemically deficient due to epistemic luck. Finally, we might – following Mogensen (2015: Ch. 3) – be more *quietist* about the *problem of counterpossibles*, construing it as semantic and metaphysical, not epistemological. In that case, there might be no need for solving it in our context.

therefore modally unstable. That insight informed a novel conception of (1): to explain our reliability in a given domain means showing that our beliefs about that domain aren't epistemically lucky. Now, can we plausibly plug that understanding of the principle that moral reliability requires explanation into above argument? Will it help us motivate (2) and allow us to conclude (3)?

To answer these questions affirmatively, the evolutionary debunker might pursue the following line of reasoning: evolutionary considerations show that robust moral beliefs *are* epistemically lucky and therefore modally unstable. Modally unstable beliefs, however, cannot constitute knowledge. So, evolutionary considerations show that we cannot have knowledge of robustly moral facts.⁴⁰ But why think that evolutionary considerations reveal robust moral beliefs to be epistemically lucky? In response, the evolutionary debunker might compare the realist's epistemic position to **Brain Lesion**, with evolutionary influence playing the role of the brain lesion, 'Gettierizing' all her moral beliefs. After all, our evolved capacities for moral thought make for an *unreliable* method of forming beliefs about robust moral facts. So, her moral beliefs, despite being true and justified (as we may assume), would be epistemically lucky. And since those beliefs are only true as a matter of luck, they are also modally precarious: they are insensitive to robust moral facts and unsafe from error. So, she cannot know robust moral facts due to her evolutionary past – just as the epistemic subject in **Brain Lesion** cannot know that honesty is a virtue due to her cognitive impairment.⁴¹

But I doubt that this line of reasoning survives critical scrutiny. To see why, we must first take another look at how exactly **Brain Lesion** works. In that case, you form the (necessarily) true, justified, yet epistemically lucky moral belief that honesty is a virtue. Since that belief is lucky, it doesn't – intuitively – count as moral knowledge. That lack of knowledge, in turn, can be plausibly explained by its lack of modal stability. For instance, your belief isn't safe from moral error. Or, more precisely, your moral belief in **Brain Lesion** is *method-unsafe*: if, in close-by possible worlds, you formed beliefs about relevantly similar moral propositions based on relevantly similar method, your beliefs

⁴⁰ Of course, that line of thinking doesn't support the argument above – but one about *knowledge* defeat. (1) and (3) would have to be reformulated accordingly. More about that follows in fn. 41 below.

⁴¹ Some might worry that this line of reasoning is a non-starter, for an appeal to epistemic luck doesn't really vindicate the argument above. After all, the argument concerns defeat of *justification*, while epistemic luck (as illustrated by Gettier-style cases) undermines only *knowledge*. I agree that, epistemically speaking, not being able to even justifiably believe robust moral facts is much worse than merely not being able to know them. Still, if the evolutionary debunker managed to establish that there is no knowledge of robust moral facts, that would still spell serious trouble for robust moral realists. So, this line of reasoning is well worth examining. And if it should succeed, we simply adjust the argument above to reflect that.

would be false. After all, you form moral beliefs based on a brain lesion with the following effect: you find intuitively compelling *any* moral proposition you entertain. So, you could easily have ended up falsely believing that, say, being insensitive to others is virtuous. That is how – in maximal detail – the knowledge-defeating luck in **Brain Lesion** works.

However, it is unclear whether the case of evolved moral belief works in the same manner. After all, for the scenario to work just like **Brain Lesion**, our evolved moral beliefs, despite being true and justified, would need to be *unsafe*. In other words, we would need to end up with false moral beliefs in close-by possible worlds. However, there is reason to think that we would not.⁴² After all, our evolutionary counterparts inhabiting those close-by possible worlds would share our evolutionary history to a non-trivial degree – and thus employ methods of (moral) belief-formation sufficiently similar to ours.⁴³ But if they employ sufficiently similar methods, they plausibly end up believing similar things about fundamental moral matters. In particular, it seems plausible that employing those methods would *not* result in holding radically different explanatorily basic moral beliefs. For instance, they wouldn't believe that pain was *pro tanto* good, that parents should neglect their children, or that suicide is morally required. But given that *our* explanatorily basic moral beliefs are true (as we may assume in analogy with the **Brain Lesion** case) and given that theirs are sufficiently similar, it looks like our evolutionary counterparts inhabiting close-by possible worlds wouldn't believe many moral falsehoods. But if they don't, we couldn't easily have ended up – evolutionarily speaking – with false moral beliefs. Therefore, our robust moral beliefs are *not* method-unsafe. And if they are not method-unsafe, they cannot count as epistemically lucky in the same way as those in the **Brain Lesion** scenario. So, the line of reasoning sketched out above doesn't withstand critical scrutiny.⁴⁴ Evolutionary considerations don't show that robust moral belief is epistemically lucky and therefore modally unstable. Thus, knowledge of robust moral facts survives.

In sum, once we factor in my worry just developed, the epistemic luck conception of the principle that reliability requires explanation seems unsuitable to serve as the key epistemic

⁴² For a similar worry about whether evolutionary luck implies epistemic luck, see Dunaway 2017.

⁴³ As we saw in §4.3, that is also what the most compelling, method-relative formulation of safety entails. When assessing whether a given belief is safe, we must factor in how it is formed – and keep that way of forming belief fixed across possible worlds. Further, if these evolutionary counterparts didn't employ sufficiently similar methods, they would inhabit farther-off possible worlds – which would no longer be relevant to safety.

⁴⁴ Perhaps, as I argue, evolutionary considerations don't impugn the modal stability of beliefs about robust moral facts. But even if they are reliably formed, aren't we lucky that we ended up with a reliable method of moral belief-formation? And wouldn't that kind of epistemic luck undermine knowledge as well? §6.1. examines such a line of response in detail.

premise of the evolutionary debunking argument above. After all, the principle doesn't plausibly combine with evolutionary considerations to yield moral scepticism. It is important to highlight the significance of this finding. As we saw in §2, many contributors to the literature – explicitly or implicitly – understand evolutionary debunking arguments as an instance of the challenge to explain why we get the robust moral facts mostly right. But if explaining moral reliability indeed consists in showing why our robust moral beliefs aren't epistemically lucky, as I have argued, that cannot be correct. However evolutionary debunking arguments work, they don't plausibly depend on the principle that reliability requires explanation. To realize their sceptical dreams, evolutionary debunkers must look elsewhere.

However, despite dodging this particular evolutionary debunking threat, robust moral realists aren't completely off the hook yet. On reflection, the challenge to explain moral reliability, even construed as an instance of the problem of epistemic luck, proves troublesome for some popular views in moral epistemology. Take, for instance, moral intuitionism. As seen above, moral intuitionism is, very roughly, the view that true, intuitively justified moral belief is sufficient for moral knowledge. However, as **Brain Lesion** shows, moral beliefs can be true and intuitively justified, yet epistemically lucky and thus not moral knowledge. So, intuiting certain moral truths cannot be sufficient for moral knowledge. On the face of it, moral intuitionism seems incorrect. Therefore, even though it only concerns the analysis of moral knowledge, 'explaining reliability' in the relevant sense proves challenging for some views in moral epistemology. However, a discussion of whether and, if so, how moral intuitionism can handle this worry about epistemic luck has to wait until Chapter 5.⁴⁵

Where does this leave us? In this section, I have argued that my conception of the principle that reliability requires explanation cannot plausibly serve as the key epistemic premise of the evolutionary debunking argument presented in §2. Still, it raises a separate, yet forceful worry about epistemic luck that robust moral realists must deal with. In the next and final section, I shall round out my discussion by defending my diagnosis against two recently advanced alternatives.

⁴⁵ Admittedly, that changes the *scope* of the reliability challenge somewhat: it is no longer a challenge for robust moral realism *simpliciter*, but to certain views in moral epistemology often endorsed by robust moral realists (such as moral intuitionism). However, in my view, that is fairly unproblematic: methodologically speaking, it seems plausible that the force of any epistemological challenges to robust moral realism ultimately depends on what specific *epistemological commitments* the view entails (above and beyond the epistemic principle connecting evolutionary explanation to moral scepticism and optimism about moral knowledge). So, we can expect the following: the better we understand a given epistemological challenge, the more *specific* it will be. Given that, it is somewhat unsurprising that the reliability challenge only applies to a subset of robust moral realist views.

6. Evolutionary Moral Scepticism at Last?

If I am right, the principle that reliability requires explanation holds little promise for evolutionary debunkers. But is that true? In this section, I critically assess two recently advanced characterizations of the principle that challenge my diagnosis. While the first one introduces a different kind of epistemic luck (§6.1), the second one develops the idea that *all* our moral beliefs can be insensitive to the robust moral facts (§6.2). In both cases, we would end up with moral scepticism about robust moral belief on evolutionary grounds.

6.1. Being Morally Reliable by Accident – and Moral Scepticism

Some might worry that my argument ignores a kind of epistemic luck that, when combined with evolutionary theory, does lead to moral scepticism. How so? Masahiro Yamada (2011: 81, 97, 98-102) has recently argued that the lucky or accidental use of a reliable method of belief-formation undermines knowledge. To illustrate, consider the following case: suppose that you would like to weigh yourself after a strenuous workout at the gym. To do so, you can choose between two scales: a blue one on the right and a red one on the left. You decide to flip a coin, end up choosing the blue one, weigh yourself and come to believe that your weight is 75 kg. However, unbeknownst to you, the red scale (that you did not use) was actually broken, making it lucky or accidental that you ended up using the functioning one. According to Yamada, this type of lucky or accidental use of a reliable method of belief-formation (namely the blue scale) undermines knowledge. After all, it does not seem right, intuitively, to say that you would *know* your own weight in such a situation. And, importantly, Yamada's kind of epistemic luck differs from the one found in Gettier-style scenarios such as **Brain Lesion**: you are lucky to have ended up with a reliable method, not lucky that your belief is true (despite being unreliably formed).

Now, why is this relevant to the principle that reliability requires explanation – and evolutionary debunking? Kieran Setiya (2013: Ch. 3) has recently suggested that this kind of epistemic luck animates the challenge to explain moral reliability faced by (robust) moral realists. But if it did, Setiya argues, the challenge would combine with evolutionary theory to imply scepticism about moral knowledge. To see how he gets there, note first that evolution did not select for moral reliability. So, moral realists cannot argue that we use our methods of moral belief-formation *because* they are reliable. But neither can moral realists argue that our moral belief-forming methods are reliable *because* we use them, given that moral facts hold independently of what we believe about them. So, it looks as

if it were *lucky* or *accidental* that we use the (reliable) moral belief-forming methods that we in fact use. But, as we have just seen, such epistemic luck undermines knowledge, implying scepticism about robust moral knowledge.

What lesson does this appear to hold for my argument in §5? Contrary to what I suggest, the epistemic luck conception of the principle that reliability requires explanation *can* serve as the key epistemic premise of the evolutionary debunking argument presented in §2. To see that, we simply need to acknowledge the existence of a particular kind of epistemic luck: being reliable about moral matters by accident. But once we do that and factor in our evolutionary history, we end up with a *genuine* evolutionary debunking argument.

But this line of criticism strikes me as implausible. Importantly, it is at best unclear whether we really intuitively lack knowledge in Yamada's case above. After all, it seems to me that as long as you end up using the reliable scale, whether luckily or not, your belief about your weight is knowledge.⁴⁶ To strengthen my point, consider an analogous case about testimony, put forth by Tamar Gendler and John Hawthorne (2005: 346): 'You enter a room and ask someone the time. She replies truthfully and correctly, and she is extremely reliable. But your informant happens to be surrounded by a roomful of compulsive liars.' In this case, according to them, you *know* what time it is, even though you are lucky to have used a reliable method, namely a sincere and competent instead of compulsively lying testifier.⁴⁷ So, it is at best unclear whether the lucky or accidental use of a reliable method undermines knowledge. But, of course, if that is correct, Setiya's suggestion does not get off the ground.

6.2. Being Forgetful About Robust Moral Reality – and Moral Scepticism

Others might worry that my argument overlooks ways of demonstrating the modal instability of moral beliefs that do not rely on epistemic luck and do lead to moral scepticism on evolutionary grounds. How so? Matt Bedke (2014: 18-29) has recently argued that all our moral beliefs are *insensitive* to the robust moral facts – and that this

⁴⁶ I am not alone. Schafer (2014) seems to share my intuition here. Further, in a slightly different context, Goldberg (2010: 116) argues that such luck would not undermine perceptual justification.

⁴⁷ Kvanvig (2004: 198) appeals to a similar case in which you luckily end up with an accurate history book. However, according to him, even though you would understand history, you would not *know* it, exactly because of epistemic luck. But, in my view, his case features *evidential* luck, not the lucky use of a reliable method. And it is almost universally agreed that evidential luck does not undermine knowledge. In support of my point, see Jenkins (2006: 147) regarding luckily obtained, truthful wartime diaries.

defeats their justification.⁴⁸ His argument proceeds in two distinct steps. First, he identifies a sufficient condition on defeat of justification, which he calls *obliviousness*. Your belief that p based on justification J is *oblivious* to the target fact p when, were p not the case (as a matter of *conceptual* possibility): (i.) you would still believe that p; (ii.) you would have the same justification J for believing that p; and (iii.) the same causal explanation for why you believe that p and have justification J would hold. Once you become aware that your belief that p is oblivious, its justification J has been defeated.⁴⁹ Second, he argues that robust moral beliefs are oblivious and thus lack justification. To establish that, Bedke appeals to a *sceptical hypothesis*: a scenario in which we are, without realizing, radically deceived about the nature of reality. In particular, he assumes that the basic moral facts are different than we *prima facie* justifiably believe them to be (e.g. pain is good, there are no moral facts at all). Then, he asks: what would we believe, what justification would we have, and what would explain what we believe in that case? His answer is sobering: we would have the same robust moral beliefs, justification and causal explanation (because we could hold the natural world fixed). So, our moral beliefs are *oblivious* to the robust moral facts – which defeats their justification.

Before explaining the relevance of Bedke’s argument for our dialectic in more detail, two features are worth highlighting. To start with, obliviousness is in effect the negation of *sensitivity*. Still, it differs in the following ways: obliviousness is a sufficient condition on defeat of *justification*, not a necessary condition on *knowledge*, and it asks us to quantify over *conceptually*, not *metaphysically*, possible worlds.⁵⁰ Further, Bedke claims that perceptual beliefs aren’t oblivious. For instance, take your belief that you have hands. In a relevantly different, yet close-by possible world, you wouldn’t have hands. But that means that the causal explanation of your perceptual belief wouldn’t be the same as in the actual world. This violates clause (iii.) of the definition of obliviousness. So, your perceptual beliefs aren’t oblivious. As a result, despite relying on a sceptical hypothesis, Bedke’s argument doesn’t imply radical scepticism about the external world.

⁴⁸ More precisely, Bedke targets robust *normative* beliefs more generally. In what follows, however, I shall only be interested in assessing whether his argument applies to a sub-set of robust normative beliefs, namely robust *moral* ones.

⁴⁹ There is further wrinkle: Bedke (2014) takes obliviousness to underlie or explicate his argument from coincidence. But since obliviousness, not coincidence, seems to do all the epistemological work, we can forget about that complication.

⁵⁰ Understanding modal conditions such as sensitivity as quantifying over conceptually possible worlds is one way of avoiding their trivial satisfaction. However, as I shall argue in this section, such an appeal is problematic. A more plausible way of avoiding trivial satisfaction is to understand modal conditions as relativized to methods or processes of belief-formation. See §4.3. above for more.

What implications does Bedke's argument have for the principle that reliability requires explanation – and evolutionary debunking? If he is correct, our robust moral beliefs – *contra* Clarke-Doane – aren't already modally stable. Rather, they are all insensitive to the robust moral facts. That, in turn, furnishes us with another epistemically problematic sense of 'explaining reliability': for their moral beliefs to remain justified, robust moral realists must explain why they aren't oblivious. If it turned out that we were indeed forgetful about robust moral reality, moral scepticism would ensue. Moreover, interpreting the principle in terms of obliviousness would render it suitable as the key epistemic premise of the evolutionary debunking argument presented in §2. After all, the principle lends support to a genuinely sceptical conclusion: if we cannot explain why our beliefs aren't oblivious, we cannot justifiably believe robust moral facts. And, importantly, to reach that conclusion, the principle plausibly combines with evolutionary considerations: at least part of the complete causal explanation of our robust moral beliefs, encoded in clause (iii.) of the obliviousness condition, will be that we evolved to hold them. Therefore, Bedke's argument from obliviousness seems to provide us with the materials for a *genuine* evolutionary debunking argument.

Of course, these implications spell trouble for the argument I developed above. In §4, I argued that robust moral beliefs can be epistemically lucky and therefore modally unstable. That informed the following epistemically problematic sense of 'explaining reliability': for their moral beliefs to amount to knowledge, robust moral realists must explain why they aren't lucky. But, as I showed in §5, the problem raised by epistemic luck doesn't plausibly combine with evolutionary considerations to yield moral scepticism. This makes the epistemic luck conception of the principle that reliability requires explanation ill-suited for running an evolutionary debunking argument. But if Bedke is right, my argument settles for too little. In particular, it ignores a sense in which 'explaining reliability' is suitable for evolutionary debunking. So, is Bedke right?

However, Bedke's argument faces a serious issue, having to do with its *scope*. In particular, it appears that many non-moral, non-normative beliefs are also oblivious (or insensitive across conceptually possible worlds) and therefore unjustified. That, however, means that Bedke's argument overgeneralizes, collapsing into more radical scepticism.

To see why, consider first perceptual beliefs about ordinary objects such as footballs.⁵¹ The property of being a football, understood as an ordinary perceptual object, supervenes (in virtue of metaphysical necessity) on various atomic and non-atomic properties. In that

⁵¹ See Clarke-Doane 2013, 2015, 2017 as well as Korman 2014 for this point.

way, *being a football* resembles moral properties such as *being wrong*, which also supervene (in virtue of metaphysical necessity) on non-moral properties.⁵² Now, ontological nihilists about ordinary objects argue that it is at least intelligible or conceptually possible that ordinary objects don't exist, while their subvening properties do.⁵³ So, if ontological nihilism about ordinary objects is true, it looks as if our perceptual beliefs about ordinary objects such as footballs are *oblivious*. After all, given the truth of ontological nihilism, even if ordinary objects did not exist (as a matter of conceptual possibility), we would probably still believe that they do, have the same justification for believing that they do and the same causal explanation for why we believe that they exist would hold.⁵⁴ So, Bedke's argument seems to generalize to perceptual beliefs about ordinary objects – and thus leads to more general scepticism. That, though, is unacceptable to both Bedke, as seen above, and anyone interested in the viability of the reliability challenge to robust moral realism.

Now, few think that ontological nihilism about ordinary objects is plausible. For that reason, let's consider an example that doesn't presuppose its truth, namely beliefs about *phenomenal* mental properties. It seems conceptually possible that there be zombies: exact neurophysical duplicates of human beings that lack phenomenal consciousness.⁵⁵ Further, we would not be able to identify zombies, given that they are exact duplicates of ordinary human beings. Finally, if we were zombies ourselves, we would probably not realize that we lack phenomenal consciousness. Rather, we would take ourselves to experience the world as it is.⁵⁶ Jointly, these claims seem to imply that our beliefs about our own phenomenal mental states might be *oblivious*. After all, given the conceivability of zombies, even if phenomenal mental properties did not exist (as a matter of conceptual possibility), we would probably still believe that they do, have the same justification for believing that they do and the same causal explanation for why we believe that they exist. So, Bedke's argument seems to overgeneralize yet again.

⁵² Of course, the exact modal nature of moral supervenience remains a question of active debate. See, for instance, Rosen forthcoming. However, for the purposes of my argument, let's just assume that the relevant modality is metaphysical necessity. This accords well with various forms of robust moral realism.

⁵³ See, for instance, Van Inwagen 1990.

⁵⁴ Admittedly, it remains questionable whether the same causal explanation of our perceptual beliefs and their justification holds even in cases in which there are no ordinary objects. In particular, that might depend on whether the subvening properties have sufficient causal powers. However, importantly, if Bedke thinks that is the case for subvening non-normative properties, it is unclear why it should not also extend to subvening non-perceptual (or non-phenomenal) properties (at least not without further argument).

⁵⁵ Anybody other than analytic reductionists and functionalists about the mental should be able to accept that. For a highly influential argument against physicalism based on the stronger claim that zombies are metaphysically possible, see Chalmers 2009.

⁵⁶ In that respect, we would be similar to the imprisoned color scientist Mary in Jackson's (1982) knowledge argument.

In light of this serious issue, it seems reasonable to reject Bedke's argument – and thus his interpretation of the principle that reliability requires explanation. Similarly, Setiya's argument, based on the idea that we are morally reliable by accident, should be rejected, for it proved problematic as well. As a result, my diagnosis developed in §4 and §5 still stands: the principle that reliability requires explanation should be interpreted as an instance of the problem of epistemic luck and thus holds little promise for evolutionary debunkers.

7. Conclusion

In this chapter, I defended the principle that moral reliability requires explanation against Clarke-Doane's criticism. More precisely, I argued that – *contra* Clarke-Doane – our robust moral beliefs aren't already modally stable. Rather, as moral instances of Gettier-style scenarios illustrate, such beliefs can be epistemically lucky – and therefore insensitive to the robust moral facts and unsafe from moral error. That, in turn, informed a novel, attractive conception of the principle that reliability requires explanation: it demands that robust moral realists offer an account for why their moral beliefs, despite being necessarily true and defeasibly justified, avoid epistemic luck.

However, while my defence of the reliability challenge highlighted a worry about epistemic luck widely overlooked by robust moral realists, it held little promise for evolutionary debunkers. After all, my conception of the principle that moral reliability requires explanation doesn't plausibly combine with evolutionary considerations to yield moral scepticism. Therefore, it cannot serve as the key premise of an evolutionary debunking argument.

Where does that leave us? For evolutionary debunkers, my discussion means that they must look elsewhere for an epistemic principle that licenses the inference from the commitment to the existence of robust moral facts and evolutionary explanations of moral belief to moral scepticism. Chapters 3 and 4 will be dedicated to that quest. In contrast, robust moral realists (and especially moral intuitionists) must tackle the worry about epistemic luck I raised in §4.1. above. Chapter 5 will discuss whether and, if so, how they could do that successfully.

CHAPTER 3: HIGHER-ORDER EVIDENCE OF ERROR

1. Introduction

Evolutionary debunking arguments attempt to show that the evolutionary theory, when combined with a commitment to robust moral objectivity, leads to moral scepticism: the view that we lack moral knowledge or that our moral beliefs are never justified. But which epistemic principle licenses that inference? In Chapter 2, we discussed and rejected the principle that reliability requires explanation. In this chapter, we focus on another promising suggestion: that higher-order evidence of error defeats justification. If evolution provided such evidence about belief in robustly moral facts, that principle would quickly lead to moral scepticism. Tomas Bogardus (2016), Andreas Mogensen (ms, 2017) and – at least on Katia Vavova’s (2014) compelling reconstruction – Sharon Street (2006) have all put forth versions of an argument along those lines. This chapter examines their argument closely.

In response, robust moral realists such as Vavova (2014) have objected that this evolutionary debunking argument is *self-defeating*. To see how that threat materializes, note first that the epistemic principle above characterizes *Conciliationism*, a view on the epistemic significance of higher-order evidence of error.¹ This view, in turn, is typically motivated by *Independence*, the principle that we should assess higher-order evidence of error with respect to *p* *independently* of our original first-order evidence, beliefs or reasoning in support of *p*.² To respect this principle when assessing *evolutionary* higher-order evidence of error, the robust moral realist would need to set aside *all* her moral evidence, beliefs and reasoning. But doing so risks setting aside *too much* to know whether she is genuinely mistaken about morality or not. So, the evolutionary debunking argument above seem to defeat itself.

The literature lacks any discussion of whether evolutionary debunkers can handle this self-defeat objection.³ My overall aim in this chapter is to argue that they cannot, thus filling that lacuna – and vindicating Vavova’s worry. To achieve my aim, I proceed in two steps.

¹ See Christensen 2007 and Elga 2007 for peer disagreement and Christensen 2010 for other kinds of higher-order evidence of error. Further references follow in §2.

² See Christensen 2011 or Lord 2014.

³ De Cruz et al. 2011, Kyriacou forthcoming and Sterpetti 2015 discuss a self-defeat worry for evolutionary debunking arguments. However, their discussion isn’t set within a higher-order evidence framework and fails to engage at all with Vavova’s (2014) specific objection.

First, I propose a novel, *prima facie* promising strategy for avoiding self-defeat. Then, I show that evolutionary debunkers face insuperable difficulties trying to successfully implement that strategy. As a result, the evolutionary debunking argument based on the principle that higher-order evidence of error defeats justification fails.

What does the strategy (that constitutes the first step of my argument) look like? It consists in endorsing an alternative (background) view on the epistemic significance of higher-order evidence of error. More specifically, to avoid self-defeat, evolutionary debunkers should replace Conciliationism with a view that (A) rejects *Independence* and therefore allows first-order moral evidence into the picture, yet (B) still yields the verdict that evolutionary higher-order evidence can defeat the justification of first-order moral beliefs. Thomas Kelly's (2010) *Total Evidence View* promises to fit the bill. According to it, your *total evidence* determines whether your belief that *p* is justified or not. The total evidence includes both your first-order evidence (in support of *p*) and the second-order evidence of peer disagreement (regarding *p*). Therefore, Kelly's (2010) view clearly satisfies (A): it rejects *Independence*, affording our first-order moral evidence, beliefs and reasoning a role in determining whether a given belief is justified or not.

Meanwhile, whether Kelly's (2010) view satisfies (B) is more complicated, as I show in the second step of my argument. To establish (B), evolutionary debunkers must argue that the *total evidence* available to the robust moral realist, consisting of her first-order moral evidence and the evolutionary higher-order evidence of error, *defeats* the justification of her moral beliefs. But, on reflection, evolutionary debunkers cannot discharge this argumentative burden. The exact reason for their failure depends on the *kind* of higher-order evidence of error that evolutionary considerations allegedly provide. Debunkers such as Street (2006), for whom evolutionary considerations supply evidence of moral unreliability, struggle with *evidential weight*. In contrast, debunkers such as Bogardus (2016) and Mogensen (ms, 2017), who construe evolutionary considerations as evidence of moral peer disagreement, are committed to a pair of *inconsistent* assumptions about evolutionary counterparts. Either way, evolutionary debunkers who rely on the epistemic principle that higher-order evidence of error undermines justification struggle to implement (B) of my proposed strategy. By implication, their arguments cannot avoid self-defeat.

This chapter is structured as follows. §2 presents the evolutionary debunking argument based on the principle about higher-order evidence of error in more detail, while §3 turns towards characterizing the self-defeat objection. Then, §4 unveils my two-part strategy for dealing with self-defeat. The remainder of the chapter explores whether evolutionary debunkers can make good on that strategy. §5 argues that the first part is easy to pull off,

but both §6 and §7 raise concerns about the implementation of the second part, having to do with evidential weight and whether evolutionary counterparts qualify as epistemic peers. §8 concludes.

2. Evolutionary Debunking and Higher-Order Evidence

Recently, some philosophers have argued that insights from evolutionary theory debunk beliefs in robust moral facts. To make their case, a few of these evolutionary debunkers rely on the principle that higher-order evidence of error defeats justification. Despite some differences in detail, the arguments sketched by Tomas Bogardus (2016), Andreas Mogensen (ms, 2017) and – at least on Vavova’s (2014) compelling reconstruction – Sharon Street (2006) share the following structure:

- (1) Higher-order evidence of error about your beliefs that *p*, *q*, etc. defeats their justification.⁴
- (2) Evolutionary considerations provide robust moral realists with higher-order evidence of error about their moral beliefs.
- (3) Therefore, evolutionary considerations defeat the justification of the robust moral realists’ moral beliefs.

This piece of reasoning, if sound, forces moral realists into moral scepticism: even if robust moral facts exist, we couldn’t ever form justified beliefs about them. In the remainder of this section, I shall explain in more detail what these premises mean, why they might be true – and how they jointly seem to entail the conclusion.

2.1. Higher-Order Evidence of Error and Defeat

To begin with, the first premise introduces the notion of *higher-order evidence of error*. Such evidence of error indicates that one suffers from some epistemic malfunction.⁵ Here are two familiar examples:

Offside Call: In my spare time, I enjoy attending football games with my best friend Julian. We are both equally good at spotting whether a forward is

⁴ Plausibly, the higher-order evidence of error would also need to be good, strong or weighty *enough*. Most common cases, including the ones below, intuitively meet that threshold. For more on the issue of evidential weight in more controversial cases such as evolutionary debunking, see my discussion in §6 below.

⁵ See Christensen 2010 and Lasonen-Aarnio 2014. Some higher-order evidence is evidence of epistemic success, not error (e.g. visiting your optometrist might confirm your visual reliability).

offside or not. Last Sunday, though, we disagreed: while Julian judged that our forward started from an offside position and the resulting goal was thus irregular, it seemed to me that our forward timed his run well and was onside.⁶

Hypoxia: While climbing the Dufour peak in the Swiss Alps, the weather suddenly turns near the summit. I stop briefly to calculate whether there is enough time to reach the peak and start the climb down before the snow storm hits. After going over my calculations several times, I am rather confident that I should be able to make it. However, I suddenly remember that, given the high altitude, I am very likely to suffer from mild hypoxia (or lack of oxygen), which undetectably impairs one's reasoning, leading to stupid, yet fatal mistakes.⁷

In the first case, I receive evidence of *peer disagreement*: Julian and I are equally good at making offside calls, but disagree about whether our forward was offside or not this time. Since we cannot both be right (but are equally good at making offside calls), one of us must be in error, which might very well be me. In contrast, the second case features evidence of *unreliability*: hypoxia makes it very likely that my reasoning (about time, in this case) is mistaken.⁸

The first premise doesn't just introduce higher-order evidence of error, but it also articulates a view about its epistemic significance. *Conciliationism* says that higher-order evidence of error defeats the *justification* of relevant first-order beliefs.⁹ Originally, this view was defended in the context of peer disagreement, recommending that one *conciliate* (hence its name) upon receiving evidence of peer disagreement such as in **Offside Call**.¹⁰ However, it can easily be generalized, resulting in a view that says that *any kind* of higher-order evidence of error defeats justification.¹¹ For instance, it would say that, in **Hypoxia**, learning of my likely reason-distortion defeats the justification of my belief that I have enough time to reach the Dufour Peak and return safely. So, Conciliationism is a

⁶ See Elga 2007 for a structurally similar case.

⁷ See Elga ms and Lasonen-Aarnio 2014.

⁸ Other kinds of higher-order evidence of error include following incorrect epistemic rules, following correct epistemic rules incorrectly, etc. For more, see Lasonen-Aarnio 2014: 315. Evolutionary debunkers focus on the two featured in the main text, though.

⁹ Nothing below hinges on characterizing Conciliationism in terms of *justification* defeat. Instead, it could be defined in other epistemic terms (e.g. rationality, reasonability, confidence).

¹⁰ See Christensen 2007, Elga 2007, Feldman 2006 and Matheson 2015.

¹¹ See, for example, Christensen 2010, Lasonen-Aarnio 2014 or Vavova 2018.

moderate form of *scepticism*: according to it, higher-order evidence of error defeats justification, yet only the justification of our *relevant* first-order beliefs, namely those we have such evidence about.¹²

Conciliationism is *prima facie* attractive.¹³ First, it accommodates our *intuitions*: in both **Offside Call** and **Hypoxia**, it seems *intuitively* appropriate to revise our beliefs in light of the higher-order evidence of error – which Conciliationism respects. Second, the view also plausibly *explains* our intuitions. For instance, given that Julian and I are equally like to get offside calls right and that there must be a mistake either on his or my part in **Offside Call**, we both have reason to think that we have made a mistake, which may well be enough to defeat the justification of our relevant beliefs. Similarly, given my medical condition in **Hypoxia**, it is significantly more likely that I have made a justification-defeating mistake in my time management. Third, Conciliationism is motivated by a seemingly plausible principle for correctly evaluating evidence. According to *Independence*, we should assess higher-order evidence of error with respect to *p* *independently* of our original first-order evidence, beliefs or reasoning in support of *p*.¹⁴ To see why that seems plausible, reconsider **Offside Call**: once I learn that Julian, an epistemic peer, disagrees with me, it would be intuitively *wrong* or *irrational* to dismiss his judgment by relying on my initial perceptual seeming that the forward wasn't offside. (Similarly, for **Hypoxia**: sticking with my original reasoning would be epistemically problematic in the face of significant risk of altitude-induced distortion of reasoning.) But once we accept *Independence*, Conciliationism straightforwardly follows: if it is rational to bracket one's first-order evidence and thus only the higher-order evidence of error matters, it will defeat the justification of our relevant first-order beliefs.¹⁵

2.2. Evolution and Higher-Order Evidence of Error

The second premise states that evolutionary considerations provide robust moral realists with higher-order evidence of error about their moral beliefs. To see why, it is helpful to take a slight detour and introduce a more general epistemological phenomenon first. Consider:

¹² For that very reason, the argument does not collapse into an argument for radical external world scepticism. After all, evolutionary considerations don't provide higher-order evidence of error about our perceptual beliefs. For more on that, see Vavova's 2014: §3 & §4.

¹³ See Matheson 2015.

¹⁴ See Christensen 2011 in the context of peer disagreement. For criticism, see Lord 2014. Both evolutionary debunkers and their critics endorse *Independence*. See Bogardus 2016: 656 and Vavova 2015: 12.

¹⁵ For more on that implication, see Christensen 2009: 758f.

Liberal Community: You grow up in a liberal community, acquiring corresponding political beliefs (such as that government should levy high taxes, that social security is important, etc.). Later in life, you realize that you could easily have grown up in a conservative community instead. In that case, you would hold conservative beliefs now (such as that government should minimize their interference with the free market, etc.). So, you hold liberal beliefs, in some sense, *just because* you grew up in a liberal community.¹⁶

The realization you have in this case strikes many as unsettling and epistemically problematic. In fact, becoming aware of an irrelevant influence on your beliefs seems, intuitively, to have *defeating force*: your political convictions no longer seem to enjoy the same level of *justification* once you realize that, in a sense, you only hold them because you grew up in a liberal community (and would hold contrary beliefs, had you grown up in a conservative community instead). This case is an instance of a more general epistemological phenomenon that has recently garnered increased philosophical attention, namely the *problem of irrelevant influence*.¹⁷ An irrelevant influence is a factor that influences our beliefs, but doesn't bear on their truth.¹⁸ For instance, in the case above, the irrelevant factor might be political *indoctrination*, which is an anti-reliable method of belief-acquisition, likely leading to a high proportion of false to true beliefs.¹⁹ So, even though those indoctrinating practices have influenced your political beliefs, they don't aim at truth – and therefore don't bear on the truth of your beliefs.

Any case of irrelevant influence will vary along (at least) three dimensions: the nature of the *factor* in question, the nature of its *influence* on our beliefs, and the class of our *beliefs* under influence. To begin with, many different factors can irrelevantly influence our beliefs, including genetics, gender, personality traits, cultural and socio-economic

¹⁶ For a similar case, see Cohen 2000. I am using the American sense of 'liberal' and 'conservative' here, not the European one.

¹⁷ Since the relevant literature has been emerging only recently, there isn't one universally accepted label for the phenomenon yet. For instance, while Mogensen (2017) calls it *contingency anxiety*, Ballantyne (2013) terms it the *historical variability problem*. In addition, the literature circumscribes the scope of the phenomenon in different ways. For instance, while some focus only on political beliefs (e.g. DiPaolo & Simpson 2016), others include any beliefs based on non-contingent background features such as genetics or species membership (e.g. Vavova 2018). In the rest of this chapter, I shall adopt Vavova's (2018) label, which seems most inclusive: the *problem of irrelevant influence*. For more, see Ballantyne 2013, Cohen 2000, DiPaolo & Simpson 2016, Dworkin 1996, Elga ms, Mogensen 2017, Rosen 2001, Schechter ms, Schoenfeld 2013, Sher 2001, (and especially) Vavova 2018 and White 2010.

¹⁸ See Vavova 2018: 134. For an alternative definition in terms of *arbitrary background factors*, see Mogensen 2017.

¹⁹ For more on a diagnosis along those lines, see DiPaolo & Simpson 2016.

background, class and education. In **Liberal Community** above, the irrelevant factor probably amounts to a combination of class, socio-economic background and education. Further, irrelevant factors can influence our beliefs in different ways.²⁰ Often, irrelevant factors are part of a *causal* explanation of why we hold certain beliefs. For instance, **Liberal Community** seems to be such a case: growing up in a liberal community at least partially causally explains why you have liberal political convictions. But that is not the only way irrelevant factors can influence our beliefs. In other cases, the relationship between irrelevant factors and our beliefs seems best described as *counterfactual*: we only believe that, say, Swiss folk music is beautiful because we were born and raised in Switzerland. Had we grown up in the USA instead, we would find it revolting or, at best, curious. Again, there is also a hint of that in **Liberal Community** above.²¹ After all, had you grown up in a conservative community instead, you would hold conservative beliefs about politics now. Finally, irrelevant factors don't just influence our political beliefs, as seen above. Rather, they can also influence our *moral*, *philosophical*, *aesthetic* and *religious* beliefs.²² For instance, you might believe that allowing your disabled newborn to die of hypothermia is morally wrong only because you don't live in ancient Sparta, that virtue epistemology is an implausible view of knowledge only because you didn't attend Rutgers, that opera is beautiful only because of your upper-middle class background, or that there is no God only because you didn't grow up in an Amish community.

But not all irrelevant influence on our beliefs is epistemically problematic. Consider:

TV Love: Your college crush happened to think that 'The Wire' was the best TV show ever made. In an effort to impress her, you binge-watched the series, which had hitherto been unknown to you, trying to keep up with the complexities of the narrative. While your crush was eventually short-lived, you have been in love with 'The Wire' ever since. Later in life, you realize that you could easily have had a crush on a different person in college – and therefore loved another TV show. So, you think that the 'The Wire' is great, in some sense, *just because* of your college crush.²³

²⁰ See Vavova 2018: 136.

²¹ Perhaps, irrelevant factors can even non-counterfactually relate to our beliefs. Consider: 'if it were not for x, I would not exist today'. For more, see Vavova 2018: 136.

²² For more on the scope of the problem, see especially Ballantyne 2013.

²³ This case is inspired by Mogensen 2017: 595.

In this case, your feelings towards your college crush irrelevantly influence your belief that ‘The Wire’ is a great TV show. After all, your crush partially explains why you hold this aesthetic belief, yet doesn’t bear on its truth.²⁴ However, there doesn’t seem anything epistemically fishy here: intuitively, realizing that you only love ‘The Wire’ because of your college crush doesn’t have any defeating force. So, not all irrelevant influences on our beliefs are epistemically problematic. This, though, leads to an important question: *when* exactly is irrelevant doxastic influence epistemically problematic – and, if so, *why*? According to most contributors to the literature, irrelevant influence on our political, moral, philosophical, religious or aesthetic beliefs is epistemically problematic when and because it provides us with *higher-order evidence of error*.²⁵ So, the epistemic difference between **Liberal Community** and **TV Love** comes down to this: while the former case features higher-order evidence of error, the latter does not.

What *kind* of higher-order evidence of error can becoming aware of irrelevant influences on our beliefs provide? There are at least two well-developed answers in the literature.²⁶ According to the first, as defended by Tomas Bogardus (2016) and Andreas Mogensen (ms, 2017), becoming aware of irrelevant influences provides evidence of possible *peer disagreement*. Their account plausibly explains some of the core cases of epistemically problematic irrelevant influence, including **Liberal Community** above. After all, once you realize that you would have held contrary political convictions if you had been raised in a conservative community instead, you also acknowledge the existence of substantive political disagreement between – what seem to be – epistemic peers. Further, their account also plausibly explains why there is nothing epistemically worrisome about **TV Love**: there is no (or only little) peer disagreement about the aesthetic qualities of ‘The Wire’.²⁷ So, you shouldn’t worry about the irrelevant influence of your crush on your aesthetic beliefs about a TV show that is universally critically acclaimed.

²⁴ At least assuming some form of aesthetic realism. However, even given a form of aesthetic response-dependence, *her* approval of the TV show probably doesn’t bear on the truth of *your* aesthetic belief.

²⁵ See especially Ballantyne 2013, Mogensen 2017 and Vavova 2018. For a dissenting voice, see Dworkin 1996.

²⁶ In addition to the two accounts discussed in this chapter, some have argued that irrelevant influences provide evidence that we are violating our own standards of reasoning (e.g. Elga ms, Schoenfield 2013), that our beliefs are causally arbitrary such that we could easily have ended up in an equally good epistemic position that would have supported contrary beliefs (e.g. Ballantyne 2013) or that we cannot explain moral reliability (e.g. Schechter ms). However, in what follows, I shall focus on the two most well-developed accounts found in the literature.

²⁷ For some – probably defeasible – evidence, see: <http://www.metacritic.com/tv/the-wire> (August 21, 2018).

According to the second answer, as defended by Katia Vavova (2014, 2018), becoming aware of irrelevant influences provides evidence of *unreliability*. Such evidence, as **Hypoxia** shows, tells us that our beliefs have been formed in ways that make them probably false. Again, her account plausibly diagnoses what is troublesome about **Liberal Community** (and other core cases): tight-knit political communities raise the risk of *indoctrination*, a highly unreliable (or even *anti-reliable*) method of acquiring beliefs.²⁸ Further, her account can explain why irrelevant influences are sometimes unproblematic. For instance, **TV Love** does not seem to feature any evidence of unreliability. After all, your crush might have *motivated* you to engage with ‘The Wire’ in the first place, she isn’t the *reason* why you think that the show is amazing. So, your college crush didn’t impair your ability to form true aesthetic beliefs – and you therefore need not worry about its irrelevant influence.

Now, it should become clear why evolution provides robust moral realists with higher-order evidence of error. After all, our discussion seems to establish that, if irrelevant influences are epistemically worrisome, they provide higher-order evidence of error. Evolution, though, seems to irrelevantly influence our moral beliefs in a problematic way. Therefore, evolution provides higher-order evidence of error. To see how evolutionary debunking can be construed as an instance of the *problem of irrelevant influence*, consider:

Evolutionary History: You believe that parents are morally obliged to look after their children (at least *ceteris paribus*). But you realize that, had humans evolved differently, you would hold different moral beliefs now. For instance, if looking after one’s offspring would *not* have boosted reproductive fitness, you would now believe that we have a moral obligation to kill our first-borns. So, you hold your moral beliefs, in some sense, *just because* we humans evolved in certain ways.

In this case, evolution works as an *irrelevant influence*: even though it at least partially explains why we hold our moral beliefs, it does not bear on their truth, since it concerns boosting reproductive fitness. The irrelevant evolutionary influence on our moral beliefs also seems to be epistemically problematic: for many of us, it has, intuitively, *defeating force* – and therefore, following the line of thought above, provides us with higher-order evidence of error.

²⁸ An anti-reliable method of belief-formation is likely to lead to false beliefs. In contrast, an unreliably formed beliefs are equally likely to be true or false. See DiPaolo & Simpson 2016.

But what *kind* of higher-order evidence of error does evolution provide the robust moral realist in **Evolutionary History**? Plausibly, that depends on which more general account of the problem of irrelevant influences we endorse. For Bogardus (2016) and Mogensen (ms, 2017), evolutionary considerations amount to evidence of *possible peer disagreement*: evidence that robust moral realists *could* disagree with their evolutionary counterparts about fundamental moral matters (such as the wrongness of incest or slavery, our obligations towards our children, that ethnicity doesn't matter to moral standing). After all, robust moral realists must realize that, had humans evolved differently, they would hold different moral beliefs now. Suppose robust moral realists believe that incest is morally wrong, based on a corresponding moral intuition. Their evolutionary counterparts might disagree: since incest did *not* hamper *their* reproductive fitness, they don't believe that it is morally impermissible. Rather, they believe that it is perfectly morally alright, based on *their* corresponding moral intuition. In that manner, evolutionary considerations amount to evidence of possible peer disagreement. So, for Bogardus (2016) and Mogensen (ms, 2017), robust moral realists find themselves in a situation similar to **Offside Call**.

In contrast, according to Sharon Street (2006), robust moral realists face a scenario similar to **Hypoxia**.²⁹ For her, evolutionary considerations provide robust moral realists with evidence of moral *unreliability*.³⁰ After all, when robust moral realists reflect on the evolutionary origin of our moral beliefs, they must realize that evolution selects for *adaptive*, not true, moral beliefs.³¹ For instance, suppose you believe that you have special moral obligations to your family, based on a corresponding moral intuition. But then you realize that we evolved to survive, not to track robustly moral truths, and that it is therefore likely that your belief is false. So, moral realists have good reason to think that their moral beliefs have been *unreliably* formed: they are the upshot of a process that was *not* designed to get at robust moral truth. In that way, evolutionary considerations provide moral realists with evidence of moral unreliability.

With both premises in place, we are now able to secure the sceptical conclusion of the argument above. To do that, we first take the view that higher-order evidence of error defeats justification. Then, we add the claim that that evolutionary considerations provide such evidence about moral beliefs, understood to be about robust moral facts. We end up

²⁹ At least in some moods. In other moods, Street's (2006, 2008) argument is more plausibly reconstructed along the lines of a reliability challenge to (robust) moral realism. See Chapter 2 for more on that.

³⁰ See also Vavova 2014. However, she ultimately thinks that robust moral realists cannot recognize evolutionary evidence of moral unreliability as such, as I shall explain in §3 below.

³¹ See Fraser 2014 for empirical details.

with the conclusion that evolutionary considerations defeat the justification of beliefs in robust moral facts. Therefore, evolutionary considerations seem to saddle robust moral realists with an uncomfortable moral scepticism. Even if robust moral facts exist, we do not and cannot form justified beliefs about them.

3. The Threat of Self-Defeat

One of the most powerful objections to the evolutionary debunking argument above is that it threatens to be *self-defeating*. To see how that threat materializes, recall that Conciliationism is typically motivated by *Independence*, the principle that we should assess higher-order evidence of error with respect to *p* *independently* of our original first-order evidence, beliefs or reasoning in support of *p*. To respect this principle when assessing *evolutionary* higher-order evidence of error, the robust moral realist would need to set aside *all* her moral evidence, beliefs and reasoning. Why? Because the debunker's idea is that evolutionary considerations call into question *all* her robustly moral beliefs. But doing so risks setting aside too much to know whether she is mistaken about morality or not. As Katia Vavova (2014: 89-93), who develops this objection most clearly and forcefully, writes: 'we cannot determine if we are likely to be mistaken about morality if we can make no assumptions at all about what morality is like' (Vavova 2014: 92).³² After all, to see whether true robustly moral and adaptive moral beliefs do indeed come apart, as the evolutionary debunker has it, we need to know something about the contents of both of those sets. If we don't know what robust morality is, how can we know whether we fall short of it? Or, if robust morality could be about anything, we have no reason to think that mind-independent moral truths and adaptive moral beliefs don't coincide or overlap. So, the evolutionary debunking argument above seems to defeat itself.

For illustration, consider an analogy with perception: to evaluate whether my perceptual beliefs about mid-sized objects in my immediate environment (e.g. tables, chairs, desk lamps, water bottles, coffee mugs) are indeed unreliably formed, I need to make some assumptions about the contents of my perception. For instance, I need to know very roughly what a chair looks like to make sure that my perceptual belief that there is a chair right in front of me is false. Similarly, Vavova (2014) points out: 'I cannot show that I am not hopeless at understanding right and wrong without being allowed to make some

³² Vavova frames her objection slightly differently: robust moral realists cannot *recognize* evolutionary evidence as good evidence of error – and that there are thus limits on our ability to get evidence of our own error, arising from the way such evidence works. However, her key move is denying that robust moral realists have a 'good *independent* reason' (92-96) to doubt their moral beliefs. So, it can be framed as an attack on *Independence* in the context of evolutionary debunking, as I do here. Thanks to Neil Sinclair for helpful discussion of this point.

assumptions about what is right and wrong' (ibid.). For instance, consider my moral belief that racism is wrong. At the same time, I am aware of evolutionary explanations of racism: it is adaptive to be suspicious of those who look different to me. Here, the adaptive and true robustly moral beliefs come apart. However, importantly, to draw that distinction, I must already assume some robustly moral truths, including that racism is morally wrong.³³

It is important to appreciate how *general* this objection is. More precisely, the worry does not depend on a *specific* interpretation of premise (2), the claim that evolutionary considerations provide higher-order evidence of error. Sure, Vavova (2014) criticizes a version of the evolutionary debunking argument according to which thinking about the evolutionary origins of their moral beliefs provides robust moral realists with evidence of unreliability. But that, in my view, artificially restricts the scope of the objection, since we can easily *generalize* it to *any* evolutionary debunking argument from higher-order evidence of error, regardless of the *kind* of evolutionary higher-order evidence. After all, any such argument subscribes to the problematic commitments that the objection capitalizes on: evolution as providing higher-order evidence of error, Conciliationism and, especially, *Independence*. For instance, suppose that evolutionary considerations amount to evidence of peer disagreement, as Bogardus (2016) and Mogensen (ms, 2017) argue. Once the robust moral realist receives such evidence, she must once again set aside all her moral evidence, beliefs and reasoning to respect *Independence*. But doing so would make it impossible for her to assess whether she, as opposed to her evolutionary counterpart, is more likely mistaken about morality. Again, the evolutionary debunker ends up with *self-defeat*. So, the threat is perfectly general – and therefore relevant to any evolutionary debunker relying on the notion of higher-order evidence of error.³⁴

³³ Vavova also addresses some preliminary responses to her objection. For instance, what if the evolutionary debunker extends the scope of their attack to include *evaluative* in addition to moral beliefs? In that case, Vavova (2014: 87-89) argues, the evolutionary debunking argument defeats itself as well: now, to respect *Independence*, the moral realist cannot even rely on her beliefs about epistemic principles, including principles about how to evaluate evidence such as *Independence* or *Conciliationism*. What if the evolutionary debunker *restricts* the scope of their attack to deontological moral beliefs? In that case, Vavova (2014: 93-95) argues, their argument either collapses into a more ambitious form (such as the one against robust moral realism) – or the evolutionary story turns out to be idle, given other worries about deontology. Finally, what if the evolutionary debunker insists that her conclusion is merely *dialectical* (such that it establishes moral constructivism, say), not sceptical? According to Vavova (2014: 89), that wouldn't get them off the hook: even if the conclusion is dialectical, the inference of the argument must still go through. But it doesn't, if her objection is correct and the argument is self-defeating.

³⁴ The self-defeat objection might be even *more* general still, arising for analogue arguments against robust construals of other putatively *apriori* beliefs about mathematics, logic, modality, epistemology or religion. Examining these analogue arguments, however, goes well beyond the scope of this chapter.

4. Avoiding Self-Defeat: A Strategy

As mentioned in the introduction, the literature doesn't feature any discussion of how evolutionary debunkers could or should respond to this self-defeat objection. My aim in what follows is to address that shortcoming. I shall propose and subsequently evaluate a two-pronged strategy. To avoid self-defeat, evolutionary debunkers should replace Conciliationism (or premise (1) of their argument) with a view that (A) rejects *Independence* and therefore allows first-order moral evidence into the picture, yet (B) still yields the verdict that evolutionary higher-order evidence defeats the justification of first-order moral beliefs.³⁵ In this section, I shall briefly motivate both prongs, while the next section introduces Thomas Kelly's (2010) *Total Evidence View* as a candidate framework that clearly satisfies (A). The final two sections will then focus on whether evolutionary debunkers could make good on (B) within Kelly's framework, discussing evolutionary evidence of moral unreliability (§6) before delving into evolutionary evidence of moral peer disagreement (§7).

To wit, three claims make up the evolutionary debunking argument: Conciliationism about higher-order evidence of error (including *Independence*), the claim that evolutionary considerations provide the robust moral realist with such evidence, and the sceptical conclusion. To avoid self-defeat, evolutionary debunkers must give up one of them. They cannot give up the last two. Without the distinctive claim about the evidential import of evolution, their argument would cease to be an *evolutionary* one. And without the sceptical conclusion, their argument wouldn't be a *debunking* one. But Conciliationism and especially *Independence* appear to be the culprits: only if the robust moral realist is forced to assess the evolutionary evidence *independently* of all her first-order moral evidence does the argument threaten to defeat itself. Therefore, evolutionary debunkers should reject Conciliationism due to its commitment to *Independence*, while holding on to the claim that evolution provides higher-order evidence of error as well as the sceptical conclusion.

But giving up on Conciliationism (or premise (1)) won't be enough, of course. Rather, evolutionary debunkers also need a *replacement*, a view that licenses the inference from the claim that evolutionary considerations amount to higher-order evidence of error to

³⁵ Why not reformulate *Independence* instead of rejecting it? Because the most plausible ways of doing so strike me as highly problematic. For instance, evolutionary debunkers might first distinguish between *substantive* and *formal* assumptions about moral truth (e.g. Gert & Gert 2016, Sinclair forthcoming: 19-22) – and then argue that *Independence* tells us to set aside only substantive, yet not formal, moral assumptions when assessing higher-order evidence of error. However, it remains unclear whether formal assumptions provide sufficient detail to assess whether moral realists are mistaken about morality. After all, those assumptions only describe the form of robust morality, not what robust morality actually looks like. We shall discuss that issue in more depth in Chapter 5.

the sceptical conclusion. More precisely, evolutionary debunkers must defend (or, at least, sketch) a view with two distinctive features: it must (A) reject *Independence* and thus allow first-order (moral) evidence into the picture, yet (B) still yield the verdict that (evolutionary) higher-order evidence defeats the justification of first-order (moral) beliefs. To successfully deal with the self-defeat objection, the view appealed to in premise (1) must satisfy both requirements. Are such views available?

5. The Total Evidence View

Fortunately for the evolutionary debunker, there are views on *peer disagreement* that promise to fit the bill. For instance, take the *Total Evidence View*, as developed by Thomas Kelly (2010: 33-44).³⁶ This view states that your *total evidence* determines whether your belief that *p* is justified or not. The total evidence includes both your first-order evidence (in support of *p*) and the second-order evidence of peer disagreement (regarding *p*). Sometimes, the total evidence justifies your original belief and thus makes it reasonable for you to stick to your guns. Suppose that you – just like me above – disagree with Julian about a football matter. While you insist that there was no delay of game, Julian claims that there was – and that, rather absurdly, the goalkeeper, who isn't on a yellow card, should be immediately sent off for it.³⁷ Here, the total evidence, consisting of your perceptual experience and knowledge of the football rule book, justifies your belief that the goalkeeper shouldn't be sent off for delay of game. But on other occasions, the total evidence may defeat the justification of your original belief – and therefore make it reasonable for you to change your outlook. Suppose that you are a sceptic about other minds, based on the relevant class of arguments. But then you find out that an overwhelming majority of professional philosophers disagree with you, having independently arrived at their view.³⁸ Here, the total evidence, consisting of the arguments, your considered judgment as well as all the considered judgments of your

³⁶ Another example would be Jennifer Lackey's (2008) *Justificationism*. However, those two views ultimately converge, according to Matheson (2015). So, I shall focus on the more widely discussed presentation. For the purposes of my argument in this chapter, I shall assume that Kelly's view is *generally* plausible. For some common objections, though, see Kelly 2010: 42-64.

³⁷ In football, delay of game is, at best, worthy of a yellow card. Players get awarded yellow cards for bad fouls that don't warrant immediate ejection. Two yellow cards, though, equal a red card, which signifies immediate rejection. This case is modelled after so-called *Extreme Restaurant Cases*. See Christensen 2007: 199-203, Elga 2007: 490f, and Kelly 2010: 41f.

³⁸ See Kelly 2010: 37. To what extent the assumption of independent convergence holds is, of course, a psychological and sociological, not philosophical, matter. But for more on its epistemic importance, see Kelly 2010: 37-41.

peers, seems to justify the common-sense view that there are indeed other minds. So, it would be reasonable for you to conciliate – or even accept the common-sense view.³⁹

The Total Evidence View clearly satisfies the first requirement or (A) above. Unlike Conciliationism, this view rejects *Independence*: it affords our first-order evidence, beliefs and reasoning a role in determining whether a given belief is justified or not. In our context, that means that the robust moral realist no longer needs to set aside *all* her first-order moral evidence, beliefs and reasoning when assessing evolutionary higher-order evidence of error. And since she is not forced to do that, she does not risk setting aside *too much* to know whether she is mistaken or not. Therefore, the Total Evidence looks like a suitable candidate for replacing premise (1) of the debunkers' argument.

But what about the second requirement or (B): does the view yield the verdict that evolutionary higher-order evidence of error defeats the justification of first-order robustly moral beliefs? Here, matters get more complicated. To establish that, evolutionary debunkers would need to argue that the *total evidence* available to the robust moral realist, consisting of her first-order moral evidence and the evolutionary higher-order evidence of error, does indeed *defeat* the justification of her moral beliefs. Can they do so?

In what follows, I shall answer that question negatively. On reflection, the total evidence available to the robust moral realist does *not* defeat the justification of her moral beliefs. By implication, evolutionary debunkers cannot satisfy the second requirement or (B) of my strategy above. And since they cannot do that, their argument defeats itself. However, as I shall argue, the exact reason for why evolutionary debunkers fail to establish defeat depends on the *kind* of higher-order evidence that evolutionary considerations allegedly provide. As I shall discuss in §6, debunkers such as Street (2006), for whom evolutionary considerations supply evidence of moral unreliability, struggle with evidential weight. In contrast, I make the case in §7 that the evolutionary disagreement argument developed by Bogardus (2016) and Mogensen (ms, 2017) rests on a pair of inconsistent assumptions.

6. Evolution and Moral Unreliability

Suppose Street (2006) is right: evolutionary considerations provide evidence of moral unreliability. Then, the total evidence available to the robust moral realist consists of both that evidence and our first-order moral evidence. Does *that* evidence defeat the

³⁹ Following Kelly (2010), my presentation of the view focuses on peer disagreement here. But, at least initially, the view seems to plausibly generalize to other kinds of higher-order evidence of error. For more on whether that impression withstands scrutiny, see §6 below.

justification of our moral beliefs? The answer seems to depend on the *weight* of the respective bodies of evidence. If the evolutionary evidence of moral unreliability *outweighed* our first-order moral evidence, our robustly moral beliefs would no longer be justified. If it didn't, our robustly moral beliefs would remain undefeated. But how should we decide which antecedent of those two conditionals holds?

To start with, we might appeal to brute intuition. After all, we seem to have widely shared intuitions about whether the first- or higher-order evidence has greater weight in some extreme cases. For instance, it is intuitive to think that the higher-order evidence of error *outweighs* our first-order evidence in **Hypoxia**. Conversely, it is natural to think that our first-order perceptual evidence *outweighs* the higher-order evidence of error gained by finding out that a very popular headache remedy caused hallucinations in 1 in 1 million test subjects. And there is some support for this line of thinking in the literature on peer disagreement. For instance, Kelly (2010) seems to advocate a form of epistemic particularism about the matter. When considering the question of whether the first- or higher-order evidence plays a greater role in fixing the reasonability of what to believe, he writes that '...the question of which counts for *more* – peer opinion, or the evidence on which the peers base their opinion? – is not, I think, a good question when it is posed at such a high level of abstraction' (34). Rather, we have to examine cases and our intuitive verdicts about them individually. Similarly, Errol Lord (2014) sketches a test for evidential weight in the context of peer disagreement that seems driven by brute intuition. To determine whether one's original reasons are strong or *weighty* enough to ground a permission to dismiss peer disagreement, we should ask: 'do [those original reasons] put you in a position to think your peer is *crazy* or otherwise epistemically *suspect*?' (Lord 2014: 376, fn. 15; emphasis mine). And perhaps, his proposal can be *generalized* to evidence of unreliability: if your first-order evidence makes the *source* of the evidence of unreliability seem epistemically suspect, the former outweighs the latter.

But I don't think that an appeal to brute intuition will help the debunker. Unlike our widely shared intuitions about evidential weight in extreme cases, our intuitions about the weight of evolutionary evidence of moral unreliability vis-à-vis our first-order moral evidence strike me much more moot. To see that, it suffices to point to the persistent disagreement in the literature about the epistemic import of evolutionary biology for robustly moral belief – which seems at least partially fueled by conflicting intuitions about evidential weight. While evolutionary debunkers share the intuition that evolutionary evidence is weightier than moral evidence and therefore undermines robustly moral beliefs, robustly moral realists tend to lack the intuition – or explain it away as irrelevant. If that is correct, we have reached a dialectical stalemate, without making any progress in the matter at hand. So, an appeal to brute intuition won't help evolutionary debunkers

such as Street (2006) to establish the claim that evolutionary evidence of moral unreliability outweighs our first-order moral evidence.

Instead, the evolutionary debunker might look towards a more *theoretical* (or *formal*) notion of evidential weight. Unfortunately, there is remarkably little literature on how to measure and compare the weight of evidence – and no literature at all on how to apply such ideas to evolutionary debunking. Still, there are some suggestions worth exploring. For instance, take James Joyce (2005: 162-5): for him, the *balance* of total evidence favors whatever ordered sequence of propositions contains an estimated higher number of truths.⁴⁰ Following his proposal, we estimate whether the first-order moral evidence or the evolutionary evidence of moral unreliability contains a higher number of truths. If we can reasonably expect the first-order moral evidence to feature more truths, the balance of total evidence favors the moral propositions making up that evidence. That would be bad news for evolutionary debunkers. Conversely, if we can reasonably expect the higher-order evidence of error to feature more truths, the balance of total evidence favors the propositions making up the higher-order evidence of error. That would be good news for debunkers such as Street.

However, I doubt that evolutionary debunkers like Street (2006) can avail themselves readily of the resources that Joyce's (2005) framework offers. First, we might worry generally that this probabilistic notion of evidential weight cannot successfully model interactions between bodies of evidence at different levels. After all, when weighing evidence from different levels, the estimates won't be independent. Rather, the estimate of how many truths the first-order evidence contains will depend to some extent on our estimate of how many truths the higher-order evidence (of error) contains. That might complicate the formation of reasonable expectations. Further, that dependency might be especially pertinent in the context of evolutionary debunking. As Kevin Brosnan (2011: 55) points out, it is impossible to estimate of how likely it is that moral truths obtain *prior* to evolutionary influence. After all, (almost) everybody accepts that our moral beliefs *evolved* – and that we thus cannot assess their truth prior to or independently of evolution. But if that is correct, how can we estimate how many truths the first-order moral evidence contains? Second, the specific application of Joyce's account to evolutionary debunking might prove problematic also in another way. In particular, we might find it hard to

⁴⁰ J. Joyce (2005) distinguishes between the *balance* and *weight* of evidence. While the *balance* of evidence concerns whether a given body of evidence 'points' towards one set of propositions over another, its *weight* corresponds to the *size* of a body of evidence. Given his usage, an evolutionary debunker therefore requires a theoretical notion of evidential *balance*, not weight. However, since the distinction isn't relevant in our discussion, I shall continue using them interchangeably.

estimate the number of basic moral truths. After all, there is plenty of disagreement about what they are and which one of them are properly basic, even amongst robust moral realists. So, how are we supposed to count the number of basic robustly moral truths? Given these significant issues, evolutionary debunkers most likely cannot borrow Joyce's way of measuring the balance of total evidence. But if they cannot do that, they once again won't be able to make good on the claim that the evolutionary evidence of moral unreliability outweighs our first-order moral evidence.

In short, debunkers who construe evolutionary considerations as evidence of moral unreliability struggle at the first hurdle. To meet the second requirement or (B) of the strategy outlined in §4 and therefore avoid self-defeat, they must establish that the evolutionary evidence of moral unreliability *outweighs* our first-order moral evidence. But doing that, in turn, requires a plausible theoretical notion of evidential weight. Even though it might not be impossible to find such a notion, our discussion (and the dearth of literature on the subject) suggests that these evolutionary debunkers have their work cut out for themselves.⁴¹ It seems fair to conclude that any evolutionary debunking argument based on evidence of moral unreliability must defeat itself.

7. Evolution and Moral Disagreement

Suppose Bogardus (2016) and Mogensen (ms, 2017) are right: evolutionary considerations amount to evidence of possible moral peer disagreement. Does the total evidence available to the robust moral realist in *that* case defeat the justification of our robustly moral beliefs?

To answer that question affirmatively, evolutionary debunkers may proceed in two steps. First, they remind us of Kelly's (2010: 42-44) diagnosis of **Offside Call**. There, Kelly argues, the total evidence available to me defeats the justification of my belief that the scorer was onside.⁴² Why that? Initially, Julian and I have different first-order perceptual evidence: while it appears to me that the forward was onside when he started his run, the opposite seems to be the case to Julian. This evidence justifies our initial perceptual beliefs, respectively. But once we become aware of our perceptual disagreement, we *pool* our first-

⁴¹ It might be unsurprising that Kelly's (2010) view requires supplementation when generalizing it to higher-order evidence of error other than peer disagreement (for which it was conceived and developed). In fact, it might even be unsurprising that *any decent* epistemic theory requires a notion of weight. (For an argument that any decent *moral* theory requires weighted notions such as normative reasons, see Lord & Maguire 2016: 1-8.) What *is* surprising, however, is that supplementing Kelly's view proves *that* hard.

⁴² Of course, the total evidence also defeats the justification of Julian's belief that the scorer was offside, as we shall see shortly.

order evidence and add the higher-order evidence of peer disagreement. The resulting *total* evidence neither supports my belief that the goal was scored onside – nor Julian’s belief to the contrary. Instead, we have a situation of *evidential symmetry* between myself and Julian. Therefore, neither his nor my initial belief is justified anymore. That is how the *Total Evidence View* diagnoses cases with the structure of **Offside Call**.

In a second step, evolutionary debunkers can argue that possible moral peer disagreements with your evolutionary counterparts share the structure of **Offside Call**. Initially, we also have two different bodies of first-order evidence: while you have the moral intuition that you have special obligations to your family members, your evolutionary counterpart has the contrary moral intuition. Again, this evidence justifies our respective initial moral beliefs. But once you discover the moral peer disagreement and pool your first-order moral evidence, the *total evidence* – just as above – supports neither your belief that you have special moral obligations to your family members nor your evolutionary counterpart’s belief to the contrary. Therefore, the justification of both your and your counterpart’s initial moral beliefs have been defeated. And, of course, the same reasoning could be employed for any other basic moral intuition that seems *prima facie* compelling and could be subject to disagreement based on divergent evolutionary histories.⁴³

By following these two steps, evolutionary debunkers who construe evolutionary considerations as evidence of disagreement could argue that the total evidence available to the robust moral realist defeats the justification of her moral beliefs. Achieving that – without even requiring a theoretical notion of evidential weight – would be no mean feat. Rather, it would show that they make good on the second requirement or (B) of my strategy outlined in §4: that the Total Evidence View yields the verdict that evolutionary higher-order evidence defeats the justification of first-order robustly moral beliefs. As a result, these evolutionary debunkers would avoid self-defeat.

To make this move work, however, evolutionary considerations must indeed amount to evidence of possible moral peer disagreement. (We cannot merely suppose that they do, as we have done up to this point.) But is that plausible at all? On reflection, I don’t think that it is: evolutionary considerations cannot amount to evidence of possible moral peer disagreement. But before developing an argument in support of my assessment, let me briefly address a red herring about whether merely possible disagreements have defeating force.

⁴³ For a similar line of thought, see Setiya 2013: Ch. 1.

Whether robust moral realists or not, we don't *actually* disagree with anyone with a different evolutionary background. Rather, the disagreement is *merely possible*. Such disagreement, some might argue, isn't epistemically significant, though. After all, we could – in principle – disagree about almost anything: that Edinburgh is the capital of Scotland, that I have hands, that $2 + 2 = 4$, and so forth. Importantly, we don't take the mere possibility of such disagreement as a reason to revise our corresponding beliefs. For instance, even if there could be someone who denies that $2 + 2 = 4$, that alone doesn't make it rational for me to stop believing in a simple arithmetic truth. So, if merely possible disagreements don't have defeating force, why does moral peer disagreement with imaginary evolutionary counterparts?

In response, Mogensen (ms, 2017) distinguishes between merely possible and arbitrarily absent disagreements. While the former lack epistemic significance, the latter still matter. How so? Importantly, even though we do not, we *could* easily disagree with an evolutionary counterpart about fundamental moral matters (such as the wrongness of incest, our obligations towards family members, and so forth). Such disagreement is just *arbitrarily absent*: its absence has nothing to do with the truth of the moral beliefs disagreed about. Therefore, it is still epistemically significant.⁴⁴ In contrast, the truth of beliefs about simple arithmetic explains the absence of disagreements about whether $2 + 2 = 4$. So, disagreements about simple arithmetic are not arbitrarily absent, but merely possible – and thus not epistemically significant. By implication, the worry that merely possible disagreements lack defeating force is a red herring, at least in our context.

However, as I shall argue now, evolutionary debunkers who construe evolutionary considerations as evidence of moral peer disagreement face a much more serious worry. For their argument to work, it must be plausible that evolutionary considerations amount to such evidence. In more detail, our evolutionary counterparts must not just disagree with us about fundamental moral matters, but also count as our epistemic and – in a sense to be explained presently – metaphysical peers. However, I doubt that entities can simultaneously satisfy both of those criteria. But if they cannot, evolutionary considerations cannot plausibly amount to evidence of moral peer disagreement. That, in turn, means that the evolutionary debunking argument from moral peer disagreement doesn't get off the ground.

⁴⁴ Mogensen (2017) calls this the *Arbitrary Absence Thesis*: 'Necessarily for any $S1, S2, p$: If $S1$ believes p and knows that $S2$ *would* believe some contrary of p if not for some condition C , which is arbitrary with respect to $S1$'s belief that p , $S1$ should be as confident of p as $S1$ ought to be of p if, all else being equal, $S1$ knew that $S2$ *does* believe some contrary of p .' (Mogensen 2017: 598; emphasis mine).

To see how my worry arises, note that any evolutionary debunking argument based on moral peer disagreement rests on two crucial assumptions. First, it assumes that our counterparts have an alternative evolutionary history – and that this explains why they hold radically different moral beliefs or disagree with us about fundamental moral matters. To illustrate, consider an example introduced in §2 above. Suppose robust moral realists believe that incest is morally wrong and that their evolutionary counterparts morally disagree. What explains their disagreement is the difference in evolutionary past: while incest hampered our ancestors' reproductive fitness, it did – by stipulation – not do so for our evolutionary counterparts.⁴⁵ As a result, they don't believe that incest is morally impermissible. Importantly, the assumption that differences in evolutionary trajectory explain differences in fundamental moral outlook is indispensable to the argument above: without it, the argument would cease to be distinctively evolutionary. Instead, it would just be a generic argument from moral peer disagreement against robust moral realism.

Second, any evolutionary debunking argument based on moral peer disagreement presupposes that our evolutionary counterparts count as our *peers*. More precisely, those counterparts must be comparable to us both *epistemically* and *metaphysically*. Epistemically, they must be equally well placed to us: their evidence must be similarly strong and their intellectual abilities comparable.⁴⁶ If their evidence was lacking or significantly impoverished and/or their reasoning capacities were impaired, we couldn't consider them our epistemic equals. That we are epistemic peers is important. For if our evolutionary counterparts weren't epistemic peers, their disagreement with us about fundamental moral matters wouldn't defeat the justification of our robustly moral beliefs. Metaphysically, those counterparts need to partake in our robustly moral reality: the moral facts that hold in their world must significantly overlap with those holding in the actual world. Or, more poetically: we both must be bound by most of the same moral laws (and seeking to uncover them). Why that? Because we couldn't have meaningful disagreement otherwise, let alone disagreement between epistemic peers. Rather, we would be talking past each other – just like two people 'disagreeing' over which ice cream flavor tastes best.⁴⁷

⁴⁵ We can imagine and spell out the empirical details in various ways: perhaps those evolutionary counterparts reproduce very differently, or incest strengthens group cohesion which, in turn, increases survival prospects, etc. For the purpose of my worry, those details don't matter.

⁴⁶ See Kelly 2005: 10. For criticism, see King 2012.

⁴⁷ Of course, this assumes – plausibly in my mind – a roughly subjectivist or response-dependent account of taste judgments. Further, note that most paradigmatic cases of peer disagreement discussed in the literature trivially meet the criterion of metaphysical peerhood – and therefore don't spell it out. Take **Offside** above: it is uncontroversial that there is a mind-independent (yet perhaps not practice/institution-independent) fact of the matter about whether the striker was offside or not.

So, the argument must assume that our evolutionary counterparts are peers, both epistemically and metaphysically.

But, on reflection, those two assumptions are inconsistent – or, at least, stand in serious tension. Suppose that the first assumption is true: our counterparts radically disagree with us about moral matters because of their radically different evolutionary history. That seems to cast doubt on their putative status as both epistemic and metaphysical peers. To begin with, if their evolved moral beliefs are radically different to ours, yet still rationally formed, they probably rest on very different bodies of moral evidence, including moral intuitions and morally relevant non-moral facts.⁴⁸ But if their evidence so significantly differs from ours that it becomes *unintelligible* to us and their moral outlook strikes us as completely alien, it seems difficult to count them as our epistemic peers.⁴⁹ So, our evolutionary counterparts don't seem equally well placed to us, epistemically speaking.

Further, the first assumption also undermines their putative status as metaphysical peers. Suppose that our counterparts hold radically different moral beliefs due to alternative evolutionary pressures. If their moral beliefs are radically different, however, it seems plausible that other, morally relevant, *non-doxastic* features of their psychology would differ from ours as well. For instance, they might not be able to experience pain (or only certain kinds of pain) or they might lack the emotion of romantic love. But if our evolutionary counterparts experience and navigate the (moral) world so differently, why think that the same robust moral facts hold for them as for us? After all, many moral facts depend on the morally relevant, non-doxastic features of our psychologies (as even robust moral realists would admit).⁵⁰ For instance, take the fact that needlessly inflicting pain on others is morally wrong. That fact only holds if there are subjects capable of experiencing pain. If our evolutionary counterparts couldn't experience pain, they would not be bound by that moral fact. Similarly, if they were incapable of experiencing romantic love, many moral laws that specifically govern romantic interpersonal relationships wouldn't apply to them. For instance, it might not be wrong for our evolutionary counterparts to cheat on each other. So, it seems plausible that, if our evolutionary counterparts had very different moral beliefs and thus different non-doxastic moral psychologies, the robustly moral facts that hold for them would differ as well. That, though, means that they cannot be our

⁴⁸ See Wedgwood 2010: 7.

⁴⁹ This point is exacerbated within the framework of the Total Evidence View: we are allowed to rely on our first-order moral evidence when assessing evidence of peer disagreement. But once we do, it becomes hard to regard our radically disagreeing evolutionary counterparts as epistemic peers – whether or not they share our evidence. Thanks to Camil Golub for this point.

⁵⁰ Not all, of course. For instance, it might be a fact that eco-systems have moral value. But their value wouldn't directly depend on any feature of our non-doxastic moral psychologies.

peers, metaphysically speaking. After all, those counterparts do not partake in our robustly moral reality (or they aren't bound by most of the same moral laws).

In sum, it appears that the truth of the first assumption, namely that our counterparts have evolved to radically morally disagree with us, undermines the second assumption, namely that our counterparts are our *peers*, both epistemically and metaphysically. As a result, those two assumptions are inconsistent – or, at least, stand in serious tension. But, importantly, both assumptions are indispensable to any evolutionary debunking argument based on moral peer disagreement. We cannot just relax them.⁵¹ So, evolutionary debunkers such as Bogardus (2016) and Mogensen (ms, 2017) face a serious worry: it does not seem plausible that evolutionary considerations amount to evidence of moral peer disagreement. That, in turn, means that their argument doesn't get off the ground.

Where are we now? In this section, I sketched how evolutionary debunkers such as Bogardus (2016) and Mogensen (ms, 2017) can avail themselves of the resources of the Total Evidence View to get around the self-defeat objection. Their key move consists in establishing that moral disagreements with evolutionary counterparts are structurally similar to **Offside Call** – and then adopt Kelly's (2010) diagnosis of such standard cases. However, that only works if we can plausibly suppose that evolutionary consideration supply evidence of moral peer disagreement. That supposition, though, strikes me as implausible. On reflection, we cannot simultaneously assume that our counterparts have evolved to radically morally disagree with us – and count them as our *peers*, both epistemically and metaphysically. So, their key move falters and, by implication, they cannot make good on the second part or (B) of my strategy. But if they cannot do that, the evolutionary debunking argument from moral peer disagreement defeats itself.⁵²

8. Conclusion

In this chapter, I closely examined an evolutionary debunking argument based on the epistemic principle that higher-order evidence of error defeats justification. More precisely, I argued that any such argument ultimately cannot avert self-defeat. After presenting both the argument and the worry in more depth, my argument proceeded in

⁵¹ As rehearsed above, if we relax the first one, the argument ceases to be evolutionary – and if we relax the second one, it ceases to be about disagreement between peers.

⁵² Of course, the worry I develop in this section might actually be so powerful that it would undermine any evolutionary debunking argument from moral peer disagreement irrespective of self-defeat. After all, my worry targets the assumption at its very core: that evolutionary considerations amount to such evidence in the first place.

two steps. First, I sketched out an initially promising strategy against self-defeat: evolutionary debunkers should change their background view on the epistemic significance of higher-order evidence of error from Conciliationism to Kelly's (2010) Total Evidence View. However, as I argued in a second step, both versions of the evolutionary debunking argument from higher-order evidence of error fail to take advantage of that switch. More specifically, both versions fall short of establishing that the total evidence available to the robust moral realist defeats the justification of their moral beliefs. Debunkers such as Street (2006), who construe evolutionary considerations as evidence of moral unreliability, lack a plausible theoretical notion of evidential weight. In contrast, the evolutionary disagreement argument developed by Bogardus (2016) and Mogensen (ms, 2017) rests on two inconsistent assumptions about the epistemic (and metaphysical) credentials of our evolutionary counterparts.

What does that imply for the broader dialectic of my thesis? For robust moral realists, my discussion is good news: they don't have to fear evolutionary debunking arguments from higher-order evidence of error. In contrast, evolutionary debunkers must look elsewhere for an epistemic principle that licenses the inference from the commitment to the existence of robust moral facts and evolutionary explanations of moral belief to moral scepticism. Chapter 4 aims to offer exactly that.

CHAPTER 4: EXPLANATORY DISPENSABILITY

1. Introduction

Evolutionary debunking arguments attempt to show that the evolutionary theory, when combined with a commitment to robust moral objectivity, leads to moral scepticism. However, it has proven difficult so far to find an epistemic principle that licenses this inference. In Chapter 2, we discussed and rejected the principle that reliability requires explanation. In Chapter 3, we closely examined the – ultimately inadequate – principle that higher-order evidence of error defeats justification. In this chapter, we focus on one final suggestion: that, for our beliefs to be justified, they must be best explained by the facts they are about. When combined with the idea that evolution, not robust moral facts, provides the best explanation of our moral beliefs, this principle implies moral scepticism: our moral beliefs are unjustified.

My aim in this chapter is to offer a systematic, comprehensive defense of an evolutionary debunking argument inspired by that line of thinking.¹ According to that specific argument, since robust moral facts are dispensable to the best (namely: evolutionary) explanation of our moral beliefs and there are no additional, non-abductive reasons to maintain them, those beliefs lose their justification. To establish its plausibility, I defend this evolutionary debunking argument against three powerful recent criticisms. More specifically, I argue that neither Andreas Mogensen's (2015, 2016) worry about levels of evolutionary explanation, nor Terrence Cuneo's (2007) companion in guilt strategy nor David Enoch's (2012, 2016) appeal to deliberative indispensability should make us doubt its soundness. The upshot of my discussion is significant: at last, we have a compelling evolutionary debunking challenge to robust moral realism, inspired by the epistemic principle that, for our beliefs to be justified, they must be best explained by the facts they are about.

Here is my plan. §2 presents the evolutionary debunking argument from explanatory dispensability in more detail. The remainder of the chapter then focuses on defending its premises. §3 argues that – *contra* Mogensen (2015, 2016) – we shouldn't think of evolutionary explanations of moral belief as complementary to explanations in terms of robust moral facts. §4 shows why – *contra* Cuneo (2007) – robust epistemic facts aren't

¹ Why only *inspired*? Because, unlike more basic explanatory dispensability arguments, it involves a *burden shift*: its sceptical conclusion is conditional on the lack of additional, non-abductive reasons for robust moral belief. §2 explains that feature in more detail.

explanatorily dispensable in the relevant sense and thus cannot serve as companions in guilt. §5 criticizes Enoch's (2012, 2016) proposal that the deliberative indispensability of robust moral facts gives us additional, non-abductive reason to believe in their existence. §6 concludes.

2. Evolutionary Debunking and Explanatory Dispensability

Recently, some philosophers have argued that insights from evolutionary theory debunk belief in robust moral facts. To make their case, a few of these evolutionary debunkers, including Richard Joyce (2006: Ch. 6), Philipp Kitcher (2005: 163-85) and Jack Woods (2016), rely on a variation of the epistemic principle that, for our beliefs to be justified, they must be best explained by the facts they are about.² Despite some differences in detail, their arguments share the following structure:³

- (1) Evolutionary, not robust moral, facts best explain why we hold certain moral beliefs and not others. (Or, robust moral facts are *dispensable* to the *best explanations* of our moral beliefs.)⁴
- (2) If our moral beliefs are not best explained by robust moral facts and there are no additional, non-abductive reasons for maintaining them, that defeats their justification. (Or, if robust moral facts are *dispensable* to the *best explanation* of our moral beliefs and there are no additional, non-abductive reasons for maintaining them, that defeats the justification of those beliefs (that purport to represent them).)
- (3) There are no additional, non-abductive reasons for maintaining moral belief, robustly construed.
- (4) Therefore, (attention to) facts about evolution defeat(s) the justification of our moral beliefs, robustly construed.

This is – what we might call – the *best evolutionary explanation argument* against robust moral realism. If sound, it forces moral realists into moral scepticism: we cannot ever form

² The original principle can be found in Harman (1977). The variation below features an additional reasons clause and is inspired by Woods (2016). See also fn. 3.

³ Woods (2016) clearly develops an argument similar to the one below. In contrast, both Joyce (2006) and Kitcher (2005) can be *charitably interpreted* as advancing something to that effect – instead of a more basic, question-begging version without appeal to the additional reasons clause in premises (2) and (3). For more discussion of evolutionary debunking from explanatory dispensability, see also Ruse 1986, Bogardus 2016: fn. 26 as well as Clarke-Doane 2015 & 2016.

⁴ Some might feel queasy about explanatory (in)dispensability talk. I employ it to formulate above argument because large swaths of the literature are couched in it. However, if you have reservations about it, simply ignore the second half of each premise and the conclusion.

justified beliefs about robustly moral facts. In the remainder of this section, I shall explain in more detail what these premises mean, why they might be true – and how they jointly seem to entail the conclusion. Doing so will set the stage for my systematic, comprehensive defense of the argument in §3 to §5, where I argue that the most powerful objections to each of the three premises turn out to be unconvincing.

2.1. Evolutionary Explanations of Moral Belief

The first premise appeals to the evolutionary explanation of moral beliefs. Very roughly, such explanation say that we believe what we believe in virtue of natural selection.⁵ For example, we believe that incest is wrong or that we have special moral obligations to family members because that proved *adaptive* or helped our ancestors survive. Evolutionary explanations share three characteristics that are crucial for evolutionary debunking.

First, evolutionary explanations don't avert to (or imply or ineliminably feature) robust moral facts or truths. After all, evolution hasn't selected for robustly *true*, but adaptive, moral beliefs. As Street (2008) puts it: 'the best explanation of why we tend to value our survival is not that it's independently true that our survival is valuable ... but rather, much more simply, that creatures who valued their survival tended to do what promoted it, and therefore left more descendants' (209). In that respect, the evolutionary explanation of moral belief differs significantly from that of our perceptual beliefs about, say, mid-sized objects. It is overwhelmingly plausible that evolution has selected for their *truth*. After all, having true perceptual beliefs would have enhanced the fitness of our ancestors, helping them locate food and detect environmental threats. Second, evolutionary explanations are – what we might call – *competitive*: they compete with other kinds of explanations for overall plausibility and, if successful, are *incompatible* with them. For instance, evolutionary explanations of moral belief cannot hold true at the same time as *moral*

⁵ As mentioned in Chapter 1, it is implausible that evolution explains *particular* moral beliefs. Rather, it explains our capacity for moral judgment. That, in turn, can be understood as either dispositions to apply moral concepts (e.g. Joyce 2001, 2006) or basic evaluative dispositions that shape our moral beliefs (e.g. Street 2006, 2008, 2011). However, this complication doesn't matter for my argument above.

explanations: explanations of moral beliefs in terms of robust moral facts.⁶ Rather, if the evolutionary explanation proves successful, that rules out moral explanations.⁷

Third and finally, evolutionary explanations of moral beliefs are *best* – or, at least, *better* than moral explanations. And that seems plausible: to start with, moral explanations cannot be causal. After all, (most) robust moral realists deny that moral facts have causal powers. But if moral explanations aren't causal, what are they – and why think of them as explanations at all?⁸ Surely, if we are unsure whether a given putative explanation is an explanation at all (or, more precisely, if we are unsure how a given putative explanation is supposed to *work*), that explanation cannot be as good, let alone better, than a well-respected scientific kind of explanation such as explanation by natural selection. Further, evolutionary explanations of moral beliefs aren't just widely accepted as explanations. Rather, they also satisfy several criteria of what makes explanations *good*. More specifically, they are more *parsimonious* than moral explanations: they don't feature *sui generis* robust moral facts in the explaining phrase or explanans.⁹ And evolutionary explanations of moral beliefs are more *general* than moral explanations: they don't just explain why we hold the moral beliefs we hold, but also other characteristics of human psychology, such as why we perceive the world as we do or why we reason the way we do. So, evolutionary forces explain why we hold certain moral beliefs and not others better than moral explanations. And given that evolutionary explanations of human behavior and thinking are generally considered superior to other kinds of explanations, it seems to follow that evolution *best* explains our moral beliefs.¹⁰ Or, put differently, robust moral facts are *dispensable* to the

⁶ There is some disagreement in the literature about the *scope* of moral explanations. Some construe them broadly, such that a moral explanation is '...an[y] explanation of a particular or type of event (or fact or state of affairs) that features moral terms in the explaining phrase' (Sinclair 2013: 1). So, moral facts might explain both other moral facts as well as non-moral facts, including facts about moral beliefs. Others have a narrower conception: for them, a moral explanation is an appeal to a moral property of a person or situation to explain the *formation of a moral judgment*. In what follows, I shall operate on the latter, narrower definition. After all, the (epistemological) problem under discussion is not that moral facts don't explain anything at all, since they may explain other moral facts. Rather, the problem is that moral facts don't 'hook up properly with our ability to detect facts' (Sayre-McCord 1988: 443), or, put differently, don't explain *non-moral facts*, especially our making moral observations and forming moral beliefs.

⁷ I discuss a prominent worry about this feature of evolutionary explanations in §3 below.

⁸ I develop that issue further in §3 below. Woods (2016) emphasizes that robust moral realists cannot avail themselves of causal explanation – and that *that* puts pressure on them. He goes as far as including this point as an explicit premise in the formulation of his Harman-style debunking argument. For reasons of simplicity, I have decided against that, instead bringing up the point in the explication of the first premise of the argument above.

⁹ See §2.2 below for more on what underwrites such a parsimony criterion.

¹⁰ Most likely, the argument doesn't require that strong of a claim to go through. Instead, we simply need the claim that evolutionary explanations of moral beliefs are *better* than moral explanations and/or that the best explanation of moral belief, evolutionary or not, doesn't appeal to robust moral facts or truths. I take it that I have established that here.

best explanations of our moral beliefs: they play no role in the best (namely: evolutionary) explanation of moral belief.

Together, these three characteristics of evolutionary explanations of moral beliefs prove crucial to evolutionary debunking: such explanations avoid robust moral truths, rule out moral explanations – and are better than them. So, as premise (1) states, they explain why we hold certain moral beliefs, not robust moral facts.

2.2. Explanatory Dispensability and Defeat

The second premise articulates an epistemic principle that ties explanatory dispensability to the defeat of justification.¹¹ More precisely, it says that if robust moral facts are *dispensable* to the *best explanation* of our moral beliefs and there are no additional, non-abductive reasons for maintaining them, that defeats the justification of those beliefs. This premise is inspired by what Woods (2016) calls ‘negative inference to the best explanation’ or *IBE-*. It reads as follows: ‘If the truth and content of our moral beliefs is not involved in the best explanation for our possession of them, then we need *additional reasons* to believe them’ (Woods 2016: 6; emphasis mine). Importantly, these additional reasons aren’t abductive, but could be grounded in deliberative indispensability, common sense or a permissive moral epistemology.¹²

IBE- seems plausible enough. To start with, it is intuitively compelling: just as figuring in the best explanation of a given phenomenon (e.g. our possession of scientific beliefs) generates a reason for belief in a given entity (e.g. scientific truths), not figuring in the best explanation of a given phenomenon (e.g. our possession of moral beliefs) raises the requirement of giving additional reasons in support of belief in a given entity (e.g. moral truths). For illustration, consider a non-moral example: belief in the axioms of set theory. Set-theoretical truths aren’t plausibly involved in the best explanation of our set-theoretical beliefs. After all, set-theoretical truths plausibly lack causal powers.¹³ Further, many sophisticated set-theoretical beliefs wouldn’t have been evolutionarily advantageous to our ancestors. So, if we aim to maintain our set-theoretic beliefs, we need additional, non-abductive reasons for them. For example, set theory might be indispensable to

¹¹ As in the rest of my thesis, I mean ‘undermining defeater’ by ‘defeater’ here, dropping the qualifier henceforth.

¹² See Woods 2016: 8. I shall assess the two most plausible candidates for such additional reasons in §5 below – and §4.3 of Chapter 5, respectively.

¹³ For the sake of the illustration, I assume that we should be robust realists about set theory. If you aren’t on board with that, pick a different domain (e.g. mathematics, modality) that you deem more plausibly robust.

mathematics – which, in turn, might be indispensable to scientific theorizing.¹⁴ So, the scientific indispensability of set theory might give us additional, non-abductive reason to believe in set-theoretic truths. Thus, it looks like *IBE-* is intuitively compelling.

In addition to its intuitive appeal, *IBE-* is also an instance of a more general, equally plausible principle, namely *Burden Shift*: ‘If our believing in certain claims of a domain D can be well explained without any appeal to their content and truth, then we acquire the epistemic burden of explaining why we should continue to believe them in spite of their theoretical superfluousness’ (Woods 2016: 7). Again, that seems plausible, given that explanatory indispensability is a very good guide to ontology¹⁵ – and if we cannot rely on that, we inherit the task of identifying another, similarly good ontological roadmap. But if *Burden Shift* strikes us as plausible, so must *IBE-*, for the latter is an instance of the former.

Before moving on, let me stress the importance of the additional reasons clause featured in the second premise. If the evolutionary debunker dropped it and opted for direct, unconditional defeat by explanatory dispensability instead, the second premise would *beg the question* against the robust moral realist. After all, most robust moral realists reject that robust moral facts possess causal-explanatory powers – and therefore also reject abductive moral epistemology.¹⁶ In other words, they hold that moral beliefs can be justified even if they are not best explained by the facts that they are about. Instead, these robust moral realists endorse non-abductive moral epistemologies such as moral intuitionism.¹⁷ As a result, if debunkers dropped the additional reasons qualifier, their argument would rule out moral justification (for robust moral realists) *ab initio*, thus assuming what it sets out to prove.

By inserting the additional reasons clause, evolutionary debunkers can avoid that complication. After all, neither *IBE-* nor *Burden Shift* defeat the justification of moral – and, more generally, *apriori* – beliefs outright, based solely on *abductive* grounds. Rather, those principles simply *shift the burden*: given the lack of abductive reasons for robust moral or, more generally, *apriori* belief, robust moral realists must provide other reasons. And if they fail to do so and thus cannot discharge this epistemic burden, *then* the justification of their moral beliefs has been defeated. So, *IBE-* and *Burden Shift* make

¹⁴ For more on the scientific indispensability of mathematics and how that grounds ontological commitment to mathematical entities, see Colyvan 2015.

¹⁵ See Cline 2016.

¹⁶ For this criticism, Setiya 2013, Sinclair 2013, and Zimmerman 2010: Ch. 5.

¹⁷ Chapter 5 will examine moral intuitionism in depth.

defeat *conditional* on the robust moral realist's inability to provide additional reasons for moral belief. And by doing so, the second premise avoids begging the question.

2.3. Non-Abductive Reasons for Robust Moral Belief and Moral Scepticism

The third premise states that there are no additional, non-abductive reasons for belief in robustly moral facts. This is, obviously, a substantive and controversial claim: it implies that many *prima facie* plausible candidates found in the literature prove ultimately unsuitable or, at least, sufficiently problematic for the argument to go through. In the remainder of my thesis, I shall closely examine two such candidates. In §5 below, I present and evaluate a recent, widely discussed proposal by David Enoch (2012, 2016), according to which the deliberative indispensability of robust moral facts gives us (non-abductive) reason to believe in their existence. In §4 of Chapter 5, I develop and criticize another attractive option, combining moral intuitionism with a third-factor explanation of moral reliability. So, a full-blown defense of premise (3) has to wait until later. For now, let me simply stress the importance of this premise: without it, the antecedent of the conditional in premise (2) won't hold. So, defending premise (3) is crucial to the aspirations of the evolutionary debunker.

Together, premises (1) to (3) imply that the robust moral realist's belief in robust moral facts is unjustified. After all, if robust moral facts are dispensable to the best (namely: evolutionary) explanation of moral beliefs and robust moral realists lack additional, non-abductive reasons for maintaining them, that seems to defeat the justification of those beliefs. Put succinctly, attention to evolutionary explanations of moral belief – in conjunction with a critique of non-abductive sources of moral justification – undermines the justification of their robust moral beliefs.

Before moving on to a comprehensive defense of each premise, let me clarify one aspect of this conclusion. As it stands, the conclusion seems ambiguous between an epistemological and a (primarily) metaphysical interpretation. The argument could either show that robust moral realists don't have *justified* moral beliefs (and thus, if knowledge requires justification, lack moral *knowledge*) – or that robust moral facts don't exist (and that corresponding beliefs are thus unjustified).¹⁸ However, according to Alex Miller (2003: Ch. 8) and Neil Sinclair (2013), this ambiguity is problematic: we should clearly

¹⁸ This ambiguity or conflation is pervasive in the large literature on Harman-style debunking arguments: contributors often write that (robust) moral realists are not *justified* in their *belief* in the *existence* of objective moral facts. See, for example, Majors 2007.

distinguish between two versions of the Harman-style debunking argument, depending on whether its conclusion is epistemological or metaphysical. My formulation above, though, blends both interpretations together impermissibly.

But I doubt that the distinction between epistemology and metaphysics matters *that* much in our context. Sure, it is possible that, given robust moral realism, we might lack justification or knowledge of moral entities, despite the fact that those entities *do* exist. That, though, doesn't seem more than a theoretical possibility: a realist view that would render moral entities systematically inaccessible to us in that manner doesn't look attractive to start with (and thus won't be worth defending). Therefore, it is unclear whether we can or need to draw the requisite distinction here. Further, suppose that we could meaningfully draw the distinction and that the conclusion of the argument therefore is merely epistemological: robust moral realists don't have justified moral beliefs (due to evolutionary forces and the lack of additional, non-abductive reasons). In that case, it seems plausible that we are licensed to also draw a metaphysical conclusion.¹⁹ After all, if there really are no genuine reasons for belief in robust moral facts, why think that they even exist? So, it once again doesn't seem problematic that the conclusion of the argument above is ambiguous in that sense. It is fine as is.²⁰

In this section, my aim has been to introduce the *best evolutionary explanation argument* against robust moral realism. More specifically, I explained in more detail what its premises mean, why we might initially think them true – and how they jointly seem to entail the conclusion that robust moral beliefs are unjustified. In the remainder of this chapter, I offer a systematic, comprehensive defense of each premise of this argument, starting with the claim that evolution best explains moral belief in §3.

3. The Nature of Evolutionary Explanations

In this section, my aim is to defend the first premise of the argument articulated in §2, namely that evolutionary, not robust moral, facts best explain our moral beliefs. The most compelling criticism of that claim focuses on whether evolutionary explanations are competitive. Recently, some have argued that they are not: even if successful, evolutionary

¹⁹ See Worsnip 2016: 228 for a similar point, yet in the context of the *positive* deliberative indispensability argument.

²⁰ Another issue: the conclusion references 'attention to' evolutionary explanations to accommodate epistemic internalism, according to which we need to be aware of defeating evidence for it to undermine our justified belief. I didn't include that internalist qualifier in the formulation of the premises for reasons of simplicity and readability.

explanations don't rule out explanations of moral beliefs in terms of robust moral facts. However, as I argue, that line of criticism proves untenable.

Andreas Mogensen (2015) has recently put forth an objection that seems to show that the first premise of the argument developed above is false. More specifically, his objection seems to call into question an important characteristic of evolutionary explanations of moral belief, namely that they are *competitive* and, if successful, *exclude* or *rule out* moral explanations. To establish that, he points out that evolutionary explanations only operate at the level of *ultimate* causes: causes that belong to evolutionary history (e.g. natural selection). In contrast, *proximate* causes of a trait operate within an organism's lifetime (e.g. immediate triggering causes, developmental factors). Or, put differently: while proximate causes explain *how* you get a trait (i.e. causal mechanism; ontogenesis), ultimate causes explain *why* it has been promoted by natural selection (i.e. adaptive function; phylogenesis). Importantly, ultimate and proximate causes do not compete: they are *complementary*. In our context then, it seems plausible that, while natural selection (as a process indifferent to robust moral truth) is the ultimate cause of our moral beliefs, robustly moral facts may be their proximate cause.²¹ So, it looks like evolutionary explanations don't rule out moral explanations after all. Premise (1) must be false: it is not clear whether evolutionary forces best explain why we hold certain moral beliefs and not others. Or, robust moral facts are *not* dispensable to the best explanation of our moral beliefs.²²

However, proponents of the evolutionary debunking argument above can readily respond to this objection. To start with, they might take inspiration from Eleanora Severini (2016).²³ Her argument proceeds in two steps. First, she introduces a minimal constraint

²¹ Mogensen (2015: 197) credits Nozick 1981: 345 with this insight.

²² Mogensen (2016) develops this line of thought further. In that article, he argues that the *Debunking Thesis*, i.e. selection pressures explain (and thus undermine) our moral beliefs (which isn't identical with, yet implies my first premise above), conflicts with the *Negative View*, a widely accepted view in the philosophy of biology. The *Negative View* says that natural selection cannot explain the traits of *individuals* (including moral beliefs), but only the frequency with which traits occur in *populations*. (Traits of individuals, in contrast, are explained by inheritance, development and other proximate factors.) Therefore, '...it is mysterious how [the lack of selection for true moral beliefs] could make any difference to our reasons for revising these beliefs, assuming that any such reasons would have to come in the form of evidence about their discreditable causal origins' (Mogensen 2016: 1805f.). However, this line of thought is – just like the more basic one of his (2015) – subject to the worries raised below and should thus be rejected.

²³ Evolutionary debunkers could also look towards Fitzpatrick (2016a). He argues – *contra* Mogensen (2015) – that evolutionary debunkers don't conflate the ultimate/proximate distinction: 'When debunkers claim that our moral beliefs are susceptible to evolutionary explanations, what they mean is not that there are evolutionary as opposed to proximate causal explanations for these beliefs, but rather that there are evolutionary explanations for these beliefs that, when fully filled-out, appeal at every level only to causal

on complementary explanations: they must be *consistent* or ‘one explanation cannot imply the negation of the other one’ (Severini 2016: 870). Second, she points out that evolutionary and moral explanations of moral belief are *inconsistent*: ‘it cannot be the case that the evolutionary explanation of moral beliefs rules out the justification of the realist claim about the existence of moral facts at the ultimately level, and, at the same time, moral realism holds up such an existence at the proximate level’ (ibid.). As a result, these evolutionary and moral explanations of moral belief *cannot* be complementary. But if they are not complementary, Mogensen’s objection above fails.²⁴

But suppose Severini’s response doesn’t succeed. Even then, as I shall argue now, evolutionary debunkers can defend the first premise of their argument against Mogensen’s (2015) objection. For suppose Mogensen is right that moral and evolutionary explanations of moral belief are indeed complementary: while natural selection serves as ultimate cause, robust moral facts serve as proximate cause. In that case, robust moral realists face a tricky question: what exactly is the *explanatory* relationship between robust moral facts and moral beliefs? Or, more precisely: in what sense are robust moral facts proximate *causes* of our moral belief?²⁵

factors [e.g. cognitive or emotional dispositions] that have nothing to do with tracking or responding to moral facts as such.’ (436) Still, Fitzpatrick thinks that evolutionary debunkers are mistaken: they overlook ‘...the possibility that although the basic psychological capacities we draw on in moral thinking originally evolved simply for Darwinian reasons, and although we may also possess certain evolved dispositions to have certain moral feelings and beliefs, we may nonetheless also be able to develop our capacities for moral reflection and reasoning in cultural contexts [following developed, learned standards internal to intelligent moral inquiry] and exercise them with a considerable degree of independence from those influences [similar to mathematics or philosophy]’ (ibid.). What Fitzpatrick ultimately has in mind is a view of moral epistemology on which ‘...we come to recognize certain moral facts by grasping the *reasons* that *justify* certain moral propositional contents (rather than simply being caused to believe things by truth-indifferent factors)’ (437). Such a view isn’t ruled out by evolutionary biology or science more generally. (See Fitzpatrick 2015 & 2016b for more on that.) Importantly, Fitzpatrick seems to concede the truth of the first premise. Instead, he challenges the third premise: he thinks that there are additional, non-abductive reasons for moral belief, having to do with a broader view on apriori knowledge on which we grasp apriori truths, etc. I shall develop and critically examine this strategy in depth in Chapter 5.

²⁴ Severini (2016: 871f.) also presses another point: proximate and ultimate causes are often not neatly distinguishable because they *interact* with each other. For instance, when constructing niches, organisms change the environment, transforming selective patterns and thus showing that ‘a proximate cause can lead to a change in selection pressure [or ultimate cause] that in turn ends up steering the evolutionary process’ (872). This, though, seems problematic for Mogensen’s objection, which presupposes that we can draw a neat ultimate/proximate distinction such that natural selection ultimately explains moral beliefs, while robust moral facts proximately cause them. However, it is unclear how plausible it is that moral beliefs evolved within a moral niche, as Severini suggests. So, evolutionary debunkers might be wary to wed their argument to such controversial commitments.

²⁵ Severini (2016) raises a similar point: if we grant that moral facts figure in proximate explanations of our moral beliefs, the realist must ‘point out the mechanism thanks to which we come to have true moral beliefs

There are compelling reasons for doubting that robust moral realists can answer that question satisfactorily. To begin with, as mentioned above, most robust moral realists cannot argue that robust moral facts literally and directly *cause* our moral beliefs, given their lack of causal powers. That rules out one plausible option from the outset.²⁶ But what about *indirect* causal explanations? After all, we sometimes form beliefs about *particular* moral facts (e.g. that it is wrong for Felicity to withhold food from her dog). In those cases, it seems attractive for robust moral realists to explain the relationship between our moral beliefs and the corresponding particular robust moral facts indirectly, namely via supervenience. According to that kind of explanation, the non-moral facts in the supervenience base *cause* our moral beliefs, while we have apriori knowledge of how moral supervenience works.²⁷ However, I worry that such an explanation doesn't provide a completely satisfactory answer. Why? Because, as many critics have argued, robust moral realists struggle with offering an account of moral (or, more generally, normative) supervenience as a metaphysically necessary relationship between metaphysically discontinuous entities.²⁸ But if they cannot offer such an account, the explanation above remains incomplete: while it specifies the relationship between moral beliefs and non-moral facts as causal, it fails to explain how those non-moral facts metaphysically relate to the robust moral facts. As a result, the issue of the exact explanatory relationship between robust moral facts and moral beliefs still hasn't been addressed satisfactorily.

In short, robust moral realists face obstacles answering the question of what the explanatory relationship between robust moral facts and moral beliefs exactly is. But if they cannot adequately address that question, it remains unclear how proximate moral explanations are supposed to work. And if that remains mysterious, it doesn't really matter whether such explanations would complement evolutionary explanations of moral belief,

and explain why such a mechanism evolved? (871). However, she motivates this explanatory need not the same way I will below. Rather, she stresses – as above – that the proximate moral explanation couldn't *contradict* the ultimate evolutionary one, which it seems to do, though, given that we would have a belief-forming mechanism that tracks moral facts at the proximate level without being able to explain this capacity at the ultimate level. (In contrast, camouflage and digestion in the insect, for example, are distinct, yet complementary, but 'tracking moral facts and not tracking moral facts are not' (871).) But since we concede Mogensen's point at this stage of the dialectic, an alternative motivation for the explanatory need is required, which I shall provide in what follows in the main text.

²⁶ Oddie (2005) thinks that robust moral facts *do* have causal powers. So, he would be an exception. But since I am concerned with paradigmatic robust moral realism (and that view denies that moral facts are causally efficacious), I shall set him aside.

²⁷ For such a story, see Zangwill (2006). Note that it might not prove attractive for beliefs about general moral facts, given that those are, according to some robust moral realists, ungrounded (and therefore don't supervene on non-moral facts). Enoch (ms) develops such a picture in detail.

²⁸ See McPherson 2012 and Vayrynen 2017.

as Mogensen (2015) alleges. So, even if Severini's response turns out to be unconvincing, the first premise of the evolutionary debunking argument above can be defended.²⁹

Where does that leave us? In this section, my aim has been to defend the first premise of the argument articulated in §2 against a recent, initially compelling objection developed by Mogensen (2015). According to him, ultimate evolutionary explanations don't rule out proximate explanations of moral belief in terms of robust moral facts. However, his objection turned out to lack bite. First, evolutionary and moral explanations of robust moral belief cannot be complementary, given their inconsistency. Further, even if evolutionary explanations are complementary, that doesn't matter as long as the workings of moral explanations remain mysterious. So, the first premise emerges unscathed: evolutionary, not robust moral, facts best explain our moral beliefs.

4. Companions in Guilt

In this section, my aim is to defend the second premise of the argument developed in §2 against a companion in guilt strategy. According to that strategy, robust *epistemic* facts are also explanatorily dispensable, yet that doesn't render them metaphysically problematic. So, the second premise must be false: even if robust moral facts are dispensable to the best explanation of our moral beliefs (and there are no additional, non-abductive reasons for maintaining them), that doesn't mean that their justification has been defeated. But, as I argue, such a companion in guilt strategy doesn't get off the ground. On reflection, robust epistemic facts aren't explanatorily dispensable in the relevant sense.

²⁹ Some might worry that my point here is too nuclear, defending the first premise by seemingly establishing that there are no moral explanations for robust moral realists. But I think that my point is more modest: establishing the complementarity of evolutionary and moral explanations – as Mogensen (2015) aims to do – won't be sufficient to undermine the first premise of the Harman-style evolutionary debunking arguments, given the difficulties involved for robust moral realists in showing *how moral explanations work* in the first place. Still, my point *does* have further substantive implications: it might hamper attempts at undermining the two other important characteristics of evolutionary explanations of moral belief that the first premise relies on (and thus shield the premise from such challenges). For instance, some robust moral realists (e.g. Roberts 2016) argue that moral explanations are *better* than non-moral, scientific explanations because they generalize and discriminate properly. (Or, put differently: such realists claim that moral facts are *indispensable* to moral explanations.) If that were correct, the first premise would be false. But my point seems to show that this cannot be: if the workings of moral explanations remain metaphysically mysterious, they cannot be better than widely-accepted scientific explanations of moral belief.

4.1. Explanatory Dispensability and Companions in Guilt

Robust moral realists might want to defend themselves against the evolutionary debunking argument developed in §2 by relying on a *companion in guilt argument*.³⁰ Such a strategy, even though not widely used to assuage explanatory concerns,³¹ has proven successful against arguments for moral error theory centering on moral categoricity.³² So, it merits closer attention.

Any companion in guilt strategy follows the same template. To start with, we have the threat to a given theory by some sceptical argument, stating that entities E1 don't exist because of (metaphysical or epistemological) feature F. In response, the companion in guilt theorist makes two claims. First, the *Parity Claim*: if entities E1 don't exist because of F, neither do entities E2, given that they share feature F. So, E1 and E2 are *on a par* with respect to F. Second, she puts forth the *Existence Claim*: we have good reason to think that entities E2 exist, despite feature F. So, having feature F isn't enough to deny the existence of a given class of entities, including E1. We thus have good reason to think that entities E1 exist as well.³³

For illustration, take robust moral realism as the theory under threat by moral error theory. Moral error theory says that that robust moral facts don't exist because they are categorical: they obtain, if they do, independently of our aims, desires or interests.³⁴ In response, the companion in guilt theorist points out that robust *epistemic* facts, i.e. facts about the normative relation between beliefs and evidence, are also categorical. Epistemic facts are on a par with moral facts in that respect. For example, you should believe that you will die before your 125th birthday, given the available statistical evidence on average life expectancies – even if you desperately don't want that to be the case.. So, if robust moral facts don't exist because of being categorical, neither do robust epistemic facts. They are companions in guilt. However, we have excellent reason to believe in robust epistemic facts, despite their categoricity. For instance, denying their existence leads to implausibly

³⁰ See Cuneo 2007, Rowland 2013 & forthcoming, and Das 2017 & forthcoming.

³¹ To my knowledge, only Cuneo (2007: 103-7) runs such a strategy in response to explanatory dispensability arguments against robust moral realism.

³² See the exchange between, on the one hand, Cowie 2014 & 2016 and, on the other, Das 2017 & forthcoming as well as Rowland 2013 & forthcoming. In my mind, Rowland and Das competently respond to Cowie's criticism of companion in guilt strategies.

³³ The terminology for the key claims follows Cowie 2014.

³⁴ At least, that is one strand of argument in support of moral error theory. See, for example, Mackie 1977 or Garner 1990. Perhaps, only reasons, not facts, can be categorical. However, that complication won't matter for our purposes here.

radical scepticism.³⁵ So, being categorical isn't enough to deny the existence of either robust epistemic or moral facts. In this manner, robust moral realism escapes the argument put forth by moral error theorists.

To make a companion in guilt strategy work as a response to the Harman-style evolutionary debunking argument advanced in §2, we therefore need to establish two things. First, it must be *prima facie* plausible that robust moral and epistemic facts are on a par with respect to explanatory dispensability. Let's call this *Explanatory Parity*. Second, we need to have good reason to believe in robust epistemic facts. Let's call this *Epistemic Existence*. In the remainder of this sub-section, I shall briefly motivate both claims as well as explain their relevance to the evolutionary debunking argument put forth in §2. Doing so will provide the background necessary for developing my objection to the strategy in §4.2.

Why think that *Explanatory Parity* holds between robust morality and epistemology? To see why robust epistemic facts might be just as problematically *dispensable* to our best explanations of epistemic beliefs as robust moral facts to our best explanations of moral beliefs, we can turn to an argument by Sharon Street (2009). For her, the best explanation of our epistemic beliefs is also an evolutionary one. Very roughly, we hold our epistemic beliefs in virtue of natural selection. For instance, we believe that induction is reliable or wishful thinking is irrational because thinking with those epistemic concepts helped our ancestors survive. But, importantly, this kind of explanation doesn't avert to robust epistemic facts. After all, evolution didn't select for *true* epistemic beliefs, but for those that are most fitness-enhancing.³⁶ Or, to quote Street: 'How does it help [for purposes of explaining our ancestors' survival] to say that creatures who made the normative judgment expressing the rule of induction grasped a fact about [robust] normative reasons?' (2009: 241). So, robust epistemic facts seem dispensable to our best, namely evolutionary, explanations of epistemic beliefs. But that seems just as metaphysically problematic for epistemic facts as it is for moral ones. In this manner, we can establish the *Parity Claim* in our context.

What about *Epistemic Existence*? We can motivate the claim that (robust) epistemic facts exist in at least two ways. And since *Epistemic Existence*, unlike *Explanatory Parity*, is

³⁵ See Rowland 2013. I shall present Rowland's argument in more detail below.

³⁶ For this to work, as Street (2009) points out, we need to make the reasonable assumptions that having roughly true beliefs promotes survival and reproduction (see Stich 1990: Ch. 3 for discussion, Plantinga 1993 for criticism) and that there is something that plays the role of belief as we know them (e.g. basis for action when combined with desires, etc.).

common to any companion in guilt strategy in support of robust moral realism, we can just borrow whatever reasons the general debate generates.

To start with, as Cuneo (2007: Ch. 4) argues, anyone denying the existence of robust epistemic facts faces a deeply unattractive *dilemma*. Either, there is epistemic reason to believe epistemic nihilism, the view that there are no epistemic facts or reasons – in which case the view is *self-defeating*. After all, the view would then presuppose the very sorts of entity (e.g. facts about what we rationally ought to believe) that it claims not to exist. Or, there is no epistemic reason to believe epistemic nihilism, despite its truth. In that case, though, there are not just no epistemic reason to believe epistemic nihilism, but no epistemic reasons to believe *anything*. In other words, this implies a radical form of epistemological scepticism, according to which no propositional attitudes can exhibit epistemic merit (e.g. justification) or demerit (e.g. irrationality). Such epistemological scepticism, just like self-defeat, is deeply theoretically unattractive. So, we should reject epistemic nihilism – and affirm the existence of (robust) epistemic facts.³⁷

Further, Richard Rowland (2013: 13-15) elaborates on why the radical scepticism entailed by epistemic nihilism seems implausible. More specifically, he argues that we *do* have epistemic reasons for belief. After all, if we do not, that means that nobody knows anything. But some people do have some knowledge! For instance, I know that I don't know everything (e.g. what sports fishing rods are made of), that bachelors aren't women (just by understanding the concept BACHELOR) or that I am having a thought right now. As Rowland puts it: 'So long as I believe that there is a thought I am right that there is thought – I could not be wrong about it – and so long as I believe that there is thought on the basis that I am thinking right now, my belief that there is thought is based in an appropriate way for me to know that there is thought' (2013: 15). So, some people do have some knowledge. However, it is almost universally accepted that if S knows p, there is some epistemic justification for believing that p.³⁸ For instance, if I know that I am currently walking through Edinburgh's Old Town, there must be some justification for that such as my perceptual experience of St. Giles cathedral or Edinburgh castle. Further,

³⁷ Cuneo (2007: 120f.) also notes a third implication of the truth of epistemic nihilism. In particular, if epistemic nihilism were true, there couldn't be any arguments for anything. This is because the premises of non-question-begging arguments offer evidential support (and thus are epistemic reasons) for the conclusion. But if there are no epistemic reasons, we couldn't have any arguments (including the one above according to which nihilism implies radical scepticism).

³⁸ Rowland is quick to note that we shouldn't confuse this claim with what the reliabilists deny. They hold that in order to know p, belief in p has to be true and reliably produced (and thus justified), yet we don't need justification for the belief that p was reliably produced (or we don't need to know that we know p). That, though, doesn't mean that, for them, knowledge doesn't require justification. Similarly, even for knowledge first theorists such as Williamson (2000), knowledge entails justification.

it also seems that if there is some epistemic justification for believing p , then there is epistemic reason for believing p . For instance, it would be odd to claim that I am justified in believing that I am currently walking through Edinburgh's Old Town, but that there is no epistemic reason at all for me to believe that. But if knowing that p entails being justified in believing that p – and that, in turn, entails there being an epistemic reason for p , that implies that if S knows p , there is an epistemic reason to believe p . And since at least some S know at least some p , there must exist some corresponding epistemic reasons for p . So, the radical scepticism about epistemic reasons entailed by epistemic nihilism must be false.

In sum, both *Explanatory Parity* and *Epistemic Existence* seem plausible. But how do they bear on the evolutionary debunking argument developed in §2 above? Jointly, those two claims appear to show that there is something wrong with the second premise of the argument. To wit, that premise states that, if robust moral facts are dispensable to the best explanation of our moral beliefs (and there are no other, non-abductive reasons for them), that defeats the justification of those beliefs. So, that premise rests on the assumption that being explanatorily dispensable is metaphysically problematic. That, however, seems to be false, in light of *Explanatory Parity* and *Epistemic Existence*: robust *epistemic* facts are also explanatorily dispensable, yet that doesn't render them metaphysically problematic. So, the explanatory dispensability of robust moral facts doesn't impose a burden on robust moral realists or threatens the justification of their moral beliefs. But if it does not, why think that the second premise of the argument in §2 is true? In this manner, the companion in guilt strategy helps robust moral realists defend themselves against evolutionary debunking.

4.2. Explanatory Dispensability and the Limits of Guilt

However, any companion in guilt strategy deployed in response to evolutionary debunking arguments from explanatory dispensability faces a fatal issue. On reflection, *Explanatory Parity* must be false: (robust) morality and (robust) epistemology cannot be on a par with respect to *problematic* explanatory dispensability. This is because the Harman-style argument against epistemic, yet not moral, facts proves *self-effacing*. (And the same is true if we swap in mathematical, logical or evaluative facts.) But if robust moral realists cannot establish *Explanatory Parity*, their companion in guilt strategy fails.

To see how the issue arises, note first that *Explanatory Parity* doesn't just require that robust moral and epistemic facts are explanatorily dispensable. Rather, it requires that both are *problematically* explanatorily dispensable. After all, if explanatory idleness wasn't *prima facie* problematic for both (and would thus make them 'guilty'), we couldn't mount

a companion in *guilt* argument: an argument to the effect that one class of entities (namely robust epistemic facts) shares the problem of another (namely robust moral facts), yet without suffering the consequences (in terms of plausibility, credibility or ontological status). To establish that moral and epistemic facts are on a par with respect to their *problematic* explanatory dispensability, we need a *prima facie* plausible argument applicable to both of them, explaining why explanatory dispensability is problematic.³⁹ Now, we already have such a dialectically appropriate argument in the moral case: the Harman-style explanatory challenge to robust moral realism developed in §2, which shows that the explanatory idleness of moral facts is problematic because it gives us reason to think that they don't exist. (Of course, that reason might not be *decisive*. Rather, as we saw in §2, there might be further, non-abductive reasons for their existence).

But what about the epistemic case? There is reason to think that there cannot be a *prima facie* plausible analogue Harman-style challenge to robust epistemic facts. To see why, note that any Harman-style debunking challenge to putatively *apriori* facts *presupposes* some epistemic facts. In particular, it presupposes facts that underwrite the epistemic principles such as *IBE-* or *Burden Shift*, which the challenge employs as explicit or implicit premises.⁴⁰ That, though, wreaks havoc for the Harman-style debunking challenge to robust epistemic realism: it would simultaneously *accept* (in the second premise, by presupposition) and *deny* (in the conclusion) the existence of robust epistemic facts. Therefore, such a challenge would be *self-effacing*.⁴¹ As a result, there cannot be a *prima facie* plausible Harman-style challenge to robust epistemic realism.

Now, what does that all mean for *Explanatory Parity*? Even though robust morality and epistemology might be equally explanatorily dispensable, that is only *problematic* for the

³⁹ Why not think that explanatory idleness is *brutely* problematic or problematic without further *argument*? Because explanatory dispensability seems like a *neutral* feature of (a given class of) entities: it is only once we combine it with an appeal to a *methodological principle* such as *IBE-*, *Burden Shift* or a parsimony requirement that the feature becomes problematic in theory construction. Combining explanatory idleness with such a principle, however, will amount to an *argument*, most likely to the conclusion that we are unjustified or have less reason in believing in the problematically explanatorily dispensable entities.

⁴⁰ Why assume that those facts should be construed *robustly*? Because a robust moral realist invested in epistemic-moral parity *must* do that. Without doing so, her companion in guilt strategy doesn't get off the ground. For instance, Cline (2016) denies the existence of epistemic reasons, robustly construed, and argues that this needn't imply that we cannot have serious epistemology. But he is a global normative error theorist!

⁴¹ See Woods 2016: 16f. In addition to being self-effacing, the analogue argument of §2 against robust epistemic facts also shares another flaw: its premise (3) would be false. After all, denying the existence of robust epistemic facts leads to self-defeating epistemic nihilism or implausible epistemological scepticism, as my discussion of *Epistemic Existence* above showed. But avoiding those outcomes might give us additional, non-abductive reason for belief in robust epistemic facts – which means that premise (3) of the analogue argument must be false.

former, yet not the latter. By implication, robust moral realists cannot establish *Explanatory Parity*: it is false that robust moral and epistemic facts are equally problematic with respect to explanatory dispensability. Rather, there are limits to the guilt of robust epistemic facts. But, of course, if robust moral realists cannot establish *Explanatory Parity*, the companion in guilt strategy must fail.⁴²

4.3. The Severity of The Worry

To illustrate the severity of my worry, let me briefly address some potential responses on behalf of robust moral realism. First, proponents of that view might want to reformulate *Explanatory Parity*. For instance, they might narrow its scope, arguing that what matters for parity is only *existential* implications or entailment: if moral facts don't exist, neither do epistemic facts. Meanwhile, the *grounds* for why this implication holds don't enter into our formulation of parity. But once we conceive of parity in that manner (or endorse – what we might call – *Existential Parity*), the fact that the explanatory idleness of moral facts is problematic, while the explanatory idleness of epistemic facts is not, doesn't undermine the parity claim.

However, *Existential Parity* strikes me as ad-hoc and implausible, at least once we consider companion in guilt strategies against moral error theory more generally. To secure the parity claim, those strategies are committed to saying that the inexistence of robust moral facts implies the inexistence of robust normative facts exactly *because* both are categorical. So, the *relevant* ground for why the existential entailment holds *does* enter into their formulation of parity. In our dialectical context, the relevant ground won't be categoricity, but problematic explanatory dispensability. And, if the above line of thought about such strategies more generally is correct, *that* ground should enter into our formulation of parity. Once it does, though, and we take my worry above on board, we see that robust epistemology and morality won't be on a par after all. So, narrowing the scope of *Explanatory Parity* to *Existential Parity* won't help the robust moral realist salvage their companion in guilt strategy.

Conversely, robust moral realists might want to broaden the scope of parity. In particular, they might argue that *Overall Parity* matters for their argument. Even if robust morality

⁴² This objection differs from Cowie's (2014) worry for companion in guilt strategies in at least two respects. First, it doesn't capitalize on the tension between *Parity* and *Epistemic Existence*, but shows that robust moral realists cannot establish *Parity*, irrespective of whether *Epistemic Existence* holds. Second, it isn't open to Rowland's (forthcoming) and Das' (2017) response to Cowie's (2014) objection on behalf of robust moral realism because it doesn't mischaracterize moral-epistemic parity as parity *simpliciter* (but rather as parity with respect to problematic explanatory idleness). More on the second point follows below.

and epistemology aren't on a par with respect to problematic explanatory idleness, they are on a par in many *other* respects (e.g. categoricity, authority, irreducibility). And, so the thought goes, *that* parity is all the companion in guilt theorist needs to run her strategy.

But that move seems once again implausible once we factor in the dialectical situation properly. More precisely, moral error theory calls in to question *categorical* moral facts. In response, companion in guilt theorists argue that this implies that there aren't any epistemic facts either, given that they are also categorical. So, in the context of (moral) error theory, all the other aspects of parity wouldn't matter if epistemic and moral facts weren't both categorical (the existence of the latter moral error theory denies). Similarly, all the other aspects of parity don't matter in *our* context if moral and epistemic facts aren't on a par in terms of problematic explanatory dispensability. In other words, to serve as a response to the Harman-style debunking argument against robust moral realism, the companion in guilt strategy *must* be formulated around parity regarding problematic explanatory idleness. So, broadening the scope of *Explanatory Parity* to *Overall Parity* won't do either.⁴³

Instead of reformulating *Explanatory Parity*, robust moral realists might – inspired by Sayre-McCord (1988: 451-3) – try to harness the self-effacement insight to salvage the companion in guilt strategy or, at least, robust moral realism.⁴⁴ How so? As we have seen,

⁴³ What about parity with respect to explanatory idleness only? But that wouldn't suffice. Recall that the aim of the companion in guilt theorist is to respond to Harman's challenge by drawing an analogy to epistemic facts and arguing that we have good reason to believe in their existence. However, drawing that analogy only works in our context if epistemic facts are subject to the same threat, i.e. if their explanatory idleness is somehow problematic. So, parity with respect to explanatory idleness alone won't do. In sum, trying to reformulate *Explanatory Parity* strikes me as hopeless.

⁴⁴ Why only *inspired*? Because it is debatable whether Sayre-McCord's (1988) proposal should be best thought of as a companion in guilt strategy. In conversation, he has stressed relevant differences: his proposal simply makes a point about the structure of reasons and normativity – without supporting any kind of moral or epistemic realism, it doesn't explicitly feature a parity claim or epistemic existence claim, and it doesn't emphasize guilt at all.

But I am unsure whether, on reflection, those differences hold up. First, Sayre-McCord's proposal cannot be completely non-committal with respect to the nature of normativity. After all, his *must* be an argument in support of moral realism generally, given that other meta-ethical views aren't troubled by Harman's challenge. Of course, that doesn't mean that his proposal supports a *specific* form of moral realism such as robust moral realism (as a standard companion in guilt argument would). Still, that shouldn't stop the robust moral realist from adopting his proposal – and us from examining it accordingly. Second, it seems to me that Sayre-McCord's proposal involves, albeit implicitly, both a parity and epistemic existence claim. After all, he stresses that evaluative, epistemic and moral facts are all *normative* such that any attack against one entails attacking the others. In other words: they are *on a par* with respect to normativity. Further, even though Sayre-McCord doesn't endorse the epistemic/evaluative existence premise, he offers a functionally equivalent claim: that Harman's challenge *presupposes* epistemic/evaluative facts – and that this is a reason

the Harman-style debunking challenge against robust epistemic realism presupposes epistemic facts – and therefore defeats itself. So, epistemic facts are insulated from it. But since robust epistemic facts are metaphysically and epistemologically on a par with robust moral facts (or: they are similarly *normative*), that insulation extends to moral facts as well. So, there is no such thing as a *stable* (and *selective*) Harman-style debunking argument against robust moral facts. Such an explanatory challenge either undermines all of normativity (thus leading to self-effacement) – or none.

However, this response suffers from an obvious flaw: it relies on *Overall Parity* (e.g. ‘similarly normative’) – which is, as we saw above, dialectically inappropriate. If it relied on *Explanatory Parity* instead, it clearly wouldn’t work. After all, robust moral facts are not on a par with robust epistemic facts with respect to problematic explanatory idleness *exactly because* the Harman-style debunking argument against moral facts doesn’t presuppose moral facts – and thus doesn’t self-defeat. So, the Harman-style debunking argument is *not* unstable when it comes to robust moral facts. The instability through self-defeat for all of normativity, including morality, in the face of explanatory challenges, just isn’t there. So, while we shouldn’t worry about robust epistemology, we should worry about robust morality in light of the evolutionary debunking argument put forth in §2 above⁴⁵.

Where does this leave us? My worry shows that companion in guilt strategies fail as a response to the Harman-style evolutionary debunking challenge to robust moral realism. As I have argued, the reason for that is straightforward: Harman’s challenge to robust epistemic realism is self-defeating, while the one to robust moral realism is not. As a result, robust epistemic and moral facts just cannot be on a par with respect to how explanatorily problematic they are. *Explanatory Parity* thus cannot be established, and the strategy remains incomplete. But if the strategy remains incomplete, we don’t have reason to doubt the truth of the second premise: that, if robust moral facts are dispensable to the best explanation of our moral beliefs (and there are no other, non-abductive reasons for them), that defeats the justification of those beliefs.

to believe in their existence. So, Sayre-McCord’s proposal seems to involve claims about both parity and evaluative/epistemic existence. Third, the fact that Sayre-McCord shuns accusations of guilt doesn’t seem relevant. After all, the companion in guilt theorists uses such emotionally charged notions merely in a loose and metaphorical way. What truly matters are the underlying argumentative mechanics – and those Sayre-McCord’s proposal seems to share.

⁴⁵ Put differently, robust moral realists just cannot hold on to both parity (with respect to problematic explanatory idleness) and epistemic presupposition, given that the latter undermines the former, as my worry above shows.

5. Deliberative Indispensability

In this section, my aim is to defend the third premise of the evolutionary debunking argument developed in §2, namely that there are no additional, non-abductive reasons for belief in robust moral facts. Recently, David Enoch (2012, 2016) has argued that robust moral facts, despite not figuring in our best scientific explanations, are indispensable to practical deliberation. And that, his thought goes, gives us non-abductive reason to believe in their existence.⁴⁶ If that was indeed correct, the third premise would be false – and my evolutionary debunking argument unsound. However, as I shall argue, Enoch’s appeal to deliberative indispensability turns out to be implausible.

5.1. The Deliberative Indispensability Argument

In more detail, Enoch’s (2012: Ch. 3, 2016) line of argument runs as follows:

- (1) If something is instrumentally indispensable to an intrinsically indispensable project, then we are epistemically justified (for that very reason) in believing that that thing exists.
- (2) The deliberative project is intrinsically indispensable.
- (3) Robustly normative truths are instrumentally indispensable to the deliberative project.
- (4) Therefore, we are epistemically justified in believing that there exist robustly normative truths.

Before assessing the plausibility of this argument, let me briefly motivate each premise and its conclusion. The first premise articulates a very general view about (the grounds of) epistemic justification and ontological commitment. To see how Enoch arrives at it, we must take several steps back. First off, note that scientific realists often use considerations of explanatory indispensability to justify ontological commitments. More precisely, they put forth *explanatory indispensability arguments*: if an entity proves indispensable to our best scientific explanations, we have reason to believe (or are justified in believing) that it exists.⁴⁷ For instance, if our best scientific theories quantify over electrons, that gives us reason to think that electrons exist. Importantly, such arguments rely on a rule of reasoning called *inference to the best explanation*: an inference from a higher-order claim

⁴⁶ As mentioned in §2, deliberative indispensability is just one initially attractive candidate. I critically examine another one, namely the combination of moral intuitionism and a third-factor explanation of moral reliability, in Chapter 5.

⁴⁷ See Colyvan 2015.

that some first-order claim C is the best explanation of the data (e.g. the vapor trail in a cloud chamber is best explained by the existence of electrons) to C itself (e.g. there are electrons).⁴⁸ This inference rule, in turn, is plausibly *fundamental* or *basic* (just like perception, memory or modus ponens). In other words, we can employ inference to the best explanation to form epistemically justified beliefs. However, we do not possess *epistemic* justification for the inference rule itself. After all, there seems to be something wrong about asking for further justification for employing inference to the best explanation. Still, while some belief-forming methods (e.g. perception, memory, inference to the best explanation) qualify as properly basic, others (e.g. wishful thinking) do not. Thus, there must be some deeper, principled, positively justification-relevant difference between them, vindicating the former while incriminating the latter. But what could that difference be?⁴⁹

According to Enoch, the difference is *pragmatic*: belief-forming methods are vindicated as properly basic if their success is the only relevant hope of successfully engaging in an extremely important project.⁵⁰ Or, more technically, basic belief-forming methods are vindicated if they are *instrumentally indispensable* to an *intrinsically indispensable* project. To illustrate, take again inference to the best explanation. Plausibly, the scientific-explanatory project is intrinsically indispensable in the sense that we cannot rationally opt out of it.⁵¹ Further, inference to the best explanation seems instrumentally indispensable to the scientific-explanatory project: given our constitution, we can only hope to make the world intelligible if this rule of inference is reasonably reliable.⁵² After all, as Enoch puts it, ‘...if not even IBE works, all is lost’ (2012: 60). So, we have a pragmatic vindication of inference to the best explanation as a basic belief-forming method.⁵³ And

⁴⁸ See Worsnip 2016: 227.

⁴⁹ See Boghossian 2000: 239 and Peacocke 2003. Of course, the vindication in question won’t be epistemic justification, since such justification comes to an end at our basic belief-forming methods.

⁵⁰ For more on how to understand the relevant pragmatic modality, see Enoch 2012: 63.

⁵¹ This non-optionality is *normative*: it involves a project we have strong reason or rationally out not to disengage from. As Enoch (2012: 84, fn. 33) points out, that only begs the question against the normative error theorist, for the premise is the normative proposition itself, not the robust realist understanding of it. Also, the normativity of the non-optionality is important: it means that we cannot discard the (e.g. explanatory-scientific) project – and must thus buy into the indispensable method (e.g. inference to the best explanation).

⁵² More precisely, ‘something is instrumentally indispensable for a project [...] just in case it cannot be eliminated without undermining [...] whatever reason we had to engage in that project in the first place’ (Enoch 2012: 69).

⁵³ But why think that basic belief-forming methods, vindicated thus, are *reliable*? According to Enoch (2012: 66), we have *evolutionary* reason to think that they are reliable, as he discusses further in Ch.7 of his book. However, we shall revisit the relationship between justification and truth below, when rehearsing Plunkett & McPherson’s (2015) criticism.

since that method generates ontological commitments to scientific entities such as electrons, we are justified in believing in their existence.⁵⁴ Given the plausibility of a similar vindicatory story about perception, memory or modus ponens, we have arrived at the first premise of the argument above: if something is instrumentally indispensable to an intrinsically indispensable project, then we are epistemically justified (for that very reason) in believing that that thing exists.

Importantly, the vindicatory account of basic belief-forming methods that the first premise articulates proves crucial to Enoch's ambition of mounting his deliberative indispensability argument for the existence of robust normative truths. After all, given the truth of the first premise, he now only needs to show that its antecedent holds: that deliberation is intrinsically indispensable and that robust normative truths are instrumentally indispensable to it. So, can he? The next two premises try to establish exactly that.

The second premise states that the deliberative project is intrinsically indispensable or rationally non-optional. That strikes me as fairly uncontroversial. After all, as Enoch points out, it might be that we are, as human beings, essentially or fundamentally deliberative creatures, given that we rely on practical deliberation to successfully navigate the world. Alternatively, it might be that deliberation can be opted out of, yet not rationally. Further, there does not seem to be a reason for thinking that the explanatory-scientific project must somehow be privileged over the deliberative one when it comes to intrinsic indispensability.⁵⁵ Just as we rationally ought not disengage from explaining the world within and around us, we rationally ought not stop deliberating about what to do.

The third premise amounts to the claim that robust normative truths are instrumentally indispensable to the deliberative project. In other words, robust normative truths cannot be eliminated without undermining the project of practical deliberation. In support of that, Enoch tells us to pay attention to the *phenomenology* of practical deliberation. Suppose you deliberate about what future career to pursue: you could either write a PhD thesis in philosophy – or take up a job at a big consulting firm in the City. What should you do? According to Enoch, when trying to make up your mind, 'it *feels* like trying to

⁵⁴ The analogue inference from vindicated basic method to existence of entities doesn't seem to hold in the deliberative case, given that inference to the presupposition of deliberation doesn't generate ontological commitment. See Cline 2016 for discussion below.

⁵⁵ As Enoch (2012: fn. 52) points out, the deliberative project may even be privileged over the explanatory one, given that explanation involves evaluation. For more, see Sayre-McCord 1988: 277-81.

make the *right* choice' (2012: 72; emphasis mine), not just an *arbitrary* one.⁵⁶ The phenomenology of deliberation, not unlike the one of finding the answer to a factual question, seems to involve discovery, not creation.⁵⁷ Deliberation seems to be about attempting to eliminate arbitrariness by discovering robust normative reasons. So, deliberation commits us to the existence of robust normative reasons relevant to deliberation.⁵⁸ Thus, it looks like robust normative reasons and truths are instrumentally indispensable to practical deliberation.⁵⁹

Together, these premises imply that we are justified in believing in the existence of robust normative truths. After all, being instrumentally indispensable to an intrinsically indispensable or rationally non-optional project confers justification. Practical deliberation has all the hallmarks of a rationally non-optional project, while robust normative reasons or truths appear instrumentally indispensable to practical deliberation. So, our belief in the existence of robust normative truths must be justified on grounds of deliberative indispensability. That, though, has important ramifications for the evolutionary debunking argument put forth in §2 above. After all, if Enoch is right, the robust moral realist has an additional, non-abductive reason to hold on to her moral belief in robust moral facts. The crucial third premise of the Harman-style evolutionary challenge seems to have been proven false – and it therefore doesn't pose a legitimate epistemological threat to robust moral realism.

5.2. Why The Deliberative Indispensability Argument Fails

How plausible is Enoch's deliberative indispensability argument? While few take issue with the rational non-optimality of practical deliberation, the other two premises have faced severe criticism. After all, it is far from uncontroversial whether indispensability confers justification (and/or generates ontological commitment) and whether practical deliberation really commits us to robustly normative reasons.

⁵⁶ Sometimes, of course, we merely make an arbitrary choice or *pick*. That, though, differs from deliberation, as Enoch (2012: 73) rightly notes: it doesn't involve a commitment to reasons – and can be done even in the believed absence of them.

⁵⁷ See Nagel 1986: 149.

⁵⁸ That doesn't mean that we must hold an explicit belief in that commitment, just as we don't hold an explicit belief in ontological commitment when employing IBE. See Enoch (2012: 75) for that point.

⁵⁹ But even if we concede that deliberation commits us to the existence of normative reasons based on phenomenological grounds, why think that those reasons are robustly normative? Enoch admits that the success of his deliberative indispensability argument depends on rejecting meta-ethical alternatives, which he motivates in various ways in other parts of his book. Still, we shall revisit that question below when discussing Lenman's (2014) objection to the third premise of the deliberative indispensability argument.

David Plunkett and Tristram McPherson (2015) have recently raised an epistemological worry for the first premise of Enoch's deliberative indispensability argument. Recall that this premise articulates a *pragmatic* vindication of the basic sources of epistemic justification. More precisely, it says that '...[o]ne complete ground for the fact that something is a source of basic epistemic justification is the fact that treating it as such a basic source is instrumentally indispensable to an intrinsically indispensable project [such as practical deliberation]' (Plunkett & McPherson 2015: 110). Initially, this pragmatic vindication looks attractive. After all, it correctly classifies intuitively basic sources (e.g. perception, memory, inference to the best explanation). Further, it explains why epistemic normativity is substantive (namely by grounding it in practical normativity). However, as Plunkett and McPherson argue, this pragmatic vindication suffers from a fatal shortcoming. In particular, it conflicts with a distinctive feature of epistemic justification, namely its *truth bias* or – what they call – *Truth-Directedness*: 'the sources of basic epistemic justification have the content that they do (in part) because of some positive connection between those sources and the truth of the beliefs that they govern' (114).⁶⁰ The pragmatic vindication violates this constraint because '...the fact that the belief that p is indispensable to our deliberative projects bears no positive relationship to the truth of p' (121).

To illustrate that, they develop several counterexamples to the pragmatic account of epistemic basicness. For example, consider Sparky, an artificial intelligence that is identical to Sally, an ordinary human being, in every respect other than one: unlike Sally, Sparky cannot deliberate about practical matters (and thus also lacks any respective introspective or mnemonic beliefs). Intuitively, Sparky seems almost always justified *to the same extent* as Sally in believing the very same propositions, including moral or more generally normative ones. But Enoch's account cannot allow for that: since Sparky lacks the project of practical deliberation, '...Sally has some defeasible epistemic justification for believing that there are ethical facts that Sparky lacks' (118). That, though, seems odd. And it seems odd exactly because whatever Sally has, yet Sparky lacks, doesn't have anything to do with the (moral) truth. And if it doesn't have anything to do with the truth, how could Sally be more justified than Sparky? Or, put differently, what explains the counter-intuitive

⁶⁰ Plunkett & McPherson (2015: 114-17) hasten to qualify *Truth-Directedness* in three ways. First, they stress that it is neutral with respect to accounts of truth. So, it is compatible with correspondence, deflationary or epistemic notions of truth. Second, *Truth-Directedness* is not a constraint on the basic sources of epistemic justification themselves. Rather, it is a constraint on *theories* that purport to explain why something is a basic source of epistemic justification. Third, the positive connection to truth should be understood ecumenically. It might involve modal notions such as safety or sensitivity, constitutive goals, or else. Further, it is also distinct, yet compatible, with epistemic constitutivism, epistemic teleology or pragmatic encroachment. The indeterminacy is dialectically crucial: 'we take *commitment* to something like Truth-Directedness to be close to common ground in many parts of epistemology' (116).

diagnosis that Enoch's account provides for cases such as this one is that it violates *Truth-Directedness*: it says that levels of justification may vary irrespective of their connection to truth. So, it looks like Enoch's pragmatic vindication fails to capture a distinctive feature of epistemic justification.⁶¹ And since it is deficient in that crucial respect, it cannot be a plausible vindicatory account of basic belief-forming methods. By implication, this story cannot underwrite the first premise of Enoch's argument – or ground deliberative dispensability as a basic source of epistemic justification.⁶²

The third premise has also attracted critical attention. James Lenman (2014) has argued that – *contra* Enoch – robust normative truths are not instrumentally indispensable to practical deliberation.⁶³ Recall that Enoch offers a *phenomenological* argument: in deliberation, it *feels* or *seems* like we are trying to make the *right* choice (e.g. about what career to pursue), based on normative reasons – as opposed to merely picking arbitrarily (e.g. between two qualitatively identical pens at the stationary store). Therefore, we must be committed to the existence of robust normative reasons. However, Lenman argues that this inference is invalid. To see that, he asks us to imagine *schemeliberation*: 'the sort of practical thinking where we try to do well with respect to standards of evaluation set by all those deeply entrenched concerns and aspirations we humans happen to have, aspirations for pleasure, flourishing, love, prosperity, peace, justice, security, country air and all that other stuff we happen to like, but remain cheerfully indifferent to normative truths as Enoch conceives them' (Lenman 2014: 839). According to Lenman, two aspects of *schemeliberation* stand out. First, we can do better or worse (or: succeed or fail) at *schemeliberating*. So, *schemeliberation* is a standard-regulated intellectual practice.⁶⁴ Second (and more importantly), the difference between deliberation and *schemeliberation* isn't like the difference between choosing, governed by the recognition of normative reasons, and arbitrary picking. After all, the *phenomenology* of *schemeliberation* isn't like the phenomenology of arbitrary picking, but just like the one of deliberation. Or,

⁶¹ In case you remain unconvinced by the Sally & Sparky counterexample, check out the other two that Plunkett & McPherson (2015: 120) present, namely Declan and Marjorie.

⁶² Plunkett & McPherson (2015: 124-29) also make the case that their worry *generalizes* to pragmatic accounts other than Enoch's that aim to vindicate deliberative indispensability. Since deliberative indispensability is unconnected to truth, anyone defending it as a basic source of epistemic justification faces the following *dilemma*: either try to vindicate deliberative indispensability as a basic source of justification and violate Truth-Directedness – or develop a vindicating account that is compatible with Truth-Directedness (e.g. Goldman's (2008) reliabilism), yet won't vindicate deliberative indispensability as a basic source of epistemic justification.

⁶³ For a similar argument, see also Olinder & Bjoernsson 2016: 105-8.

⁶⁴ In response, Enoch might press the Kantian line of asking where normativity gets into the mere desires. However, for Lenman, the space of reasons is just (and unmysteriously) the space of one's own, first-person (self- and other-regarding) desires – and normativity is just the '...force that attaches to the motivation we feel from unalienated desires, desires with which we stably reflectively identify' (2014: 840).

schmeliberating doesn't *feel* arbitrary like picking, but rather just like deliberation. As a result, Enoch's phenomenological argument fails: the phenomenology of deliberation isn't rich enough to license the inference to the existence of *robustly* normative reasons. Rather, the phenomenology of deliberation seems to support the existence of normative reasons *simpliciter*. For all we know, however, those reasons might as well reduce to complex human desires or those 'deeply entrenched concerns and aspirations we humans happen to have'.⁶⁵ So, robustly normative truths are not instrumentally indispensable to practical deliberation.⁶⁶

But even if the first and third premise escape these worries, some have recently questioned the argument's *validity*. Brendan Cline (2016) has argued that deliberative indispensability cannot be a legitimate, independent guide to (normative) ontology. That is so even if – *pace* Plunkett & McPherson (2015) – deliberative indispensability is vindicated as a basic source of justification and – *pace* Lenman (2014) – robust normative truths are instrumentally indispensable to practical deliberation.⁶⁷ After all, even if some entity (such as robustly normative truth) is deliberatively indispensable, it is still an open question whether that entity *exists*. That question, though, need be settled by additional *explanatory* work, given that it doesn't remain open if an entity is indispensable to an actual explanation of a phenomenon. So, deliberative indispensability cannot independently ground ontological commitments – and thus secure the conclusion of Enoch's argument.

Cline (2016) formulates his objection in two steps. First, he reminds us *how* inference to the best explanation (or, henceforth: IBE) generates ontological commitment. IBE treats explanatory considerations as a guide to inference: it helps us spot the actual explanation

⁶⁵ Lenman is a (Humean) naturalist moral realist. As a result, he is primarily interested in showing that Enoch's phenomenological argument doesn't rule out normative reasons understood as complex human desires. But I think that Lenman's point doesn't just exonerate naturalist moral realism, but any meta-ethical view committed to the existence of normative reasons. After all, his point is simply that the phenomenology of deliberation (unlike the one of arbitrary picking) seems to commit us to the existence of normative reasons – but (*contra* Enoch) doesn't put any further constraints on their nature. So, his point is also good news for, say, moral constructivists or dispositionalists or conventionalists, for whom normative reasons derive from moral attitudes, dispositions or institutional norms.

⁶⁶ For those unconvinced by his point about the shared phenomenology of deliberation and schmeliberation, Lenman also offers a thought-experiment in support of this conclusion. He surmises that if Enoch renounced his belief in robust normative realism, he wouldn't stop deliberating and '[n]otwithstanding the loss of his belief in robustly real normative truths, he still cares about his own prosperity and happiness, about the wellbeing of his family and friends and the continued flourishing of his relationships with them, about the success of his various projects (finishing that book, building that outhouse), about living a whole life that will bear his survey in the light of such ideals of the person as he stably reflectively endorses (all this of course just a load more "mere" desires)' (841).

⁶⁷ Bjoernsson & Olinder 2016: 106 develop a similar worry.

by making us collect plausible *potential* explanations and examine them for their indicators of *accuracy* (such as coherence, predictive power, simplicity, generality, precision, etc.).⁶⁸ So, IBE helps us move from potential to actual explanations.⁶⁹ Importantly, for scientific realists, only (approximately) *true* hypotheses *actually* explain anything.⁷⁰ So, if an actual explanation is committed to a particular kind of entity, that entity must exist (because the hypothesis wouldn't be true otherwise). Actual explanations, then, involve true existential claims. IBE thus generates ontological commitment by linking potential to actual explanations which, in turn, involve true existential claims. Or, put differently, 'this link between actual explanations and reality is what makes explanatory indispensability well suited to play a privileged role in ontological construction' (Cline 2016: 3245).

Second, Cline shows why, unlike IBE, deliberative indispensability cannot generate ontological commitment. More precisely, suppose there is an inference rule called *Inference to the Commitments of Deliberation* (ICD): 'we should try to discern and then embrace entities that are deliberately indispensable' (3245). However, that rule – unlike IBE – is *not* ontologically committing because, importantly, something can be deliberately indispensable, yet *false*. To illustrate, consider free will. Many hold that libertarian free will is deliberately indispensable, given that 'we cannot deliberate if we just sit back and try to see what happens' (3246).⁷¹ Still, even if libertarians are right about our sense of freedom and its incompatibility with determinism, it remains an *open question* whether determinism is true and whether we are free in the relevant sense. Put differently, it is possible that deliberation simply commits us to falsehoods.⁷² Or, we can coherently acknowledge the deliberative indispensability of free will and simultaneously doubt it.⁷³ So, even if we correctly characterize the indispensable commitments to deliberation, more work is needed to get to the bottom of things. And, given the points above, that work is plausibly *explanatory*, not deliberative: we should look for explanatory considerations to

⁶⁸ See Harman 1965. As Cline points out, given the notion of actual explanation below (according to which only *true* hypotheses are actual explanations), the label 'IBE' is misleading: it cannot be IBAE, since that would be trivial and presupposes that we already know which explanation is true. Rather, following Lipton (2004: 59), we should introduce potential explanation (i.e. a claim that would, if true, provide an actual explanation of some phenomenon) – and read IBE as IBPE. Nothing in what follows hinges on that, though.

⁶⁹ Of course, as Cline reminds us, there is still a risk of misstepping (because the actual explanation might rate poorly or won't be on our list of potential explanations) – inviting IBE-scepticism. Enoch's story about the justification of IBE aims to address sceptical worries about IBE, but also leads to his overlooking that, in successful cases, the best-looking potential explanation will be the actual explanation – which is ontologically committing.

⁷⁰ See Lipton 1991: 129, 174 and Field 1989: 15.

⁷¹ See, for example, Searle 2001: 13.

⁷² Searle 2001: 17, 67, 71, 278f.

⁷³ See Nagel 1986: 117.

examine the commitments of deliberation. In Cline's words: 'Thus ICD, even when successful, is in principle incapable of settling ontological disputes. This is why ICD is not fit to serve as an independent guide in our quest to work out what there is' (3247). In short, deliberative indispensability is not a legitimate guide to (normative) ontology.⁷⁴ But if it is not, then the conclusion of Enoch's argument doesn't follow, even if it is true that deliberative indispensability is a basic source of epistemic justification, deliberation is rationally non-optional and robust normative reasons are indispensable to it. It still won't be true that robust normative reasons *exist*.

Where does that leave Enoch's deliberative indispensability argument? The worry arising from the truth-bias of epistemic justification (Plunkett & McPherson), the instrumental dispensability of robust normative truths to practical deliberation (Lenman), and the failure of deliberative indispensability to generate ontological commitment (Cline) all seem to render the argument implausible. But if Enoch's argument is implausible, an appeal to deliberative indispensability won't help robust moral realists in response to the Harman-style evolutionary debunking argument developed in §2. After all, deliberative indispensability doesn't give us non-abductive reasons to believe in the existence of robust moral facts. By implication, the third premise of the debunking argument still holds true. Robust moral realists must look elsewhere to avoid moral scepticism.

6. Conclusion

In this chapter, I offered a systematic, comprehensive defence of an evolutionary debunking argument inspired by the epistemic principle that, for our beliefs to be justified, they must be best explained by the facts they are about. According to the more specific argument, since robust moral facts are dispensable to the best (namely: evolutionary) explanation of our moral beliefs and there are no additional, non-abductive reasons to maintain them, those beliefs lose their justification. As my discussion showed, this argument proves compelling: on reflection, it withstands the most powerful challenges to the truth of its premises.

To begin with, I argued that – *contra* Mogensen (2015, 2016) – evolutionary and moral explanations of moral belief are not compatible (and that, even if they were, it wouldn't matter, given that the latter remain metaphysically mysterious). So, evolution *does* best explain moral belief. Further, I made the case that – *contra* Cuneo (2007) – robust epistemic facts aren't *objectionably* explanatorily dispensable and thus cannot serve as

⁷⁴ Cline stresses – rightly in my mind – that deliberative indispensability might still be a source of preliminary evidence that we can subject to scrutiny (comparable to *perceptual* indispensability).

companions for robust moral facts. So, robust moral facts remain metaphysically problematic, which defeats our belief in them, unless we have some other, non-abductive reasons for their existence. Finally, I rejected Enoch's (2012, 2016) proposal that the deliberative indispensability of robust moral facts gives us exactly such reasons.

Where does that leave us? This chapter concludes the second step of my investigation. After two failed attempts in Chapters 2 and 3, we finally hit on a compelling evolutionary debunking challenge to robust moral realism, based on an independently plausible epistemic principle. That, of course, is good news for evolutionary debunkers. In contrast, robust moral realists seem to be in deep epistemological trouble. If they cannot find another way of faulting the argument developed in this chapter, they are saddled with implausible moral scepticism. Chapter 5 amounts to the third and final step of my investigation into evolutionary debunking. It focuses on moral intuitionism, the most popular, best-developed, non-abductive moral epistemology available to robust moral realists. More precisely, Chapter 5 addresses the following question: does moral intuitionism have the conceptual tools to tackle the worry about epistemic luck raised in Chapter 2 – as well as offer additional, non-abductive reasons for robust moral belief (thus staving off the sceptical worry developed in this chapter)?

CHAPTER 5: EVOLUTION AND THE POSSIBILITY OF ROBUSTLY MORAL KNOWLEDGE

1. Introduction

While Chapter 1 set the stage, Chapters 2 to 4 took the second step in my investigation. In them, I developed and critically assessed evolutionary debunking arguments based on independently plausible epistemic principles, informed by cutting-edge debates in contemporary epistemology. My discussion showed that, upon closer inspection, robust moral realism faces two distinct epistemic worries. The first, developed in Chapter 2, is loosely inspired by evolutionary debunking arguments and centers on knowledge-undermining epistemic luck. Meanwhile, the second, as discussed in Chapter 4, is directly raised by evolutionary considerations, focusing on whether robust moral realists have compelling non-explanatory reasons for maintaining their moral beliefs.¹

Chapter 5 amounts to the third and final step. Its overall aim is to argue that, while robust moral realism can tackle the worry about epistemic luck, it cannot offer compelling non-explanatory reasons for robustly moral belief. As a result, the evolutionary debunking argument developed in Chapter 4 succeeds, saddling robust moral realism with implausible moral scepticism: we cannot justifiably believe robustly moral facts. However, achieving that overall aim necessitates a slight shift in *methodology*. While Chapters 2 to 4 remained largely neutral with respect to the correct first-order moral epistemology for robust moral realism, those minimal assumptions won't suffice for figuring whether (and, if so, how) robust moral realism can tackle the residual worries raised (or inspired) by evolutionary debunking arguments.² Rather, doing so requires an idea of how we form true, justified beliefs about and acquire knowledge of moral facts. Only then can we meaningfully gauge whether robust moral realism survives evolutionary debunking. Now, contemporary moral intuitionism is the most popular and best-developed moral

¹ Recall that, upon closer inspection, the worry about irrelevant evolutionary influence on robustly moral belief discussed in Chapter 3 did not prove compelling. So, robust moral realists don't need to address it.

² In fact, all I assumed in Chapters 2 to 4 was that robust moral realists are non-sceptics: they hold that we can have knowledge of robust moral facts, at least under epistemically favorable conditions. Assuming that little was enough to assess whether the three epistemic principles outlined in Chapter 1 proved generally plausible and, more specifically, whether they combined with evolutionary explanations to undermine robustly moral knowledge.

epistemology available to robust moral realism.³ It therefore offers the best chance of dealing with the worries about epistemic luck and the putative lack of non-explanatory reasons for robustly moral belief. And if even moral intuitionism failed to do so, robust moral realism would be in serious trouble. So, my inquiry in this chapter henceforth focuses on moral intuitionism.⁴

The overall aim of this chapter breaks down into two more specific goals. First, I intend to argue that, on reflection, moral intuitionism isn't vulnerable to the worry about epistemic luck developed in Chapter 2. However, examining closely why that is helps uncover the most compelling version of moral intuitionism. According to that view, you know (an intuitive moral) *p* if and only if you (truly) believe that *p*, justified on the basis of a veridical intellectual seeming that *p* that results from adequately understanding *p*. Second, I aim to show that even this version of moral intuitionism cannot stave off the worry raised by best evolutionary explanation argument in Chapter 4. As I argue, moral intuitionism, by itself, cannot give us additional, non-explanatory reasons for robustly moral beliefs. Rather, the view must be supplemented with a third-factor explanation of moral reliability. To avoid begging the question against the evolutionary debunker, however, any such explanation can only rely on formal moral intuitions: intuitions that impose formal constraints on ethical viewpoints, yet without evaluating anything positively or negatively. I argue, however, that these intuitions cannot underwrite many substantive moral beliefs, even when combined with rational reflection. Therefore, moral intuitionism cannot give us additional, non-explanatory reasons for anything beyond our beliefs about the *form* of morality. As a result, the best evolutionary explanation argument in Chapter 4 succeeds. At last, robust moral realism is saddled with unacceptable moral scepticism: we cannot justifiably believe substantive facts about morality, robustly construed. While evolution doesn't imperil the mere possibility of robustly moral knowledge, it undermines any claim to know what robust morality is substantively like.

Here is my plan. §2 sets the stage by introducing moral intuitionism in broad strokes. §3 develops the most compelling version of this view in response to the worry about epistemic luck. §4 argues that moral intuitionism cannot accommodate the best evolutionary explanation argument against robust moral realism. In more detail, it makes the case that the view requires a third-factor explanation of moral reliability (§4.1 & 4.2), that any such

³ With respect to popularity: Shafer-Landau (2003), Huemer (2005) and Parfit (2013) are all, *inter alia*, prominent contemporary robust moral realists and moral intuitionists. That moral intuitionism is well-developed as a view in moral epistemology will become apparent in §2 and §3 below.

⁴ Importantly, I henceforth don't assume that moral intuitionism is *true* or *correct*. Rather, I simply (and charitably) assume that it is the best option available to robust moral realism. But, as I shall argue at length in this chapter, I don't think even that option pans out.

explanation must rely on formal moral intuitions only to avoid begging the question (§4.3), and that such intuitions cannot underwrite belief about the substance of robust morality (§4.4). §5 summarizes the discussion of Chapters 2 to 5 and develops some further insights into evolutionary debunking arguments and theorizing in moral epistemology more generally.

2. A Sketch of Moral Intuitionism

In this section, my aim is to introduce moral intuitionism in broad strokes, thus providing the background necessary for developing the most compelling version of this view in response to the worry about epistemic luck in §3. I start by briefly characterizing and motivating moral intuitionism in general terms (§2.1.), before presenting two more specific versions of it, based on diverging conceptions of moral intuitions (§2.2.).

2.1. What It Is and Why It Seems Attractive

Moral intuitionism, as defended by Robert Audi (1999), Michael Huemer (2005), Russ Shafer-Landau (2003) and Philip Stratton-Lake (2016), is the epistemological view that moral knowledge is true moral belief, justified on the basis of moral intuition.⁵ Or, more precisely, you know that *p* if you have a true belief that *p*, justified by a corresponding moral intuition that *p*. For instance, consider the true moral proposition that honesty is a virtue.⁶ According to moral intuitionism, *intuiting* or having the intuition that honesty is a virtue is sufficient for being *justified* in believing and therefore *knowing* that honesty is a virtue. So, moral intuitionism aims to provide a partial *analysis* of moral knowledge: it is in the business of stating sufficient conditions on it. By doing so, it provides us with a partial answer to one of the central questions of moral epistemology, namely the question of what moral knowledge is.⁷

Before proceeding, two qualifications are worth making. First, the conditional above only concerns basic moral propositions: those that can be known by moral intuition. In other

⁵ My exposition and subsequent dialectic focus on *contemporary* moral intuitionism exclusively. For more on its connection to classical forms such as Price's (1758/1969) view, see Stratton-Lake 2014: §1.

⁶ This example is due to Huemer 2005: 102. See *ibid.* and Shafer-Landau 2003: 248 for other allegedly true and intuitively compelling moral propositions. All these propositions should be understood *ceteris paribus*.

⁷ Why only a *partial* analysis of moral knowledge? Because some moral intuitionists such as Shafer-Landau (2003) allow for *non-intuitive* moral knowledge. So, for them, intuition cannot also be *necessary* for moral knowledge. A full analysis, however, would feature both necessary *and* sufficient conditions. For my purposes, the complication doesn't matter too much. The epistemic luck problem discussed in §3 below primarily concerns the sufficiency condition – and the anti-luck condition I advocate in response is necessary only for *intuitive* moral knowledge. So, I am happy to allow for non-intuitive moral knowledge.

words, moral intuitionism does not state that we can know *any* moral proposition by way of moral intuition. For instance, as Huemer (2005: 102) points out, we could not know by moral intuition whether abortion is morally permissible or not. Rather, we would know that based on inference from other moral and non-moral beliefs (e.g. about moral criteria for personhood, about whether a human fetus meets these criteria). Of course, it is a matter of dispute whether and, if so, which moral propositions can be known by moral intuition. But, for the purposes of my argument in this chapter, I shall simply grant that some can. I will have more to say about the nature of those basic moral propositions in §4 below.

Second, intuiting or having an intuition that *p*, even where *p* is a moral truth, need not always mean being *justified* in believing *p* – and therefore *knowing* *p*. Rather, the moral intuition that *p* also has to remain *undefeated*. For instance, if you were in possession of better evidence that not-*p* or suitably good evidence that your intuition that *p* is misleading or formed unreliably, you would no longer be justified in believing *p*.⁸ A more precise formulation of moral intuitionism would acknowledge that. As we shall see, this complication proves irrelevant for our discussion of the epistemic luck problem in §3 below, given that scenarios used for illustrating it don't feature any such defeaters.⁹ However, the no-defeater clause will take center stage in our discussion of the best evolutionary explanation argument in §4 below. If successful, that very argument forces robust moral realists into moral scepticism by articulating an undermining defeater for their moral beliefs, showing that none of their (atomic) moral intuitions ever remains undefeated.

Moral intuitionism is an *attractive* positive, non-sceptical moral epistemology for robust moral realists. To begin with, it nicely fits with their moral metaphysics. One of the core metaphysical intuitions animating robust moral realism is that the normative is *sui generis* or just 'too different' to be reduced to or identified with anything in the world described by the natural sciences.¹⁰ But if the normative is metaphysically different in that way, it seems plausible that we get to *know* about normativity a similarly epistemically different way. More specifically, we don't access normative reality with the epistemic tools used to gain knowledge of the natural world, such as perception or scientific theorizing. Rather,

⁸ See, for instance, Audi 1999: 219f. and Huemer 2005: §5.3.

⁹ Or, more precisely, there is a potential defeater in scenarios such as **Brain Lesion**, namely the unreliability of moral intuition, but – by stipulation – the subject in my case lacks evidence for it. As a result, her belief remains justified. So, the complication proves irrelevant in the end.

¹⁰ See Enoch 2012: Ch. 5.

we do so using moral intuitions. In that manner, moral intuitionism, as a moral epistemology, fits nicely with the metaphysics of robust moral realism.

Further, contemporary moral intuitionism simultaneously manages to demystify moral intuitions. They are not the product of an ad-hoc, obscure faculty of moral intuition, but a kind of mental state essential to how we acquire *apriori* knowledge more generally.¹¹ After all, it seems plausible that we come to know at least some logical (e.g. the Law of the Excluded Middle), mathematical (e.g. that $2 + 2 = 4$), modal (e.g. that necessity entails possibility) and philosophical (e.g. that true, justified belief isn't knowledge) facts by having corresponding logical, mathematical, modal and philosophical intuitions. Intuitive moral knowledge is therefore nothing mysterious, but a species of intuitive *apriori* knowledge.

2.2. The Nature of Moral Intuitions

We can distinguish two dominant strands of contemporary moral intuitionism, depending on what *moral intuitions* are (or what it means to *intuit* a moral proposition). According to the first account, defended by Audi (1999) and Shafer-Landau (2003), moral intuitions are *beliefs* in *self-evident* moral propositions, where self-evident propositions are propositions that can be known on the basis of *adequate understanding* alone.¹² For instance, according to Shafer-Landau (2003: 248), it is a self-evident (or self-evidently true) that pain is bad or that one should not punish the innocent.¹³ So, on this account, having a moral intuition means holding a belief about a moral truth that one adequately understands.

But what exactly is *adequate understanding*, according to Audi or Shafer-Landau? Neither gives us a full analysis. Still, Audi (1999: 207f.) helpfully contrasts adequate understanding with a variety of *defective* understandings such as mistaken, insufficient, distorted and clouded understanding. For instance, thinking that there is a minimal level of epistemic confidence appropriate to knowledge would mean distortedly understanding (the sentence expressing) the proposition that knowledge entails true belief. After all, even though the former is compatible with the latter, it is not entailed by it. Further, he distinguishes adequate understanding from mere semantic understanding: '[a]dequacy here implies not only seeing what the proposition says but also being able to apply it to

¹¹ See Audi 1999 and Huemer 2005: Ch. 5.

¹² It's important to keep moral intuitions and self-evident propositions apart. See Stratton-Lake 2014: 4f.

¹³ What is self-evidence? For Audi, self-evidence is '...a kind of manifest truth of a proposition in itself...' (1999: 206).

(and withhold its application from) an appropriately wide range of cases, and being able to see some of its logical implications, to distinguish it from a certain range of close relatives, and to comprehend its elements and some of their relations' (208). So, adequately understanding that, say, honesty is a virtue means more than simply understanding what the proposition means: it involves not applying the proposition in cases in which honesty is not a virtue (e.g. when talking to a friend about their dire job prospects), observing logical implications (e.g. that honesty should be cultivated because it is a virtue), not confusing it with close relatives (e.g. being trustworthy) and comprehending its elements (e.g. that virtues are positive character traits, that honesty involves not lying or cheating).

When we integrate the self-evidence account of moral intuitions into the general characterization offered above, we get the *self-evidence version of moral intuitionism*. According to it, '*p* is self-evident provided an adequate understanding of it is sufficient for being *justified* in believing it and for *knowing* it if one believes it on the basis of that understanding' (Audi 1999: 206; emphasis mine). For instance, consider again the self-evident proposition that honesty is a virtue. According to the self-evidence version, when you adequately understand that honesty is a virtue, you are *justified* in believing it (barring defeaters) – and, if you form the corresponding belief based on that understanding, you also *know* that honesty is a virtue.

Yet not all construe moral intuitions as beliefs in self-evident moral propositions. Instead, inspired by George Bealer's (1999) work on intuitions more generally, Michael Huemer (2005) and Philip Stratton-Lake (2016) take moral intuitions to be *intellectual seemings*: non-inferential, non-doxastic mental states that result from thinking about certain moral propositions.¹⁴ For instance, it might *seem* true to you that enjoyment is better than suffering, that it is unjust to punish a person for a crime she did not commit, or that it is morally impermissible to kill one healthy patient to save the lives of five terminally ill patients in need of organ transplants.¹⁵

¹⁴ Note that, for Huemer (2005: §5.1) not all seemings are intellectual: there are also perceptual, introspective and mnemonic seemings. Further, not all intellectual seemings are moral seemings. Rather, there are logical, mathematical and modal seemings, too. For instance, it might seem to you that an argument is valid or that $2 + 2 = 4$. Finally, intellectual moral seemings can differ along several dimensions: their strength, how widely they are shared, how simple or complex their contents are, etc. For more, see Huemer 2005: § 5.2.

¹⁵ For further examples, see Huemer 2005: §5.2. As mentioned above, all the intuitive moral propositions are supposed to be understood as holding true *ceteris paribus*. For examples of non-intuitive moral propositions, see Huemer 2005: *ibid.*

When we combine this intellectual seemings account with the view that true moral beliefs, justified by moral intuition, are sufficient for moral knowledge, we get the *intellectual seemings version of moral intuitionism*. According to it, if you form the true moral belief that p, justified on the basis of the intellectual seeming that p, you know that p. Or, to quote Stratton-Lake: '[Intuitive] propositions are truths such that (a) a clear intuition [or intellectual seeming] of them is sufficient justification for believing them, and (b) believing them on the basis of a clear intuition [or intellectual seeming] of them entails knowing them' (Stratton-Lake 2016: 16).¹⁶ For instance, take the true moral proposition that enjoyment is better than suffering. According to the intellectual seemings version of moral intuitionism, if you form the true moral belief that enjoyment is better than suffering, justified on the basis of the corresponding clear intellectual seeming, you *know* that enjoyment is better than suffering.

Of course, not all intellectual seemings are *veridical*. For instance, suppose that act-consequentialism is in reality our best moral theory. If so, even though it *seems* morally impermissible to kill one healthy patient to save the lives of five, it would not in fact be so. Accordingly, your corresponding moral belief would be justified, yet *false*. Still, many intellectual seemings are probably veridical: they concern moral truths. For instance, it seems plausible to say that the intellectual seeming that, all things being equal, enjoyment is better than suffering is most likely veridical. In those cases, the beliefs you form on their basis are not just justified, but amount to knowledge.

Which version of moral intuitionism should the robust moral realist endorse? In the next section, I turn towards discussing the problem of *epistemic luck*, as developed in Chapter 2. And as I shall suggest, one version of moral intuitionism is particularly well-placed to address that worry – and should thus be preferred as the most compelling one available.

3. Moral Intuitionism and Epistemic Luck

In this section, my aim is to argue that moral intuitionism isn't vulnerable to the worry about epistemic luck raised in Chapter 2. However, examining closely why that is helps uncover the most compelling version of the view. That, in turn, prepares us well for the

¹⁶ Stratton-Lake (2016: 16) calls these truths 'self-evident', despite having previously argued *against* the notion of self-evidence, which seems misleading. I therefore call them 'intuitive' truths instead. What does a *clear* moral intuition amount to, though? According to Stratton-Lake, '[b]y getting a proposition clearly in view I mean pretty much what Audi calls having an adequate understanding of it' (Stratton-Lake 2016: 13, fn.25). So, clear intellectual seemings concern (moral) propositions that we also adequately understand. For more on the – probably problematic – role of adequate understanding in Stratton-Lake's account, see §3 below.

discussion in §4 about whether moral intuitionism can handle the best evolutionary explanation argument developed in Chapter 4. I start by briefly re-introducing the problem of epistemic luck (§3.1.), before proposing an intuitionist response based on adequate understanding (§3.2.). Then, I argue that a specific form of the intellectual seemings version of the view can avail itself most readily of my response (§3.3).

3.1. The Problem of Epistemic Luck

In Chapter 2, I argued that there are widely overlooked instances of true, justified, yet *epistemically lucky* moral belief that fall short of moral knowledge. To illustrate that, I developed the following Gettier-style case:

Brain Lesion: suppose that you highly reliably intuit (and thus know) various moral truths. However, curiously enough, you have never thought about whether honesty is a virtue or not. One night, without realizing, you suffer a brain lesion, which has the following effect: you find intuitively compelling *any* moral proposition you entertain. Luckily for you, though, the first moral proposition you entertain the next morning and form a corresponding belief about is that honesty is a virtue, which is necessarily true.

Here, your belief that honesty is a virtue is true because honesty is indeed a virtue (and necessarily so). Further, your belief is justified: you have the moral intuition – which you take to be veridical or reliably formed – that honesty is a virtue.¹⁷ Still, you don't appear to *know* that honesty is a virtue. After all, you were lucky that the first moral proposition that you entertained was true. Had your moral intuition been about the false moral proposition that courage is a vice instead, you would have ended up with a false moral belief.¹⁸ So, it is a matter of luck that your belief is true.¹⁹

¹⁷ Note that the unreliability of your moral intuition (due to the brain lesion) does not *defeat* your justified moral belief that honesty is a virtue. By stipulation, you do not have any *evidence* about it in **Brain Lesion**. However, such evidence would be required, given that moral intuitionism is a form of epistemic internalism about justification. For more on moral intuitionism and epistemic internalism, see Audi 1999: 213f.

¹⁸ I don't think you could really intuit that courage is a vice (or any other moral falsehood) if moral intuitions were beliefs in self-evident propositions. After all, a self-evident proposition is a proposition that you *adequately understand*. Adequate understanding, though, is factive (at least on some conceptions): you can only adequately understand moral truths. Therefore, you could not adequately understand – and thus not intuit – that courage is a vice. More on that below.

¹⁹ In Chapter 2, I also offered a further explanation as to why epistemic luck undermines knowledge: it renders beliefs *unsafe*, i.e. in most nearby possible worlds in which you employ the same method of belief-formation, you end up with false beliefs.

What relevance does epistemic luck have for moral intuitionism? The Gettier-style case above illustrates that epistemic luck undermines knowledge: you can form a true moral belief that *p*, justified on the basis of a moral intuition that *p*, without *knowing* that *p*. So, there is something wrong with moral intuitionism as an *analysis* of moral knowledge: it is not the case that you *know* that *p* if you form the true moral belief that *p*, justified on the basis of the moral intuition that *p*. True moral belief, justified by moral intuition, is insufficient for moral knowledge. In Chapter 2, I called this challenge the *problem of epistemic luck*.

3.2. Adequate Understanding as a Potential Solution

How should moral intuitionists respond? In my view, their most promising response involves an appeal to *adequate understanding* (or conceptual competence). Recall that, following Audi's (1999) definition introduced in §2.2. above, adequately understanding a (moral) proposition implies '...not only seeing what the proposition says but also being able to apply it to (and withhold its application from) an appropriately wide range of cases, and being able to see some of its logical implications, to distinguish it from a certain range of close relatives, and to comprehend its elements and some of their relations' (208). However, it is questionable whether you adequately understand that honesty is a virtue in **Brain Lesion** above.²⁰ Even though it seems plausible that you understand what the proposition says (namely that honesty is a virtue), the brain lesion appears to undermine your ability to both distinguish it from close relatives and comprehend its elements and some of their relations. More specifically, due to the brain lesion, you could easily have had a moral intuition about the false moral proposition that, say, insensitivity is a virtue (and thereby confused honesty with a close relative) or the moral falsehood that pride is a virtue (and thereby failed to comprehend what a virtue is). So, it is questionable whether you adequately understand that honesty is a virtue in **Brain Lesion**. If you do not and adequate understanding is a necessary condition on intuitive moral knowledge, though, you don't know that honesty is a virtue in **Brain Lesion**. So, the case cannot serve as a counterexample to moral intuitionism. After all, the verdict the view yield aligns with our intuition about that case, namely that you lack moral knowledge. On reflection, moral intuitionism therefore doesn't seem vulnerable to the problem of epistemic luck.

²⁰ Riaz (2015: 115-21) seems to share my intuition with respect to *moral* understanding (of contingent moral proposition) and veritic luck: she argues – *pace* Hills (2009) – that we don't *morally* understand *p* in Gettier-style cases where belief in *p* is veritically lucky. However, she takes that to be evidence that moral understanding is a *species* of contingent moral knowledge. Of course, I don't think we should draw the analogous conclusion regarding *apriori* moral knowledge. Adequate understanding is a necessary condition, not species, of *apriori*, intuitive moral knowledge.

But the discussion doesn't end here. Suppose that it is indeed plausible that you don't adequately understand that honesty is a virtue in **Brain Lesion**. Still, a further question remains: which version of moral intuitionism can most readily avail itself of this response? Now, you might think that the answer is obvious (no pun intended): it must be the self-evidence version of the view, according to which adequately understanding that *p* is central to having a moral intuition that *p*. But that, as I shall argue now, would be too hasty. Instead, to answer that question properly, we must figure out in what sense adequate understanding might be *most plausibly* construed as necessary for intuitive moral knowledge. And there are at least two ways of integrating adequate understanding into the architecture of the intellectual seemings version of moral intuitionism as well. More specifically, adequate understanding could either serve as a condition on the justificatory force of intellectual seemings – or as a wholly separate anti-luck condition on moral knowledge. Upon closer examination, it is this third and final candidate, namely adequate understanding as an anti-luck condition, that proves most compelling – and thus helps us adjudicate the dispute between different versions of moral intuitionism in favor of the intellectual seemings view. Adjudicating that dispute, meanwhile, is important: ultimately, we are interested in developing the strongest possible version of moral intuitionism before confronting it with the best evolutionary explanation argument in §4 below. That way, we respect charity and can be sure that, if even that version fails, there is little hope for moral intuitionism – and robust moral realism.

3.3 What Role Should Adequate Understanding Play Within Moral Intuitionism?

According to the first proposal, adequately understanding that *p* is central to what it means to have a moral intuition and thus being justified in believing that *p* in the first place. This seems to be what Audi (1999) and Shafer-Landau (2003) have in mind. After all, for them, moral intuitions are *beliefs* in *self-evident* moral propositions, where self-evident propositions are propositions that can be known on the basis of *adequate understanding* alone. Further, they both hold that adequately understanding that *p* is sufficient for being justified in believing *p*. But if you don't adequately understand that honesty is a virtue, as is plausibly the case in **Brain Lesion**, you cannot justifiably believe it either. And if moral knowledge requires justification, you don't know in **Brain Lesion**. In that manner, self-evidence theorists such as Audi and Shafer-Landau could solve the problem of epistemic luck.

However, adequate understanding isn't a plausible condition on justification. As Philip Stratton-Lake (2016: 4-9) has recently argued while criticising the self-evidence view, *adequately understanding* a self-evident moral *p* cannot justify the corresponding moral

belief in *p*. For instance, ‘because I understand that pleasure is better than agony’ seems an odd response to the question ‘Why do you believe that pleasure is better than agony?’. Or, it seems odd to *justify* the belief that pleasure is better than agony by appeal to adequately understanding it. Where does this oddity come from? According to Stratton-Lake, epistemic justification for beliefs in synthetic propositions requires *evidence*, where evidence is ‘...something that raises the (epistemic) probability of the truth of the proposition for which it is evidence’ (Stratton-Lake 2016: 5). But our understanding of a synthetic *apriori* proposition does not provide evidence for its truth.²¹ So, understanding a synthetic (moral) proposition cannot justify believing it.

It is important to stress that this issue runs deep and cannot be fixed easily. For instance, suppose self-evidence theorists just claimed that moral intuitions about *apriori* synthetic moral propositions – instead of adequate understanding – provide evidence for their truth. But that doesn’t address the issue. After all, as Stratton-Lake points out, ‘...our intuition of a self-evident proposition cannot justify us in believing that proposition [...] because an intuition is, on [Audi’s] account, a certain type of belief, and my belief that *p* cannot justify my belief that *p*.’ (Stratton-Lake 2016: 9). So, it remains unclear what justifies moral beliefs on the self-evidence version of moral intuitionism. That, though, proves highly problematic for the view, especially given the key role the justifying force of adequate understanding is supposed to play in addressing the worry about epistemic

²¹ Apart from some self-referential cases such as ‘I understand this proposition’ (cf. Stratton-Lake 2016: 5). But these cases don’t bear on self-evident *moral* propositions.

luck.²² So, it looks like the view faces what we might call the *justification problem*.²³ And that might be enough to start looking for alternatives.²⁴

The second proposal takes adequate understanding not as a justifier, but as a necessary condition on the justificatory power of intellectual seemings. So, the idea is that the intellectual seeming that honesty is a virtue justifies the corresponding moral belief *only if* it is grounded in adequate understanding. This seems to be what, on a charitable interpretation, Stratton-Lake (2016) suggests. In more detail, he writes that ‘...understanding is a necessary condition of a seeming having the sort of justificatory force it has by figuring in the right sort of explanation of why that proposition seems true.’ (Stratton-Lake 2016: 17). So, for your intellectual seeming that p to justify your belief that p, you must adequately understand that p.²⁵

²² Of course, Stratton-Lake’s argument might not be water-tight. Here are two ways of pushing back. First, self-evidence theorists might point out that responding ‘Because I believe p’ to the question ‘Why do you believe (self-evident) p?’ loses some of its oddity when we spell out exactly what adequately understanding that p entails. After all, adequate understanding, at least on some conceptions, might be factive: if you adequately understand p, p is true. So, adequately understanding that p seems to raise the epistemic probability of the truth of p – and therefore counts as evidence in favour of p. For instance, suppose I ask you why you believe that $5 + 7 = 12$. You answer my question by saying that you adequately understand that $5 + 7 = 12$. By that, you don’t just mean that you know what it says, but that you are able to correctly apply it, see some of its logical implications, and so forth. In that case, it strikes me that your adequate understanding of $5 + 7 = 12$ gives you at least some evidence for believing it to be true. Second, another issue arises for Stratton-Lake’s probabilistic account of evidence. Synthetic apriori propositions are either true or false as a matter of (conceptual or metaphysical) necessity. If they are necessarily true, their probability must be 1. Conversely, if they are false, their probability must be 0. Given that their probability must either be 1 or 0, it makes little sense to understand evidence for synthetic apriori propositions as raising the probability of their truth. As a result, it remains somewhat moot why adequate understanding of p could or should be evidence for p. (See Clarke-Doane forthcoming 2017 for a similar worry.) However, at the end of the day, I don’t need Stratton-Lake’s argument to be water-tight. Rather, it suffices if it shows that self-evidence theorists struggle with justification, while other version of moral intuitionism don’t. That alone is reason to prefer the latter over the former.

²³ According to Stratton-Lake, that is not the only problem for the view. Rather, the self-evidence view also struggles to accommodate the *recalcitrance* of moral intuitions: sometimes, we intuit p without believing that p. For instance, a moral error theorist might intuit that pleasure is better than suffering, yet fail to believe so due to her theoretical commitments. See Stratton-Lake 2016: 9-15 for further discussion.

²⁴ According to Stratton-Lake, this means that self-evidence and adequate understanding shouldn’t play a distinctive role in intuitionist moral epistemology. That, though, turns out to be slightly misleading: as we shall see shortly, he seems to think that adequate *can* play a role in moral intuitionism, just not the role of a justifier.

²⁵ Elsewhere, Stratton-Lake calls such justifying intellectual seemings *clear*, i.e. ‘[b]y getting a proposition clearly in view I mean pretty much what Audi calls having an adequate understanding of it’ (Stratton-Lake 2016: 13, fn.25), and insists that only moral beliefs formed on the basis of such intellectual seemings suffice for knowledge. However, as I shall explain below, this combination of theoretical commitments renders Stratton-Lake’s view unattractive.

Given Stratton-Lake's proposal regarding the role adequate understanding should play, the intellectual seemings version of moral intuitionism quite straightforwardly accommodates Gettier-style counterexamples like **Brain Lesion**. After all, if you don't adequately understand that honesty is a virtue and justification requires adequate understanding, you are not justified in believing that honesty is a virtue in **Brain Lesion**. But, since, on any form of moral intuitionism, justification is necessary for *intuitive* moral knowledge, you don't *know* that honesty is a virtue. In that manner, Stratton-Lake's view avoids the problem of epistemic luck. Further, unlike the self-evidence version of moral intuitionism just discussed, Stratton-Lake's form of the intellectual seemings view doesn't face the *justification problem*. After all, as both Huemer (2005: §5.1.) and Stratton-Lake (2016: 13) stress, intellectual seemings are comparable to *perceptual seemings* – and perceptual seemings can justify corresponding perceptual beliefs. For instance, if it seems to you that the Swiss flag features a white cross and you form a perceptual belief on the basis of such a perceptual seeming (and you are unaware of any defeaters), your perceptual belief that the Swiss flag features a white cross is *justified*. Similarly, if you form a moral belief based on your intellectual seeming that enjoyment is better than suffering (and you are not aware of any defeaters), your corresponding moral belief is *justified*. So, it looks like the intellectual seemings version of moral intuitionism avoids the problem with justification that the self-evidence version of the view faces.²⁶

Contrary to initial appearances, however, we should still not opt for Stratton-Lake's view – and thus the second way of integrating adequate understanding into moral intuitionism. How so? Stratton-Lake's view seems to run into – what I shall call – the *conflation problem*: it conflates the distinction between justification and knowledge. To see that, it suffices to reconsider the **Brain Lesion** scenario: according to Stratton-Lake, the epistemic subject is neither justified in believing nor knows that honesty is a virtue, given that she doesn't adequately understand – and adequate understanding is a necessary condition on the justificatory power of intellectual seemings. That verdict, though, seems implausible: intuitively, the subject in that scenario *does* have a justified belief, yet not knowledge, that honesty is a virtue. After all, she has the corresponding intellectual seeming – and isn't aware of any defeaters for it. So, why wouldn't her moral belief be justified? While Stratton-Lake's proposal might address the problem of epistemic luck, it does so only at the cost of blurring a widely accepted distinction between justification and knowledge. That cost, of course, might not be decisive. Still, it strikes me as providing motivation enough to seek a form of moral intuitionism that doesn't incur that cost.

²⁶ The intellectual seemings view also neatly accommodates the recalcitrance of moral intuitions: we need not believe what seems true to us. See Stratton-Lake 2016: 9-15 for further discussion.

The third and final way of integrating adequate understanding into moral intuitionism construes adequate understanding as a distinctive necessary anti-luck condition on intuitive moral knowledge. In other words, for you to know by intuition that, say, honesty is a virtue, you must adequately understand it. Importantly, unlike the other two options examined above, adequate understanding neither justifies nor is required for having a justified moral belief based on a corresponding intellectual seeming. The resulting view says that you know an intuitive moral proposition *p* if you (truly) believe that *p*, justified on the basis of a veridical intellectual seeming that *p*, which results from adequately understanding *p*.

This seems to be what Michael Huemer (2005: 123-27) has in mind. To address the problem of epistemic luck or ‘...to explain how it would be anything more than chance if my moral beliefs were true’ (123), Huemer sketches a more general account of *apriori* knowledge that prominently features adequate understanding as an anti-luck condition. His account contains four key elements (cf. 2005: 124-26): (i.) that mind-independent universals (including moral properties and relations) exist necessarily, (ii.) that having an *adequate* (i.e. consistent, clear, and determinate) concept constitutes the *grasping* of a mind-independent universal, (iii.) that adequately grasping a universal puts you in a position to see that it has certain properties and/or relations to other universals that one adequately grasps, and (iv.) that all *apriori* knowledge derives from knowledge of universals. Together with *plenitude*, i.e. the claim that having an adequate concept suffices for grasping a universal and thus that ‘there is no possibility of one’s failing to refer to anything...’ (126),²⁷ (i.) to (iv.) allow Huemer to argue that ‘...when one’s intuitions are caused (only) by clear, consistent, and determinate understanding [...] the internal process by which one forms beliefs guarantees their truth’ (2005: 126). So, adequate understanding effectively serves as an anti-luck condition on intuitive *apriori* knowledge. And it is easy to see how this condition allows Huemer to handle cases involving *moral* belief such as **Brain Lesion**: your intellectual seeming that honesty is a virtue is caused by your brain lesion, not adequate understanding. Therefore, you cannot have intuitive moral knowledge.

Importantly, Huemer’s way of integrating adequate understanding into moral intuitionism neither faces the *justification problem* nor the *conflation problem*. To begin with, intellectual seemings about moral propositions still serve as the justifier for corresponding moral beliefs. So, unlike the self-evidence version of moral intuitionism,

²⁷ Werner (2018) alerts us to the crucial role of plenitude in Huemer’s account, which he characterizes as follows: ‘For every possible *adequate* (consistent, clear, and determinate) concept, there is a corresponding mind-independent universal’ (622). We shall return to discussing plenitude in §4 below.

we have a plausible story about what justifies basic moral beliefs. Further, unlike Stratton-Lake's version of the view, it doesn't conflate the distinction between justification and knowledge. After all, according to Huemer, you *do* have a justified belief in **Brain Lesion**, given your corresponding intellectual seeming. Still, you don't *know*, given that your intellectual seemings isn't caused by adequate understanding. So, Huemer's view respects our intuitive epistemic verdict about **Brain Lesion** – and thus doesn't conflate justification and knowledge. That, though, is good news: it shows that there is a version of moral intuitionism, namely Huemer's (2005), that addresses the problem of epistemic luck, yet doesn't do so at the cost of complicating moral justification or violating central tenets of epistemological theorizing. Therefore, if we want to be intuitionists about moral knowledge, we should endorse a view according to which you know (an intuitive moral) *p* if you (truly) believe that *p*, justified on the basis of a veridical intellectual seeming that *p* that results from adequately understanding *p*.

Where does this leave us? In this section, I have argued that there are conceptual tools available to moral intuitionists to handle the problem of epistemic luck raised in Chapter 2, namely an appeal to adequate understanding. But, as I have shown above, it isn't obvious what role adequate understanding should play within the architecture of intuitionist moral epistemology. In response, I have made the case that adequate understanding is most plausibly construed as a necessary anti-luck condition on intuitive moral knowledge, since that avoids complicating moral justification and its relationship to moral knowledge. As a result, we end up with a version of moral intuitionism close to the one defended by Huemer (2005). But can this view double as an effective response to the best evolutionary explanation argument that I developed in Chapter 4? The next section addresses that question in depth.

4. Moral Intuitionism and the Best Evolutionary Explanation Argument

In this section, my aim is to examine whether moral intuitionists can respond to the best evolutionary explanation argument identified in Chapter 4. I first set the stage by re-introducing the argument, outlining the desiderata on any plausible response and arguing that moral intuitionism, by itself, doesn't meet them (§4.1). Then, I suggest that supplementing the view with a third-factor explanation of moral reliability seems to do the trick instead (§4.2). However, as I argue, any such explanation must rely only on formal moral intuitions to avoid begging the question (§4.3) – and those intuitions cannot underwrite substantive belief about morality (§4.4). So, moral intuitionists cannot respond to the best evolutionary explanation argument of Chapter 4. By implication, robust moral realists are saddled with implausible moral scepticism: they cannot make

sense of moral knowledge or even justified beliefs about morality, construed as a realm of robustly moral facts.

4.1. Setting the Stage

In Chapter 4, I argued that the following is a powerful abductive argument against robust moral realism:

- (1) Evolutionary, not robustly moral, facts best explain why we hold certain moral beliefs and not others. (Or, robustly moral facts are *dispensable* to the *best explanations* of our moral beliefs.)
- (2) If our moral beliefs are not best explained by robustly moral facts and there are no additional, non-abductive reasons for maintaining them, that defeats their justification. (Or, if robust moral facts are *dispensable* to the *best explanation* of our moral beliefs and there are no additional, non-abductive reasons for maintaining them, that defeats the justification of those beliefs (that *purport* to represent them).)
- (3) There are no additional, non-abductive reasons for belief in robust moral facts.
- (4) Therefore, (attention to) facts about evolution defeat the justification of our moral beliefs, robustly construed.

Importantly, this argument, as seen in Chapter 4, shifts the argumentative burden onto robust moral realism. The burden amounts to an epistemological challenge, based on the principle that ‘[i]f the truth and content of our moral beliefs is not involved in the best [evolutionary] explanation for our possession of them, then we need *additional reasons* to believe them’ (Woods 2016: 6; emphasis mine). And if robust moral realists cannot meet the challenge by providing such additional, non-abductive reasons, they acquire an undermining defeater for the justification of their moral beliefs. Moral scepticism follows.

What does it take to resist this argument? In Chapter 4, I suggested that the most promising way involves questioning the truth of premise (3) by offering plausible non-abductive reasons for maintaining robust moral belief. That way, robust moral realists could meet the challenge head-on and straightforwardly discharge the argumentative burden. But what do such reasons look like? In Chapter 4, I critically examined a recent, widely discussed proposal by David Enoch (2012, 2016): that robust moral facts are *indispensable* to practical *deliberation* – and that this gives us reason to believe in their existence. In my view, that proposal fails. Its failure, though, proves instructive for our purposes. More precisely, I take it to reveal a minimal standard of plausibility for any non-

explanatory reason to maintain robust moral beliefs. How so? As David Plunkett and Tristram McPherson (2015) point out, deliberative indispensability, construed as a pragmatic vindication of the basic sources of epistemic justification,²⁸ conflicts with a distinctive feature of epistemic justification, namely its *Truth-Directedness*.²⁹ That feature states that ‘the sources of basic epistemic justification have the content that they do (in part) because of some positive connection between those sources and the truth of the beliefs that they govern’ (2015: 114). Deliberative indispensability violates this constraint because ‘...the fact that the belief that p is indispensable to our deliberative projects bears no positive relationship to the truth of p’ (2015: 121). What does that mean for us? It means that any plausible non-explanatory reason for maintaining robust moral beliefs must at least respect *Truth-Directedness*: it must explain how our basic moral belief-forming methods are positively connected to the moral truth. If it failed to do so, it wouldn’t qualify as a plausible non-explanatory reason for robust moral belief. And if it didn’t give us such a reason, robust moral realists couldn’t resist the argument above by denying premise (3). So, the failure of deliberative indispensability helps us articulate a desideratum for any successful response to the best evolutionary explanation argument on behalf of robust moral realism.

In §3 above, I argued for a Huemer-style view as the most compelling version of contemporary moral intuitionism. Can that view double up as a response to the critical argument just rehearsed? In other words, maybe some of our robust moral beliefs are based on veridical intellectual seemings, resulting from adequate understanding. Is that a plausible non-explanatory reason to maintain them? At first glance, we might think so, given that Huemer-style moral intuitionism seems to respect *Truth-Directedness*. After all, according to Huemer, ‘...when one’s intuitions are caused (only) by clear, consistent, and determinate understanding [...] the internal process by which one forms beliefs guarantees their truth’ (2005: 126). Clearly, if everything goes right on his view, moral intuitions are positively connected to the moral truth. So, as long as our beliefs are based on moral intuition in the right way, we have reason to maintain our robust moral beliefs – and premise (3) of the argument above must therefore be false.

²⁸ More precisely: they take issue with the first premise of the argument in support of deliberative indispensability, which says that ‘...[o]ne complete ground for the fact that something is a source of basic epistemic justification is the fact that treating it as such a basic source is instrumentally indispensable to an intrinsically indispensable project [such as practical deliberation]’ (Plunkett & McPherson 2015: 110).

²⁹ They write that the positive connection to truth should be understood ecumenically (e.g. in terms of modal notions such as safety or sensitivity) and that this indeterminacy is dialectically crucial: ‘we take *commitment* to something like Truth-Directedness to be close to common ground in many parts of epistemology’ (116).

On reflection, however, that move proves too hasty. To see why, we can lean on a recent observation by Preston Werner (2018: 622-24). In short, he argues that Huemer's moral intuitionism, as an account of moral knowledge, doesn't just have to guarantee that our normative beliefs, when based on veridical intuitions grounded in adequate understanding, are true. Rather, moral intuitionism must also connect our thus formed beliefs to relevantly *robust* normative truths. After all, given the contingency of our normative conceptual framework, 'the normative concepts that we've developed may not have latched onto the robustly normative properties [but rather only onto 'merely' normative properties]' (2018: 620). To develop his worry, Werner first distinguishes between *mere* and *robust* normativity.³⁰ The rules of chess, etiquette and legal proceedings are merely normative: they offer no practice-independent reason to care and aren't authoritatively binding. In contrast, the edicts of morality are robustly normative: they have genuine normative force and are intrinsically binding.³¹ That, in turn, motivates a constraint on any adequate, non-sceptical moral epistemology that robust moral realists put forward: it cannot just exclude epistemic luck for some of our true, justified, paradigmatically normative beliefs. Instead, any such epistemology also needs to explain how some of our normative beliefs are about genuinely robust normative truths. Werner calls this requirement *Content Success*: '...at least some of our justified, first-order, and paradigmatically normative beliefs contain robustly normative contents' (2018: 619).³²

Is there reason to think that Huemer's moral intuitionism cannot make good on *Content Success*? To see that there is, Werner invites us to imagine Lucy and Carol. Both of them are in the epistemically speaking good case, by Huemer's lights: they form moral beliefs on the basis of veridical intellectual seemings, resulting from adequate understanding. In other words, their moral concepts are *adequate*, allowing them to successfully refer to and thus grasp mind-independent moral properties. They both have intuitive moral knowledge. But here is the rub: by stipulation, Lucy and Carol have moral intuitions with inconsistent normative contents. While Lucy has the intuition that lying is intrinsically bad, Carol has the intuition that lying is not intrinsically bad.³³ How can we make sense

³⁰ Werner (2018) calls it 'formal', not 'mere', normativity. However, I reserve 'formal' for a slightly different notion introduced in §4.3 below.

³¹ More precisely: 'An entity is *robustly normative* iff it is either fundamentally intrinsically binding (in the way that [merely] normative entities are not), or not fully explicable without reference to some fundamentally intrinsically binding entity' (Werner 2018: 618).

³² Werner also stresses that any adequate moral epistemology must meet the anti-luck condition and *Content Success* simultaneously with respect to many of the same beliefs. In other words, it must show that *Overlap* is true: 'at least some of our justified, first-order, and paradigmatically normative beliefs are *both* non-accidentally true and contain robustly normative contents' (619).

³³ As Werner (2018: 623f.) points out, Huemer cannot just respond by denying the possibility of such cases. After all, according to him, the defining characteristics of adequately grasping, namely consistency, clarity

of this incompatibility on Huemer's picture? Plausibly, rather than being mistaken, they both have different moral concepts altogether, namely BADNESS_(Lucy) and BADNESS_(Carol). They wouldn't genuinely disagree but talk past each other. That diagnosis, though, raises an important question: which moral concept adequately characterizes the *robustly* normative property? According to Werner, Huemer's moral intuitionism '...has provided no reason to accept that Lucy's normative concepts actually refer to the robustly normative properties, and thus no reason to accept that her normative beliefs were of the robustly normative facts' (624).³⁴ In other words, Huemer's moral epistemology, despite excluding epistemic luck, doesn't meet *Content Success*. It thus isn't a fully adequate moral epistemology for robust moral realism.

What implications does that have for our dialectic? Werner's discussion shows that Huemer's moral intuitionism isn't sufficient as a response to the best evolutionary explanation argument.³⁵ Sure, the view offers a positive, non-accidental connection to moral truth – and thus, unlike deliberative indispensability, appears to satisfy *Truth-Directedness*. However, as it stands, it struggles to meet *Content Success*, failing to explain why our moral beliefs connects to the *right kind* of moral truth. In other words, Huemer's moral intuitionism might give us a non-abductive reason to maintain our moral beliefs, yet not those in *robust* moral facts. So, to call into question the truth of premise (3) and sabotage the argument above, moral intuitionism requires further revision. But what might that look like?

4.2. Third-Factor Explanations

In this sub-section, my aim is to argue that moral intuitionists should supplement their view with a third-factor explanation of moral reliability. Such an explanation promises to satisfy the remaining desideratum articulated above: it explains why our moral beliefs, based on the right kind of intuition and grounded in conceptual competence, connect to *robustly* moral truths.

and determinacy, are *intrinsic* or belong to the concepts themselves. So, as long as those characteristics are present, an epistemic subject would grasp moral properties, irrespective of whether they are ultimately robust or not.

³⁴ In response, why couldn't Huemer pack into Lucy's concept that it refers to robustly normative badness, if it refers at all? Because this, as Werner helpfully points out, means rejecting *Plenitude*, which implies that *Non-Accidentality* is no longer met, given that '...there becomes a non-trivial chance that an adequate concept will fail to target any property' (2018: 624).

³⁵ To some extent, that is unsurprising. Sure, moral intuitionism features a no-defeater clause. However, it doesn't tell us in detail when and why that clause remains met. Adding adequate understanding as an anti-luck condition takes care of one source of defeat, namely knowledge-undermining epistemic luck. But it doesn't address other kinds of defeaters, including the one articulated by the best evolutionary explanation argument. So, moral intuitionism requires supplementation, which I shall provide below.

In response to evolutionary debunking arguments, robust moral realists such as David Enoch (2010, 2012: Ch. 7), Knut Olav Skarsaune (2011), Erik Wielenberg (2010) and others have recently defended so-called *third-factor explanations* of moral reliability.³⁶ For them, the challenge raised by those arguments consists in explaining the correlation between our moral beliefs and the corresponding robust moral truths. As we saw in Chapter 2, explaining that correlation is complicated by two prominent features of robust realist metaphysics. On the one hand, moral beliefs cannot be caused by moral truths, given their causal inefficacy and irreducibility to natural truths. And on the other, moral beliefs cannot ground moral truths, given their objectivity. So, those realists propose a *third factor*: a putative normative truth such as that survival is good (Enoch), pain is bad (Skarsaune) or that we, as human beings, have certain moral rights (Wielenberg). That normative truth, despite being shaped by evolutionary forces, *causes* our moral beliefs, while at the same time *constituting* or *cohering* with the robust moral truths. In that manner, the third factor explains the correlation (or, in Enoch's terminology: pre-establishes the harmony) between our moral beliefs and their corresponding robust moral truths. So, the mystery of moral reliability has been dispelled.

Admittedly, third-factor explanations were originally developed in the context of the reliability challenge to robust moral realism. As I argued in Chapter 2, that challenge – interpreted modally – does not amount to more than the problem of epistemic luck and is therefore unlikely to motivate evolutionary debunking. So, there might be no need for third-factor explanations in response to the reliability challenge. That, though, doesn't mean that we cannot combine them with moral intuitionism and put them to work in our context. After all, that combination promises to meet the two desiderata we discussed above on any plausible response to the best evolutionary explanation argument.

First off, recall *Truth-Directedness*, which says that 'the sources of basic epistemic justification have the content that they do (in part) because of some positive connection between those sources and the truth of the beliefs that they govern' (Plunkett & McPherson 2015: 114). As we saw above, Huemer's moral intuitionism already seems to satisfy that requirement. At least when one's moral intuitions result from adequate understanding, they are a basic moral belief-forming method positively connected to the moral truth. Still, adding a third-factor explanation might render the account even deeper and powerful, situating moral intuitions squarely within a broadly evolutionary framework.

³⁶ See also Brosnan 2011 and Schafer 2010.

Second and more importantly, appealing to a third-factor explanation promises to help moral intuitionists with respect to Werner's (2018) *Content Success*. This requirement states that any suitably non-sceptical moral epistemology needs to explain how some of our justified normative beliefs are about genuinely robust normative truths. As we saw above, moral intuitionism alone doesn't offer anything to address that, even once we add adequate understanding as an anti-luck condition. However, a third-factor explanation might help. Suppose it is a robust normative fact that survival is good (as Enoch suggests) – and robust moral realists are entitled to that supposition. Further, suppose that, as any third-factor explanation states, this fact *causes* our corresponding normative belief. In that case, we have least one normative belief that is about a robustly normative fact, namely the goodness of survival. Further, that robustly normative fact makes up, entails or coheres with other robustly normative facts such as, say, that avoiding danger is (*pro tanto*) good. Then, by figuring out how those facts relate to each other, we can acquire further beliefs about robustly normative facts. In this manner, third-factor explanations can help moral intuitionism give an account of how some of our normative beliefs, based on veridical intellectual seemings resulting from adequate understanding, concern genuinely *robust* normative truths. *Content Success* can be met after all.³⁷

But how plausible is that strategy overall? Many have argued that third-factor explanations beg the question against the evolutionary debunker. In the remainder of §4, I shall tackle this issue in depth. As I suggest, such explanations aren't problematically circular, yet fail to underwrite most of our substantive moral beliefs. By implication, third-factor explanations cannot stave off the sceptical worry raised by best evolutionary explanation argument in Chapter 4.

4.3. Formal Moral Intuitions and Begging the Question

Third-factor explanations crucially hinge on a normative claim such as that, say, survival is good. But doesn't that beg the question against the evolutionary debunker? In this sub-section, I argue that it does not. But to avoid begging the question, such explanation must rely on *formal* moral intuitions exclusively.³⁸ As the first step of my argument, I clarify the

³⁷ In fact, David Enoch (2012: 177-84) sketches such a line in response to the worry that robust moral realists cannot explain *semantic* access to moral properties.

³⁸ Behrends (2013) also develops a third-factor explanation along those lines. However, he simply *assumes* that relying on formal moral (or normative) assumptions does not beg the question in the context of third-factor explanations of moral reliability. That assumption, however, is itself highly controversial and in need of detailed defense. And that defense is exactly what I offer in this sub-section. For more on Behrends (2013), see §4.3.4 below.

issue by distinguishing defeater-deflectors from defeater-defeaters (§4.3.1.). Then, I suggest that formal moral intuitions, such as that betterness is transitive, can serve as defeater-deflectors, given their likely exemption from noxious evolutionary influence (§4.3.2.). Finally, after addressing a complication with the origins of moral reliability (§4.3.3.), I propose a third factor derived purely from formal moral intuitions (§4.3.4.).

4.3.1. Begging The Question and Deflecting Defeat

Many critics of third-factor explanations argue that they beg the question.³⁹ After all, such explanations crucially depend on a normative claim such as that survival is good, that pain is bad, and so forth. Without that claim, nothing pre-establishes the harmony between moral beliefs and robustly moral facts. But, of course, the best evolutionary explanation argument articulates a potential undermining defeater for exactly such claims. Or, more precisely, the argument asks for non-explanatory reasons for robustly moral beliefs. In response, robust moral realists cannot just *assume* that some of their moral beliefs are true – and that this gives them such a reason to maintain their robustly moral beliefs. Doing so would be dialectically inappropriate. So, by depending on a normative claim, third-factor explanations assume *ab initio* what they set out to prove. Now, the validity of this criticism has been hotly debated in the literature.⁴⁰ In fact, Katia Vavova (2015) goes so far as to call it the ‘heart of the debate between the realist and the debunker’ (111). Still, no resolution has been reached yet. So, how can we make progress on the issue?

To make progress, we should heed the lessons from a parallel debate in religious epistemology about Alvin Plantinga’s (1993) evolutionary argument against naturalism. According to Plantinga, those who accept naturalism and evolutionary theory have a defeater for the belief that their cognitive faculties are reliable. By implication, all their beliefs are defeated, including their beliefs in naturalism.⁴¹ For critics, it is crucial to determine which beliefs they may use to respond to this potential evolutionary defeater. As Andrew Moon (2017) suggests, those critics face a situation similar to the one that proponents of third-

³⁹ See Fraser 2014: 471, Street 2008 or Vavova 2015: 111. In response, Berker 2014, Schafer 2010 and White 2010 argue that circularity is inherent in the explanation of reliability of any class of beliefs. Note that those contributions are primarily concerned with versions of the evolutionary debunking argument based on the reliability challenge or higher-order evidence of error. However, I take their worry to carry over to third-factor explanations as a response to the best evolutionary explanation argument.

⁴⁰ The discussion of the self-defeat objection in Chapter 3 also touches on this issue, albeit in a different theoretical framework, namely in terms of defeat by higher-order evidence of error.

⁴¹ For discussion, see Fitelson & Sober 1998, Fodor 2002 or Ginet 1995.

factor explanations find themselves in vis-à-vis evolutionary debunking arguments.⁴² Interestingly, most contributors to the debate about Plantinga's argument agree that relying on F-faculty to avoid potential defeat of F-beliefs is not *always* problematically circular. To see that, Moon (2017: 213) urges us to consider the following scenario:

XX Defeater-Deflector: You learn that a pill, called 'XX', destroys the cognitive reliability of 95% of those who ingest it. Before taking the pill, however, a scientist you know to be trustworthy informs you that you are one of the 5% who is immune to the drug. You then take XX while knowing that you are one of the immune 5% and $P(R/I've\ ingested\ XX\ and\ I\ am\ one\ of\ the\ immune\ 5\%)$ is high.

Usually, believing that you have ingested XX and that XX destroys cognitive reliability gives you a defeater for R: that human cognitive faculties are generally reliable. But above, things are different: the scientific testimony you acquire before taking XX serves as – what Moon (2017) and others call – a *defeater-deflector*.⁴³ A defeater deflector is a belief (or proposition) that prevents a potential defeater from becoming an actual defeater. Here, your belief that you are one of the immune 5% deflects the potential defeater that you have ingested XX and $P(R/You've\ ingested\ XX)$ is low. Put differently, the potential defeater '...would never gain defeating power in the first place' (Moon 2017: 213), or you never have reason to doubt your cognitive reliability. Epistemically speaking, there is nothing to worry about. Importantly, the case illustrates that avoiding potential defeat of F-beliefs by relying on the F-faculty itself can sometimes be epistemically innocuous. In the scenario above, you would not beg the question by insisting that you were XX-immune or cognitively reliable. That, however, is an important insight, for it tells us how robust moral realists must interpret their third-factor explanations. For those not to beg the question or be objectionably circular, the normative claim at their heart must serve as a defeater-deflector. Or, put differently: for the third-factor explanations to work, robust moral realists must plausibly be in an epistemic situation comparable to the **XX Defeater-Deflector** case above.⁴⁴

⁴² Moon's (2017) discussion focuses on the evolutionary debunking argument based on the principle that reliability requires explanation developed in Chapter 2. However, I take his suggestion to generalize to arguments based on other principles, including that that justification requires explanation.

⁴³ See Plantinga 2002: 224, fn. 30, and Moon 2017: 211.

⁴⁴ Moon (2017: 221) is ultimately *agnostic* about whether the third-factor explanations provide robust moral realist with a defeater-deflector, but has a hunch that they do not.

However, we might worry that robust moral realists don't find themselves in anything close to *XX Defeater-Deflector*. As Michael Klenk (2017) suggests, their situation differs significantly. The scenario above prominently features trustworthy scientific testimony to corroborate one's immunity to the noxious, cognition-destroying effects of the XX pill. But nothing plausibly plays that role for the robust moral realist in an evolutionary debunking scenario. After all, all their moral beliefs have an evolutionary history. Thus, they lack an independent, reliable source of moral information.⁴⁵ But without such a source, why think that robust moral realists are immune to distorting evolutionary influence? At best, they find themselves in the equivalent of the following scenario:

XX Defeater-Defeater: You learn that a pill, called 'XX', destroys the cognitive reliability of 95% of those who ingest it. Two hours *after* having taken the pill, a scientist you know to be trustworthy informs you that you are one of the 5% who is immune to the drug. You come to believe that you are one of the immune 5% and that $P(R/I've\ ingested\ XX\ and\ I\ am\ one\ of\ the\ immune\ 5\%)$ is high.⁴⁶

Here, unlike above, believing that you have ingested XX and that XX destroys cognitive reliability *does* give you a defeater for R: that human cognitive faculties are generally reliable. After all, the scientific testimony you acquire *after* taking XX cannot serve as a defeater-deflector. Rather, it would have to play the role of a *defeater-defeater*: a belief (or proposition) that defeats something that is already an actual defeater.⁴⁷ But, importantly, the defeater-defeater is clearly inadmissible in this context: given that you knowingly took the reliability-destroying XX pill, you already have reason to distrust any belief you form going forward, including the belief that you are XX-immune, based on putative scientific testimony.⁴⁸ So, an appeal to scientific testimony would beg the question in **XX Defeater-Defeater**. And if the robust moral realist's epistemic position is relevantly similar, so would their appeal to a third factor. Accordingly, proponents of third-factor explanations wouldn't avoid the charge of begging the question.⁴⁹

⁴⁵ In Klenk's (2017) own words: '...in the moral case, we would have to accept another extremely controversial assumption: namely, that the scientist's testimony is reliable in the first place' (236).

⁴⁶ See Moon 2017: 213.

⁴⁷ See Moon 2017: 211.

⁴⁸ Or, more technically: 'any potential defeater-defeater will itself already be defeated by the original defeater for R' (Moon 2017: 210).

⁴⁹ Klenk (2017: 232-36) also criticizes the **XX Defeater-Deflector** case in another regard: according to him, it is committed to epistemic externalism (about defeat), which is incompatible with robust moral realism. Since I argue below that robust moral realists don't occupy a position relevantly similar to the **XX Defeater-Deflector** case anyway, Klenk's criticism misses the mark. Thus, I won't engage with it further here.

4.3.2. Deflecting Defeat with Formal Moral Intuitions

But Klenk's worry proves unfounded, as I shall argue now. Sure, robust moral realists, and especially moral intuitionists, may not be in a situation as favorable as the **XX Defeater-Deflector** case. But neither are they in a position as epistemically dire as **XX Defeater-Defeater**. To see that, it suffices to highlight a significant difference between the nature of XX pills and evolutionary influence. The XX pill – by stipulation – destroys *all* of one's cognitive reliability and calls *all* of one's beliefs into question, except for the beliefs that one took the pill and that it had exactly those noxious epistemic effect.⁵⁰ The potency of the pill neatly explains why gaining (putative) scientific testimony *after* taking the pill doesn't improve one's epistemic position. In contrast, evolutionary influence isn't an all-or-nothing affair. Rather, it is overwhelmingly likely that it didn't disrupt all of our moral reliability equally. While some *kinds* of moral belief are at epistemic risk due to their evolutionary origin, others won't be as much. Further, we might be able to tentatively categorize moral beliefs correspondingly and make educated guesses about how that risk is distributed. If that line of thought had promise, robust moral realists would most definitely not be as badly off as anyone stuck in **XX Defeater-Defeater**. Rather, they would have a potential defeater-deflector, namely those moral beliefs less likely affected by disrupting evolutionary influence.

But which moral beliefs are least likely to be disrupted by evolutionary influence? Following Michael Huemer (2008), I think we should distinguish three kinds of moral intuitions – and their respective susceptibility to epistemically bad evolutionary (and other) influence. First, there are *concrete*, single-case moral intuitions. Examples include intuitions about Philippa Foot's (1967) trolley cases, Peter Singer's (1972) shallow pond thought-experiment, special kind obligations or the prohibition against incest. Such intuitions are most likely to be vulnerable to – what Huemer (2008: 376) calls – 'biological programming', given their obvious link to adaptive behavior. In addition, concrete moral intuitions are also susceptible to other *biases*, especially cultural indoctrination, emotions and personal interest. After all, they often concern culturally sensitive, emotionally charged or personally significant situations.⁵¹ As a result, we should treat concrete moral intuitions with caution in moral theorizing.

⁵⁰ In the **XX Defeater-Deflector** case, it also leaves intact one's belief or memory about XX-immunity, based on trustworthy scientific testimony acquired *before* taking the pill.

⁵¹ In Huemer's words: '...when we are emotional about moral issues or when our own interests are at stake, the danger that our intuitions will be biased is too great for us to be justified in relying on those intuitions, in the absence of independent corroboration' (2008: 378).

Second, we have *mid-level* moral intuitions: intuitions about moral principles of intermediate generality. For instance, consider the intuition that, all things equal, one ought to keep one's promises or that adultery is wrong. Such intuitions share the epistemic risks of concrete moral intuitions: they are equally susceptible to biological, cultural and emotional biases. But, additionally, mid-level moral intuitions are prone to the problem of *overgeneralization*: judging the truth of a generalization in terms of typical cases, thus overlooking atypical ones. For instance, we might judge that we must keep our promises, despite the existence of atypical cases in which breaking promises seems morally permissible, perhaps because doing so would save numerous human lives. Given that mid-level moral intuitions are prone to both bias and overgeneralization, we should avoid them in moral theorizing.

Third and finally, there are *abstract* moral intuitions: intuitions about very general moral principles. For example, take the intuition that it is wrong to treat individuals as mere means or that the betterness relation is transitive.⁵² Unlike concrete or mid-level moral intuitions, abstract moral intuitions are less likely to be influenced by evolutionary, cultural, emotional or personal bias. After all, they concern moral matters that are not obviously adaptive, culturally informed, emotionally charged or personally meaningful. Instead, they are plausibly the upshots of rational reflection. They also avoid overgeneralization, given that they don't arise from considering merely typical cases.⁵³ Since they are largely exempt from sources of error that plague concrete or mid-level moral intuitions, these intuitions are unusually trustworthy – and should thus guide our moral theorizing. Amongst abstract moral intuitions, we should particularly focus on *formal* moral intuitions: intuitions that impose formal constraints on moral theories without positively or negatively evaluating anything. Despite not generating a substantive moral system, these intuitions may still help us adjudicate moral disagreements by ruling out otherwise attractive combinations of moral commitments.⁵⁴ Therefore, according to Huemer, '...formal ethical intuitions should be given special weight in moral reasoning' (2008: 387).

How is this typology of moral intuitions (and their associated epistemic risks) relevant for our discussion – and, in particular, for whether the normative claim at the heart of third-factor explanations can deflect evolutionary defeat? As we have just seen, it is plausible

⁵² See Huemer 2008: 386 for more examples.

⁵³ By implication, as Huemer (2008: 386f.) points out, putative counterexamples lead to a *paradox*, as opposed to forcing us to give up those principles straight away. See, for instance, Rachels' (1998) putative counterexamples to the transitivity of 'better'.

⁵⁴ Take, for example, Huemer's (2003) own criticism of welfare egalitarianism or Parfit's (1984: 419-30) discussion of the repugnant conclusion.

that not all moral beliefs are equally subject to problematic evolutionary influence. Instead, some moral beliefs, namely those formed on the basis of *formal* moral intuitions, are most likely exempt. That, though, is great news for the advocate of third-factor explanations. After all, if robust moral realists managed to appeal to *those* moral (or normative) beliefs as their third factor, they could use them as a *defeater-deflector* – instead of a defeater-defeater. How so? Because those formal moral beliefs, like any paradigmatic defeater-deflector, would prevent the potential evolutionary defeater from becoming an actual one. With them in place, robust moral realists never gain reason to doubt their reliability with respect to *formal* moral matters. To illustrate, they would find themselves in an epistemic situation relevantly similar to this:

YY Color Defeater-Deflector: You learn that a pill, called ‘YY’, destroys the reliability of color vision in 100% of people who ingest it. Two hours *after* taking the pill, a scientist you know to be trustworthy informs you that, upon further inquiry, it turns out that the pill doesn’t affect all color vision equally: while it renders the perception of blue and green tones unreliable, it leaves perception of red tones intact. You come to form a corresponding belief about the partial reliability of your color vision.⁵⁵

Here, just like in the problematic **XX Defeater-Defeater** case, you receive the trustworthy scientific testimony *after* taking the pill. However, this case differs in two crucial, redemptive respects.

To begin with, the YY pill doesn’t destroy your cognitive reliability totally or globally. Rather, it only negatively affects your color vision. As a result, it is epistemically permissible for you to rely on scientific testimony and any cognitive capacity other than color vision. Similarly, evolutionary influence doesn’t negatively affect all the robust moral realist’s cognitive faculties. Most likely, she can trust her basic perceptual, inductive, arithmetical, logical and even some other *apriori* belief-forming processes, despite their evolutionary history. Importantly, that implies that robust moral realists can also help themselves to the typology of moral intuitions (and their associated epistemic risks) above. After all, devising the typology combines scientific theorizing (about the scope of evolutionary explanations), non-moral philosophical reasoning (about their epistemic implications) and the minimal moral phenomenology required to identify kinds of moral

⁵⁵ This is very loosely inspired by a case about color vision introduced by Moon (2017: 219f.) and discussed by Klenk (2017: 238-40). Importantly, their case doesn’t feature partial reliability of color vision – and doesn’t spell out the analogy with moral belief at all.

intuitions. None of those cognitive process has been negatively affected by evolutionary influence.⁵⁶

Second, **YY Color Defeater-Deflector** above has yet another crucial, redemptive feature: once you rely on the scientific testimony, you learn that not all color beliefs are suspect. The pill actually only defeats blue and green color beliefs, while leaving red ones undefeated. Similarly, once robust moral realists take the typology of moral intuitions (and their associated epistemic risks) into account, they recognize that not all moral intuitions are dubious. Rather, noxious evolutionary influence only concerns concrete and mid-level moral intuitions, while formal moral intuitions are off the hook. That, in turn, means that robust moral realists don't occupy epistemic conditions as dire as **XX Defeater-Defeater**. In fact, they have access to a defeater-deflector, namely beliefs based on formal moral intuitions. And if those moral beliefs could play the role of a third factor, the resulting explanation would not beg the question against the evolutionary debunker. So, the worry introduced in §4.3.1 above proves unfounded.

4.3.3. A Complication: Contesting the Origins of Moral Reliability

Before wrapping up my argument that third-factor explanations do not always beg the question, let me address one complication. You might issue the following complaint: the typology of moral intuitions (and their associated epistemic risks) only shows that a certain sub-set of moral beliefs are unlikely to be defeated by evolutionary influence. But, importantly, the typology doesn't establish that those beliefs are *reliable*.⁵⁷ (Analogously, we might ask in the **YY Color Defeater-Deflector** introduced above: why think that undefeated red beliefs are reliable in the first place?) So, some distinctively non-moral *vindication* of moral reliability seems still required. But, you might continue, it is hard to see where that would come from.⁵⁸ For instance, take evolutionary considerations. While such considerations might help vindicate perceptual, logical or epistemological beliefs as reliably formed, the same cannot be said for robustly moral beliefs.⁵⁹ After all, as we have

⁵⁶ The capacity for moral phenomenology (i.e. seemings with distinctively moral contents) is most likely adaptive, even if having *veridical* seemings or corresponding *true* robustly moral beliefs is not.

⁵⁷ Klenk (2017: 238-42) seems to voice a complaint along those lines. He argues that robust moral realists can employ undefeated sources of moral belief as a defeater-deflector '...only if they can tap into a source of information that is both distinct from (i.e. not a product of) the deliverances of moral cognition and yet indicative of the reliability of moral cognition' (Klenk 2017: 240). However, it remains unclear whether Klenk criticizes as nuanced a proposal as the one I developed above. So, the following complaint is a charitable reconstruction.

⁵⁸ See Chapter 2 for what makes explaining moral reliability tricky.

⁵⁹ See de Cruz et al. 2011 for an evolutionary vindication of epistemology, Schechter 2013 for logic and Boudry & Vlerick 2014 for perception.

heard by now, evolutionary explanations of moral beliefs don't avert to their truth. In addition, robust moral realists typically don't accept empirical vindications anyway, given their commitment to robustly moral properties.⁶⁰ So, it looks like robust moral realists cannot just appeal to formal moral intuitions as defeater-deflectors. Rather, they face a further, significant hurdle: explaining why those moral intuitions are reliable in the first place.

On reflection, however, this complaint can be dismissed rather straightforwardly. First, the robust moral realist might reject the demand for an evolutionary vindication of moral reliability as *dialectically inappropriate*. To argue for that, we distinguish two kinds of tasks in epistemological theorizing: vindicating a source of information – and merely showing that it remains undefeated. Importantly, our context seems to require only the latter, not the former. To see that, recall **XX Defeater-Deflector**: initially, you consider yourself cognitively reliable. Then, there is a potential defeater for cognitive reliability, namely the XX-pill. Finally, before ingesting the pill, you receive trustworthy scientific testimony that the defeater won't apply to you. Importantly, the scientific testimony doesn't have to establish or vindicate your cognitive reliability. Rather, it only has to inform you that the defeater won't apply. Similarly, consider the epistemic situation that robust moral realists find themselves in, namely a moral analogue of the **YY Color Defeater-Deflector** developed above. To wit, the typology of moral intuitions (and their associated epistemic risks) results from a combination of scientific theorizing, non-moral philosophical thinking as well as moral phenomenology. Importantly, those forms of reasoning, themselves undefeated and potentially evolutionarily vindicated, don't need to establish that formal moral intuitions are *reliable*. Rather, they only have to make the case that the evolutionary defeater isn't likely to apply – which they seem to do in compelling fashion. So, the complaint that formal moral intuitions haven't been shown to be reliable risks conflating an important distinction between vindication and showing the absence of defeat. As a result, the complaint formulates a demand that is dialectically unfitting.

Second, suppose that the demand for vindicating moral reliability is legitimate, contrary to what I have just argued. Even then, the robust moral realist isn't lost, for she can plausibly accommodate the demand. After all, it is very likely that formal moral intuitions – unlike concrete and mid-level ones – are the product of a more general capacity for rational reflection or *apriori* reasoning. The reliability of that capacity, in turn, can be vindicated by appeal to evolutionary considerations.⁶¹ Either way, then, the robust moral

⁶⁰ See Klenk 2017: 241.

⁶¹ For more on that line of defense, see Fitzpatrick 2015 & 2016b.

realist can address the complaint that formal moral intuitions aren't demonstrably reliable.⁶²

4.3.4. Closing the Gap: From Formal Moral Intuitions to Third Factors

At last, I am in a position to complete my defense of third-factor explanations in the face of the begging the question charge. Above, I have argued that, when constructing their explanation, the robust moral realist must appeal to formal moral intuitions exclusively. That alone, however, won't get them home dry. After all, they cannot just appeal to *any* formal moral intuition. Rather, they must find formal moral intuitions that motivate a corresponding normative claim fit to play the role of third factor. In particular, that claim must not just derive from purely formal moral intuitions, but also be plausibly shaped by evolutionary forces (or be consistent with the overall aim of evolution). So, what might those formal moral intuitions in question be? To address that issue, Jeff Behrends (2013: 493-6) has recently put forward the following argument:

- (5) If we have reason to do something or other, then we have reason to pursue the means necessary to doing it.⁶³
- (6) Our existence is necessary for our doing anything.
- (7) Therefore, if we have reason to do something or other, then we have reason to pursue our own existence. (1, 2).
- (8) If robust normative realism is true, then we have reason to do something or other.
- (9) Therefore, if robust normative realism is true, then we have reason to pursue our own existence. (3, 4).

Importantly, both key premises of the argument, namely (1) and (2), are based on purely *formal* moral intuitions. Neither generate substantive moral claims or evaluate actions, states of affairs or character traits. Rather, they impose constraints on the *form* of any given moral theory. For instance, (1) '...seems to be required by the very nature of normative reasons, by the concept of something's being a reason' (495). If a theory didn't accept (1), it would hardly qualify as an action-guiding system of moral rules. Similarly, (4) is

⁶² Of course, a third option would be to re-run my argument from Chapter 2: the challenge to explain why formal moral beliefs are reliably formed – construed modally – boils down to a more general worry about epistemic luck.

⁶³ Behrends (2013: 494f.) mentions two caveats. First, the (instrumental) reason isn't all-things-considered, yet could be outweighed by competing demands. Second, since it doesn't invoke desires, it isn't Humeanism – and is, in fact, neutral about the grounds of normative reasons.

uncontroversial: it is equivalent to the conditional that ‘if [robust normative] realism is true, then there exists at least one fact that counts in favor of some action or other on our part’ (Behrends 2013: 495). That, however, simply expresses the core metaphysical commitment of robust normative realism, which we can take for granted in this context. After all, the best evolutionary explanation argument grants that robust moral facts *exist*, yet questions whether we can ever justifiably form beliefs about them. Together, (1) – (4) entail that, given the truth of robust normative realism, we have reasons to pursue our own existence or survival. That normative claim, in turn, can serve as a plausible third factor: a robustly normative truth that evolutionary forces would have pushed us towards. Or, in Behrends’ (2013) words: ‘...evolutionary forces have pushed us toward holding the belief that we have reason to pursue our own survival, and as it turns out, the truth of [robust normative] realism entails that we do have such a reason’ (494). So, the conclusion of the argument above offers exactly what we were looking for: a robustly normative claim, derived from purely formal moral intuitions, yet also plausibly shaped by evolutionary forces. Of course, the argument might not be sound as it stands. In particular, (2) might require some finessing, as even Behrends readily admits.⁶⁴ Still, it suffices to illustrate my point: that it is *in principle* possible to derive a third factor from purely formal moral intuitions. The gap can be closed. And if that is possible, robust moral realists can construct a non-question-begging third-factor response.

Where does that leave us? In this sub-section, I have argued that third-factor explanations don’t beg the question – as long as they rely exclusively on formal moral intuitions. To establish that, I first clarified the issue by distinguishing defeater-deflectors from defeater-defeaters. Then, I suggested that formal moral intuitions can serve as defeater-deflectors, given that they are likely exempt from noxious evolutionary influence. After addressing a complaint about the reliability of these intuitions, I concluded by deriving a third factor from purely formal moral intuitions.

But does that mean that robust moral realists can respond to the best evolutionary explanation argument? In the remainder of this chapter, I argue that it does not: on closer examination, the non-question-begging third-factor explanation just developed cannot underwrite belief about the *substance* of morality, even when combined with rational reflection.

⁶⁴ For further discussion, see Behrends 2013: 495-98.

4.4 Rational Reflection, Formal Moral Intuitions and The Scope of Intuitive Moral Knowledge

Suppose that the third-factor explanation developed above does indeed not beg the question. In that case, we have normative reasons to pursue our own survival. But what are those reasons exactly? Typically, third-factor explanations assume – rather optimistically – that *rational reflection* can answer that question. By thinking rationally, we can bridge the gap between the third factor and substantive knowledge of robustly normative (or moral) reasons. In this sub-section, my aim is to argue that this optimism is unfounded: formal moral intuitions cannot underwrite many beliefs about the *substance* of morality, even when combined with rational reflection. That, however, is bad news for robust moral realists: it forces them into moral scepticism. My argument starts with some stage-setting (§4.4.1), before rejecting extreme forms of optimism (§4.4.2) as well as pessimism (§4.4.3) about rational reflection. It concludes by settling for moderate pessimism about our rational powers (§4.4.4) – and highlighting its sceptical implications for robustly moral knowledge.

4.4.1. Third-Factor Explanations and Rational Reflection

Suppose that, for the sake of argument, the third-factor explanation developed above is plausible. Then, robust moral realists or moral intuitionists still haven't quite averted moral scepticism yet. After all, the explanation only gives us an approximately true normative starting point, namely that there are reasons to pursue our own existence. *What* those reasons are, however, the explanation doesn't say. To bridge the gap between the third factor and substantive knowledge of those normative reasons, proponents of third-factor explanations appeal – either explicitly or implicitly – to *rational reflection* or *reflective equilibrium*. Most prominently, David Enoch (2012) writes that '...given a starting point of normative beliefs that are not too far-off, presumably some reasoning mechanisms (and perhaps some other mechanisms as well) can get us increasingly closer to the truth by eliminating inconsistencies, increasing overall coherence, eliminating arbitrary distinctions, drawing analogies, ruling out initially justified beliefs whose justificatory status has been defeated later on, etc.' (166). In a similar vein, Jeff Behrens (2013) argues that '...more normative knowledge can be gained so long as it is plausible that our knowing that we have reason to pursue our own persistence (partially) justifies us in making inferences to further normative claims' (498). So, rational reflection enables us to see how the normative truth that plays the role of third factor makes up, entails or informatively coheres with other normative truths. In that manner, robust moral realists can allegedly underwrite significant amounts of robustly normative knowledge.

But is that optimism about our reflective powers warranted? In what follows, I shall address that question, both generally and specifically in light of the appeal to formal moral intuitions discussed above. To begin with, I shall look at and ultimately reject more extreme answers, before settling for a form of moderate pessimism that robust moral realists must find unacceptable.

4.4.2. Radical Optimism

Jeff Behrens (2013) subscribes to *radical optimism*: rational reflection is the right tool for discovering robust moral truths – and takes us from the third factor (nearly) all the way to substantive moral knowledge. In more detail, he discusses the objection that his third-factor explanation doesn't gain much epistemological ground and forces robust moral realists into '...an objectionably impoverished amount of normative knowledge' (2013: 498). In response, he articulates two kinds of consideration.

First, he sketches out how to *generate* further normative knowledge out of the knowing that we have reason to pursue our continued existence. For instance, we might be justified in inferring that we have normative reason to pursue those means necessary to securing our persistence. That, in turn, allows us to arrive at many normative truths: that we have reason to exercise and eat well, avoid danger, co-operate with others, and so forth. In that manner, we could construct a crude version of social contract theory exclusively out of the third factor that we have reasons to survive, together with empirical claims about our psychology and surroundings.

Second, Behrens emphasizes that it is easy to underestimate the epistemic significance of his third-factor explanation. After all, the evolutionary challenge (in the shape of the best evolutionary explanation argument, in our case) formulates a potential undercutting defeater for any positive realist moral epistemology, including the kind of moral intuitionism developed in §2 and §3 above. His third-factor explanation, if plausible, prevents that potential undercutting defeater from actualizing. But then, '...realists get to help themselves [absent further objections] to the justificatory story that they take to be most plausible [e.g. moral intuitionism]' (2013: 499). And it is this justificatory story that primarily explains why we have reason to think that the normative facts are one way rather than another. So, the non-question-begging third-factor explanation amounts to a highly significant epistemological achievement: it allows robust moral realists to simply proceed with *first-order normative theorizing*, whatever that looks like. (In our case, moral intuitionism would play a crucial role in that endeavor.) And it is that theorizing that gives us epistemic reason for thinking that the normative facts have a particular content, as opposed to any other.

However, I worry that Behrends' radical optimism is misplaced for two reasons. To begin with, the normative knowledge inferred from Behrends' third factor, namely that we have reasons to survive, still seems *insular* or *limited* in crucial respects. After all, whatever it entails seems to primarily cover self-regarding moral duties, while leaving open other-regarding moral duties (beyond the mere moral demand to cooperate with others to survive) and how to balance the former with the latter. At this stage, Behrends might suggest that first-order moral epistemology and normative theorizing can help resolve this under-determination. But I remain sceptical: as we saw above, to avoid begging the question, the proponent of third-factor explanations may only rely on formal moral intuitions. That, though, will significantly constrain both first-order moral epistemology and normative theorizing, making it difficult to comprehensively specify the content of first-order moral theory. So, the normative knowledge recaptured still seems limited or insular.

My second reason for doubting Behrends' radical optimism has to do with Werner's (2018) *Content Success* requirement: '...at least some of our justified, first-order, and paradigmatically normative beliefs contain robustly normative contents' (2018: 619). In particular, it remains questionable whether Behrends' third-factor helps moral intuitionists meet that desideratum. To see that, recall the disagreement between Lucy and Carol. While one believes that lying is intrinsically bad, the other holds the contrary. Now, it seems to me that Behrends' third factor cannot resolve that disagreement and tell us whose concepts latch onto the robust moral truth. After all, that we have reasons to survive seems consistent with either stance – as long as they both agree that lying is often *instrumentally* bad. But if that is correct, Behrends' strategy cannot meet *Content Success*: when it comes to many disagreements about mid-level and concrete moral matters, it fails to help us figure out whose intuitively justified moral beliefs have robustly normative contents. In conclusion, Behrends seems to underestimate the *revisionary* nature of his proposal with respect to common sense morality. Therefore, his radical optimism about rational reflection seems unwarranted.

4.4.3. Radical Pessimism

In stark contrast, some doubt that rational reflection is *ever* the right tool for discovering robust moral truths, at least in the context of evolutionary debunking. They subscribe to – what we might call – *radical pessimism* about rational reflection. Max Hayward

(forthcoming) has recently put forth an argument in support of this position.⁶⁵ He understands rational reflection as the method according to which ‘...[w]e revise our opinions by using some beliefs to evaluate others, checking for consistency between individual judgments, and searching for greater coherence and systematicity within our belief-set as a whole’ (2017: 1). Then, he argues that this method ‘...shouldn’t be expected to guide us to truth in ethics, even from somewhat correct starting points’ (2017: 2). Why? His pessimism seems based on two lines of criticism.⁶⁶

First, he takes issue with those who hold – like Thomas Scanlon (2014: 82, 84) and others – that rational reflection or reflective equilibrium only applies to ‘considered judgments’ or intuitively ‘plausible’ beliefs. After all, our intuitive assessments of plausibility are just as likely to be unreliable as our moral beliefs, given their evolutionary history. In fact, such a capacity for plausibility verdicts might well be *counter-adaptive*, given that it is often adaptive to have false moral beliefs. So, we must expect our plausibility judgments to be a ‘mixture of truth and error’.

Second, even if, following Kagan (1989), rational reflection is about searching for *coherence* (beyond mere logical consistency) as well as other *theoretical virtues* (e.g. systematicity, simplicity, explanatoriness), there is no *apriori* reason to think that the robust moral truth must mirror those theoretical virtues.⁶⁷ To insist that robust moral truth must be coherent and conflict-free just begs the question: we must assume that meta-moral beliefs about the structure of the moral truth are a mixed bag of truths and falsehoods, just like moral beliefs and plausibility judgments. Therefore, there is ‘...no reason to expect any methodology of rational reflection that goes further than simply avoiding logical contradiction to bring us closer to the truth’ (Hayward forthcoming: 11).

⁶⁵ A note on the dialectical background: Hayward discusses the role of rational reflection against the backdrop of the *modest* evolutionary debunking argument. According to that argument, we cannot assume that biological and cultural evolution gave us totally reliable dispositions, given that moral truth didn’t causally regulate evolution and natural selection favors some immoral dispositions (e.g. nepotism). Rather, our moral starting points are probably error-riddled and our dispositions only partially reliable, or that ‘we should expect to start with a pretty mixed bag [with respect to moral truths and falsehoods]’ (Hayward forthcoming: 6). Still, to avoid general scepticism, we may assume that moral truth is roughly as we take it to be. Hayward’s dialectical situation seems comparable to ours: Behrends’ (2013) third-factor guarantees that the moral truth is roughly as we take it to be, but it remains an open question whether we can reclaim substantive moral knowledge.

⁶⁶ A third line of criticism concerns another account of rational reflection, namely conceived as a method for internal value assessment and resolution of inconsistencies. For more, see Hayward forthcoming: 8-10. However, since that doesn’t appear to be the method of reflection that proponents of third-factor explanations have in mind, I shall set it aside for now.

⁶⁷ See Griffin 2015 against simplicity, Williams 1973: Ch. 11 on moral coherence and conflict, or Scanlon 2014 on moral indeterminacy.

However, Hayward's pessimism strikes me as too extreme. To begin with, it is unclear whether our intuitive assessments of plausibility are just as likely to be unreliable as our moral beliefs, given their evolutionary history. After all, as we saw in §4.3 above, formal moral intuitions (and corresponding judgments of plausibility) are less likely to be the product of problematic evolutionary pressure. In fact, they might very well originate from the same general *apriori* reasoning capacities as logic, mathematics and modality, which hardly count as counter-adaptive.

Further, I am ready to admit that there is probably no *apriori* reason to think that the robust moral truth must mirror theoretical virtues such as coherence, systematicity or simplicity. That, however, doesn't mean that there are no reasons *whatsoever* to think that human morality has a certain structure. To see that, just consider the complex defence of formal moral intuitions as exempt from noxious evolutionary influences rehearsed above, based on scientific theorizing, non-moral philosophical reasoning and some moral phenomenology. That defence should offer a reason to assume that meta-moral beliefs about the structure of moral truth are *not* just a mixed bag of truths and falsehoods. In addition, suppose that we are not allowed to make *any* assumptions about the structure of robust moral truth. In that case, we seem to just invite global moral scepticism – which is something that Hayward is at pains to avoid himself. In short, it looks like Hayward must grant us some formal constraints on morality. Given that both lines of criticism that Hayward develops can be dealt with, his radical pessimism about rational reflection appears unreasonably extreme.

4.4.4. Moderate Optimism – or Pessimism?

That leaves us with more moderate views about the powers of rational reflection. One such view is *moderate optimism*, as presented in Huemer (2008). Even though he doesn't defend a third-factor explanation, his account of moral reasoning is highly relevant to our purposes. According to him, rational reflection is the right tool, at least when guided by formal moral intuitions, for discovering *some* substantive truths about robust morality. Yet, importantly, rational reflection cannot take us all the way towards common sense morality – or towards *all* the items of robustly moral knowledge that robust moral realists want to take for granted. So, unlike Hayward (forthcoming), Huemer puts faith in the powers of rational reflection. But unlike Behrens (2013), he doesn't consider them all-conquering.

In more detail, Huemer advocates shifting the methods of ethics from narrow reflective equilibrium towards – what he calls – 'cautious critical intuitionist methodology' (2008: 381). That methodology lays out the following criteria for moral intuitions:

- (1) Seek a substantial, coherent body of moral intuitions;
- (2) Avoid intuitions that are only shared by a few, culturally specific and patriotic;
- (3) Avoid intuitions that favor reproductive fitness, especially if they don't cohere with other intuitions;
- (4) Avoid intuitions that differentially favor oneself;
- (5) Avoid intuitions that line up with strong emotions.

As we saw above, once we adhere to those criteria, we end up privileging formal moral intuitions, since they are most likely to avoid various biases. Unlike concrete or mid-level ones, those intuitions don't positively or negatively evaluate anything. Yet, they impose formal constraints on ethical theories. And we can then assess combinations of moral commitments in light of them, thus resolving moral disagreements. According to Huemer, an account of rational reflection informed by both formal constraints and the rejection of unreliable moral intuitions is likely to be *revisionary*. How far that moral revision will go remains an open question. Still, Huemer singles out sexual morality and deontology as two moral issues ripe for a shake-up. After all, common moral intuitions about sexuality are susceptible to emotional, cultural and biological biases. So, '...one's initial, intuitive opposition to those arrangements or forms of sexual activity ought not to be treated as serious evidence of their wrongness' (Huemer 2008: 388). And similar diagnoses might be developed for core deontological commitments.⁶⁸ Thus, Huemer seems *moderately optimistic* about the prospects of rational reflection. Reliance on moral reasoning helps us gain some moral knowledge, yet significantly less than the overwhelming majority of robust moral realists typically take for granted. So, if *moderate optimism* is correct, robust moral realists cannot justifiably believe many robustly moral facts. That, by itself, should worry them.

However, I suspect that Huemer's optimism, despite already being tempered, might still be *too excessive*. To see why, let's briefly recapitulate the three key steps of Huemer's proposed moral methodology: (1) eliminate questionable or probably unreliable concrete and mid-level moral intuitions (based on the criteria above); (2) generate views (on various moral issues) based on the residual intuitions; (3) appeal to formal moral intuitions to adjudicate between competing or conflicting views. There are three kinds of worries associated with this procedure. To begin with, (2) might be problematic, especially if the residual intuitions don't feature enough *content* to generate informative views on the key moral questions of that domain. For instance, take Huemer's own example, namely sexual

⁶⁸ For a sketch, see Huemer 2008: 389f.

morality. It could well be the case that, once are done rejecting moral intuitions about sexual activity with dubious pedigree, that few moral intuitions will be left to inform our view on the domain. And that seems problematic, given the domain's significance to leading a good human life. Second, there might be trouble with (3). More specifically, formal moral intuitions might be too formal or permissive to help us adjudicate between competing views on a given moral issue. To see why and for illustration, it pays to revisit Werner's (2018) conflict between Lucy and Carol about the intrinsic badness of lying. Formal moral intuitions might just not tell us who is right on that score. Third and finally, suppose the procedure outlined above is successful, i.e. we have gone through steps (1) – (3) without fail, setting the two worries just raised aside. Even then, we might end up in an *epistemically sub-optimal* situation: the moral intuitions that provide a disproportional amount of content to a given moral view will always be the less formal ones – and thus the less reliable ones. It might well be that we end up with views about a given moral domain that are based on barely acceptable moral intuitions, but cannot be ruled out by our moral certain formal moral intuitions.

Together, these three worries suggest that Huemer's methodology, on closer inspection, isn't just revisionary, but properly *sceptical* (and surely more sceptical than he is prepared to admit). If that is correct, however, we shouldn't be moderate optimists about rational reflection. Rather, we should be *moderate pessimists*: given the robust normative truth that we have reasons to survive, we won't be able to recover much *substantive* moral knowledge. To do so would require relying on concrete or mid-level moral intuitions, yet most of those are epistemically suspect. Therefore, rational reflection is a good tool for discovering *some* robust moral truths, yet decidedly less than robust moral realists usually want to take for granted.

4.4.5. What It All Means

Where does that leave us? In this sub-section, I examined whether rational reflection allows us to bridge the gap between the third factor derived in §4.3 above, namely that we have normative reasons to pursue our own survival, and substantive knowledge of further robustly normative (or moral) reasons. As I argued, we should be *moderate pessimists*: even assuming we know the supposedly robust normative truth that we have reasons to survive, we won't be able to recover much *substantive* moral knowledge. After all, doing so would require relying on concrete or mid-level moral intuitions, yet most of those are epistemically suspect, for precisely the reasons debunkers argued. So, the third-factor explanation developed above might not beg the question against the evolutionary debunker, but it fails to underwrite many substantive moral beliefs, even when combined with rational reflection.

What does that mean for the fate of robust moral realism more generally? My discussion shows that even the most compelling version of moral intuitionism, supplemented by a non-question-begging third-factor explanation for moral reliability, fails to identify additional, non-explanatory reasons for much beyond our beliefs about the *form* of morality. Thus, the view cannot stave off the sceptical worry raised by best evolutionary explanation argument in Chapter 4. By implication, robust moral realism is saddled with unacceptable moral scepticism: if they are conceived as metaphysically robust, we cannot justifiably believe many facts about the *substance* of morality.

5. Conclusion

This chapter concludes my comprehensive, three-step assessment of evolutionary debunking arguments against robust moral realism. Chapter 1 took the first step, distinguishing three independently plausible epistemic principles that, combined with evolution, seemed to render knowledge of objective moral facts problematic. These principles were informed by cutting-edge debates in contemporary epistemology, from the nature of reliability to the epistemic significance of higher-order evidence and the relevance of explanation to justification. Chapters 2 to 4 then took the second step, developing and critically assessing refined evolutionary debunking arguments based on those epistemic principles. My discussion showed that, although not all forms of the evolutionary debunking argument are successful, robust moral realism faced two distinct epistemic worries. (The worry about irrelevant evolutionary influence on robustly moral belief discussed in Chapter 3 did not prove compelling.) The first, developed in Chapter 2, was loosely inspired by evolutionary debunking arguments and centered on knowledge-undermining epistemic luck. Meanwhile, the second, as discussed in Chapter 4, was directly raised by evolutionary considerations, focusing on whether robust moral realists have compelling non-explanatory reasons for maintaining their moral beliefs.

This chapter amounted to the third and final step. I argued that, while robust moral realism can tackle the worry about epistemic luck, it cannot offer compelling non-explanatory reasons for robustly moral belief. As a result, the evolutionary debunking argument developed in Chapter 4 succeeds, saddling robust moral realism with implausible moral scepticism: we cannot justifiably believe robustly moral facts. To establish that conclusion, I closely examined moral intuitionism, the most popular and best-developed moral epistemology available to robust moral realism. First, I demonstrated that moral intuitionism isn't vulnerable to the worry about epistemic luck developed in Chapter 2. Still, its immunity helped me uncover the most compelling version of the view, according to which adequate understanding serves as a necessary anti-

luck condition on intuitive moral knowledge. Then, I argued that even this version of moral intuitionism cannot stave off the worry raised by best evolutionary explanation argument in Chapter 4. After all, the view required a third-factor explanation of moral reliability: by itself, moral intuitionism doesn't provide non-explanatory reasons for belief in robust moral facts. However, any such explanation must depend on formal moral intuitions only. Otherwise, it risks begging the question against the evolutionary debunker. But these formal intuitions, as it turned out, cannot underwrite many substantive moral beliefs, even when combined with rational reflection. Therefore, moral intuitionism cannot give us additional, non-explanatory reasons for anything beyond our beliefs about the *form* of robust morality. Robust moral realist thus become moral sceptics: they cannot justifiably believe facts about the *substance* of robust morality.

But suppose that you are not entirely convinced by this chapter's argument. Even then, my discussion holds some valuable insights. To conclude, I shall briefly elaborate on three of them. First, by tackling the issue of epistemic luck raised in Chapter 2, §3 developed the most compelling version of moral intuitionism. On the resulting view, you know an intuitive moral proposition *p* if you truly believe that *p*, justified on the basis of a veridical intellectual seeming that *p*, which is grounded in adequately understanding *p*. Importantly, the view features adequate understanding as a separate necessary anti-luck condition on intuitive moral knowledge. By doing so, this version of moral intuitionism avoids complicating moral justification and honors the difference between justification and knowledge. So, even if they don't accept my radical, sceptical conclusion, there is a lesson here for moral intuitionists: they should endorse the version of their view developed in §3.

Second, §4 also featured two lessons about how to most effectively defend third-factor explanations of moral reliability. To begin with, such explanations do not beg the question, contrary to what many critics allege. But to avoid doing so, the third factor can only be derived from formal moral intuitions that are likely exempt from distorting evolutionary pressures. So, this is the structure such an explanation should take.⁶⁹ That, in turn, is helpful for any robust moral realist turning towards a third-factor explanation in the face of the best evolutionary explanation argument (as I argue they should). In addition, my discussion indicated how to complete this defense of third-factor explanations. In particular, such explanations must be combined with an argument

⁶⁹ This also means that – contra Behrends (2013) and Enoch (2010, 2012) – third-factor explanations aren't plausibly neutral with respect to first-order moral epistemologies. Rather, they are best paired with a form of moral intuitionism that allows us to distinguish between concrete, mid-level and abstract moral intuitions.

showing how rational reflection can systematically bridge the gap between formal moral beliefs and ordinarily held beliefs about the substance of morality. So, developing such an account of moral reasoning is what robust moral realists should focus their resources on.

Third and finally, §4 contributed towards the most accurate taxonomy of responses to evolutionary debunking. Typically, contributors to the literature distinguish between exemption and alignment strategies.⁷⁰ While exemption strategies aim to show that robustly moral beliefs are exempt from noxious evolutionary influence, alignment approaches such as third-factor explanations aim to demonstrate that evolved moral beliefs – against first appearances – actually align with the robust moral truth. However, my discussion of third-factor explanations in §4 gives us reason to believe that there is no deep distinction along those lines. Rather, the best strategy available to the robust moral realist against the evolutionary debunker combines an alignment response, namely a third-factor explanation, with an exemption response, namely an appeal to formal moral intuitions. So, in light of my discussion, we should revise how we ordinarily taxonomize responses to evolutionary debunking arguments.

But what are the implications of my discussion in Chapters 2 to 4 for meta-ethics and epistemology more generally? In Chapter 1, I mentioned that my overall argument should also be more broadly relevant beyond evolutionary debunking and robust moral realism. After all, the three general epistemic principles framing my investigation of evolutionary debunking arguments are central to any foundational debate in the epistemology of the normative and, more generally, the *apriori*. Further, applying these general epistemic principles extensively to the evolutionary debunking of robust moral belief promised ways of *refining* them. So, what insights into how to formulate those principles most compellingly can we glean from Chapters 2 to 4? To begin with, consider the principle that reliability about a given domain requires explanation – and that, if we fail to provide one, our beliefs about that domain are unjustified. Chapter 2 taught us that the most plausible interpretation of the principle threatens knowledge, not justification. What's more, 'explaining reliability' doesn't really amount to showing that our moral, normative or *apriori* beliefs are true. Rather, it involves demonstrating why a given account of moral, normative or *apriori* knowledge isn't vulnerable to epistemic luck. Further, Chapter 3 held a lesson on how to properly apply the principle that higher-order evidence of error undermines justification. As we saw, the principle defeats itself when the higher-order evidence of error in question concerns all beliefs about a given domain. That suggests that we should restrict its application to kinds of higher-order evidence of error that target

⁷⁰ See, for instance, Fitzpatrick 2016b or Wielenberg 2016.

particular beliefs or beliefs about sub-domains only.⁷¹ For instance, we could rely on the principle when assessing higher-order evidence of error about specific color beliefs, yet not all color beliefs. Finally, Chapter 4 focused on the principle that, for our beliefs to be justified, our having them must be best explained by the facts they are about. My discussion suggested how to best formulate that principle, too. In particular, we must supplement it with a clause stating that we lack additional, non-abductive reasons for belief in moral, normative or *apriori* facts. Otherwise, the principle would beg the question against any non-abductive moral, normative or *apriori* epistemology. Of course, this is but a rough sketch. Still, it illustrates that my overall argument informs how to most plausibly formulate those general epistemic principles. As a result, the dialectic in Chapters 2 to 4 should be of value to anyone working on the epistemological foundations of the normative and, more generally, the *apriori*.

⁷¹ Vavova (2014) argues that debunking arguments based on that principle are governed by – what she calls – the *Inverse Rule of Debunking*: ‘The potential strength [i.e. how much it requires us to revise] of a debunking argument is inversely proportional to its ambition [i.e. how much it targets]’ (98). She recommends – in line with my suggestion in the main text – that we thus focus on local, not global, defeat by higher-order evidence of error.

BIBLIOGRAPHY

- Audi, Robert (1999). Self-evidence. *Philosophical Perspectives* 13 (s13): 205-228.
- Ballantyne, Nathan (2013). The Problem of Historical Variability. In: *Disagreement and Skepticism* 46: 239-59.
- Barkhausen, Max (2016). Reductionist Moral Realism and the Contingency of Moral Evolution. *Ethics* 126 (3): 662-689.
- Bealer, George (1999). A Theory of the Apriori. In James Tomberlin (ed.), *Epistemology*. Blackwell. 29-56.
- Becker, Kelly & Black, Tim (eds.) (2012). *The Sensitivity Principle in Epistemology*. Cambridge University Press.
- Bedke, Matthew S. (2009). Intuitive non-naturalism meets cosmic coincidence. *Pacific Philosophical Quarterly* 90 (2): 188-209.
- Bedke, Matthew S. (2014). *No Coincidence?.* Oxford Studies in Metaethics. Oxford University Press.
- Behrends, Jeff (2013). Meta-normative Realism, Evolution, and Our Reasons to Survive. *Pacific Philosophical Quarterly* 94 (4): 486-502.
- Benacerraf, Paul (1973). Mathematical Truth. *Journal of Philosophy*, Vol. 70: 661-679.
- Berker, Selim (2014). Does Evolutionary Psychology Show That Normativity Is Mind-Dependent? In: Justin D'Arms and Daniel Jacobson (eds.), *Moral Psychology and Human Agency: Philosophical Essays on the Science of Ethics*. Oxford University Press.
- Besson, Corine (2009). Logical knowledge and Gettier cases. *Philosophical Quarterly* 59 (234): 1-19.
- Björnsson, Gunnar & Francén Olinder, Ragnar (2016). Enoch's Defense of Robust Meta-Ethical Realism. *Journal of Moral Philosophy* 13 (1): 101–112.

Bogardus, Tomas (2016). Only All Naturalists Should Worry About Only One Evolutionary Debunking Argument. *Ethics* 126 (3): 636-661.

Boghossian, Paul (2000). Knowledge of Logic. In: Paul Boghossian & Christopher Peacocke (eds.), *New Essays on the A Priori*. Oxford University Press.

Boudry, Maarten & Vlerick, Michael (2014). Natural Selection Does Care about Truth. *International Studies in the Philosophy of Science* 28 (1): 65-77.

Brogaard, Berit & Salerno, Joe (2013). Remarks on Counterpossibles. *Synthese* 190 (4): 639-660.

Brosnan, Kevin (2011). Do the evolutionary origins of our moral beliefs undermine moral knowledge? *Biology and Philosophy* 26 (1): 51-64.

Chalmers, David J. (2009). The Two-Dimensional Argument Against Materialism. In: Brian P. McLaughlin & Sven Walter (eds.), *Oxford Handbook to the Philosophy of Mind*. Oxford University Press.

Chrisman, Matthew (2016). *What is This Thing Called Metaethics?*. Routledge.

Christensen, David (2007). Epistemology of disagreement: The good news. *Philosophical Review* 116 (2): 187-217.

Christensen, David (2009). Disagreement as evidence: The epistemology of controversy. *Philosophy Compass* 4 (5): 756-767.

Christensen, David (2010). Higher-Order Evidence. *Philosophy and Phenomenological Research* 81 (1): 185-215.

Christensen, David (2011). Disagreement, Question-Begging, and Epistemic Self-Criticism. *Philosophers' Imprint* 11 (6).

Cline, Brendan (2016). Against deliberative indispensability as an independent guide to what there is. *Philosophical Studies* 173 (12): 3235-3254.

Cogburn, Jon & Roland, Jeffrey (2011). Anti-Luck Epistemologies and Necessary Truths. *Philosophia* 39 (3): 547-561.

- Comesaña, Juan (2005). Unsafe knowledge. *Synthese* 146 (3): 395 - 404.
- Copp, David (2008). Darwinian skepticism about moral realism. *Philosophical Issues* 18 (1): 186-206.
- Clarke-Doane, Justin (2012). Morality and Mathematics: The Evolutionary Challenge. *Ethics* 122 (2): 313-340.
- Clarke-Doane, Justin (2013). Moral Epistemology: The Mathematics Analogy. *Noûs*.
- Clarke-Doane, Justin (2015). Justification and Explanation in Mathematics and Morality. Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics, Vol. 10*.
- Clarke-Doane, Justin (2016). Debunking and Dispensability. In: Uri D. Leibowitz & Neil Sinclair (eds.), *Explanation in Ethics and Mathematics: Debunking and Dispensability*. Oxford University Press.
- Clarke-Doane, Justin (2017). What is the Benacerraf Problem? In: Fabrice Pataut (ed.), *New Perspectives on the Philosophy of Paul Benacerraf: Truth, Objects, Infinity*.
- Clarke-Doane, Justin (forthcoming). Debunking Arguments: Mathematics, Logic, and Modal Security. In: Robert Richards and Michael Ruse (ed.), *Cambridge Handbook of Evolutionary Ethics*. Cambridge University Press.
- Cohen, G. A. (2000). If you're an egalitarian, how come you're so rich? *Journal of Ethics* 4 (1-2): 1-26.
- Cohen, Stewart & Comesaña, Juan (2013). Williamson on Gettier Cases and Epistemic Logic. *Inquiry : An Interdisciplinary Journal of Philosophy* 56 (1): 15-29.
- Colyvan, Mark (2015). Indispensability Arguments in the Philosophy of Mathematics, *The Stanford Encyclopedia of Philosophy* (Spring 2015 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2015/entries/mathphil-indis/> (August 24, 2018).
- Cowie, Christopher (2014). Why Companions in Guilt Arguments Won't Work. *Philosophical Quarterly* 64 (256): 407-422.

Cowie, Christopher (2016). Good News for Moral Error Theorists: A Master Argument Against Companions in Guilt Strategies. *Australasian Journal of Philosophy* 94 (1): 115-130.

Cuneo, Terence (2007). *The Normative Web: An Argument for Moral Realism*. Oxford University Press.

Das, Ramon (2017). Bad News for Moral Error Theorists: There Is No Master Argument Against Companions in Guilt Strategies. *Australasian Journal of Philosophy* 95 (1): 58-69.

Das, Ramon (forthcoming). Why companions in guilt arguments still work: Reply to Cowie. *Philosophical Quarterly*.

de Cruz, Helen ; Boudry, Maarten ; de Smedt, Johan & Blancke, Stefaan (2011). Evolutionary Approaches to Epistemic Justification. *Dialectica* 65 (4): 517-535.

DiPaolo, Joshua & Simpson, Robert Mark (2016). Indoctrination anxiety and the etiology of belief. *Synthese* 193 (10).

Dunaway, Billy (2017). Luck: Evolutionary and epistemic. *Episteme* 14 (4): 441-461.

Dworkin, Ronald (1996). Objectivity and Truth: You'd Better Believe It. *Philosophy and Public Affairs* 25 (2): 87-139.

Elga, Adam (2007). Reflection and disagreement. *Noûs* 41 (3): 478–502.

Elga, Adam (unpublished manuscript). Lucky to be rational. URL = www.princeton.edu/~adame/papers/bellingham-lucky.pdf (February 27, 2017).

Enoch, David (2010). The epistemological challenge to metanormative realism: How best to understand it, and how to cope with it. *Philosophical Studies* 148 (3): 413-438.

Enoch, David (2012). *Taking Morality Seriously: A Defense of Robust Realism*. Oxford University Press.

Enoch, David (2016). Indispensability Arguments in Metaethics: Even Better than in Mathematics? In: Neil Sinclair & Uri Lebowitz (eds.), *Explanation in Ethics and Mathematics*. Oxford University Press: 236-254.

Enoch, David (ms). How Principles Ground. Unpublished Manuscript. URL = http://www.academia.edu/34219914/How_Principles_Ground (August 24, 2018).

Feldman, Richard (2006). Epistemological puzzles about disagreement. In Stephen Hetherington (ed.), *Epistemology Futures*. Oxford University Press.

Field, Hartry (1989). *Realism, Mathematics, and Modality*. Oxford: Blackwell.

Fitelson, B. and Sober, E. (1998). Plantinga's Probability Arguments Against Evolutionary Naturalism. *Pacific Philosophical Quarterly* 79: 115–29.

Fitzpatrick, William (2008). Robust ethical realism, non-naturalism, and normativity. *Oxford Studies in Metaethics* 3: 159-205.

FitzPatrick, William J. (2015). Debunking evolutionary debunking of ethical realism. *Philosophical Studies* 172 (4): 883-904.

FitzPatrick, William J. (2016a). Misidentifying the Evolutionary Debunkers' Error: Reply to Mogensen. *Analysis* 76 (4): 433-437.

FitzPatrick, William (2016b). Morality and Evolutionary Biology. In: *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2016/entries/morality-biology/> (August 18, 2018).

Fodor, Jerry (2002). Is Science Biologically Possible? in J. Beilby (ed.) *Naturalism Defeated?* Ithaca, NY, Cornell University Press: 30–42.

Foot, Philippa (1967). The problem of abortion and the doctrine of double effect. *Oxford Review* 5: 5-15.

Fraser, Benjamin James (2014). Evolutionary debunking arguments and the reliability of moral cognition. *Philosophical Studies* 168 (2): 457-473.

Garner, Richard T. (1990). On the genuine queerness of moral properties and facts. *Australasian Journal of Philosophy* 68 (2): 137 – 146.

Gendler, Tamar Szabó & Hawthorne, John (2005). The real guide to fake Barns: A catalogue of gifts for your epistemic enemies. *Philosophical Studies* 124 (3): 331-352.

- Gettier, Edmund (1963). Is Justified True Belief Knowledge? *Analysis* 23 (6): 121-123.
- Ginet, C. (1995). Comments on Plantinga's Two-Volume Work on Warrant. *Philosophy and Phenomenological Research* 55: 403–8.
- Goldberg, Sanford (2010). *Relying on Others: An Essay in Epistemology*. Oxford University Press.
- Goldman, Alvin I. (1976). Discrimination and perceptual knowledge. *Journal of Philosophy* 73 (November): 771-791.
- Goldman, Alvin I. (2008). Immediate justification and process reliabilism. In: Quentin Smith (ed.), *Epistemology: New Essays*. Oxford University Press: 63-82.
- Golub, Camil (2017). Expressivism and realist explanations. *Philosophical Studies* 174 (6): 1385-1409.
- Greco, John (2000). *Putting Skeptics in Their Place: The Nature of Skeptical Arguments and Their Role in Philosophical Inquiry*. Cambridge University Press.
- Greco, John (2003). Virtue and Luck, Epistemic and Otherwise. *Metaphilosophy* 34 (3): 353-366.
- Griffin, James (2015). *What Can Philosophy Contribute to Ethics?*. Oxford University Press UK.
- Harman, Gilbert H. (1965). The inference to the best explanation. *Philosophical Review* 74 (1): 88-95.
- Harman, Gilbert (1977). *The Nature of Morality: An Introduction to Ethics*. Oxford University Press.
- Hayward, Max Khan (forthcoming). Non-Naturalist Moral Realism and the Limits of Rational Reflection. *Australasian Journal of Philosophy*: 1-14.
- Hills, Alison (2009). Moral testimony and moral epistemology. *Ethics* 120 (1): 94-127.
- Huemer, Michael (2003). Non-egalitarianism. *Philosophical Studies* 114 (1-2): 147-171.

Huemer, Michael (2005). *Ethical Intuitionism*. Palgrave MacMillan.

Huemer, Michael (2008). Revisionary intuitionism. *Social Philosophy and Policy* 25 (1): 368-392.

Ichikawa, Jonathan Jenkins & Steup, Matthias (2018). The Analysis of Knowledge. *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2018/entries/knowledge-analysis/> (August 28, 2018).

Jackson, Frank (1982). Epiphenomenal qualia. *Philosophical Quarterly* 32 (April): 127-136.

Jenkins, C. S. (2006). Knowledge and explanation. *Canadian Journal of Philosophy* 36 (2): 137-163.

Joyce, James M. (2005). How probabilities reflect evidence. *Philosophical Perspectives* 19 (1): 153–178.

Joyce, Richard (2001). *The Myth of Morality*. Cambridge University Press.

Joyce, Richard (2006). *The Evolution of Morality*. MIT Press.

Joyce, Richard (2008). Précis of The Evolution of Morality. *Philosophy and Phenomenological Research* 77 (1): 213-218.

Joyce, Richard (2014). Taking moral skepticism seriously. *Philosophical Studies* 168 (3): 843-851.

Kagan, Shelly (1989). *The Limits of Morality*. Oxford University Press.

Kahane, Guy (2011). Evolutionary Debunking Arguments. *Noûs* 45 (1): 103-125.

Kelly, Thomas (2005). The epistemic significance of disagreement. In: John Hawthorne & Tamar Gendler (eds.), *Oxford Studies in Epistemology, Volume 1*. Oxford University Press: 167-196.

Kelly, Thomas (2010). Peer disagreement and higher order evidence. In: Alvin I. Goldman & Dennis Whitcomb (eds.), *Social Epistemology: Essential Readings*. Oxford University Press: 183-217.

King, Nathan L. (2012). Disagreement: What's the Problem? or A Good Peer is Hard to Find. *Philosophy and Phenomenological Research* 85 (2): 249-272.

Kitcher, Philip (2005). Biology and ethics. In: David Copp (ed.), *The Oxford Handbook of Ethical Theory*. Oxford University Press.

Klenk, Michael (2017). Can Moral Realists Deflect Defeat Due to Evolutionary Explanations of Morality? *Pacific Philosophical Quarterly* 98 (S1): 227-248.

Korman, Daniel Z. (2014). Debunking Perceptual Beliefs about Ordinary Objects. *Philosophers' Imprint* 14.

Kvanvig, Jonathan L. (2004). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge University Press.

Kyriacou, Christos (forthcoming). Are Evolutionary Debunking Arguments Self-Debunking? *Philosophia*: 1-16.

Lackey, Jennifer (2008). A justificationist view of disagreement's epistemic significance. In: Alan Millar Adrian Haddock & Duncan Pritchard (eds.), *Social Epistemology*. Oxford University Press: 145-154.

Lasonen-Aarnio, Maria (2014). Higher-Order Evidence and the Limits of Defeat. *Philosophy and Phenomenological Research* 88 (2): 314-345.

Lenman, James (2014). Deliberation, schmeliberation: Enoch's indispensability argument. *Philosophical Studies* 168 (3): 835-842.

Lewis, David K. (1973). *Counterfactuals*. Blackwell Publishers.

Lipton, Peter (2004). *Inference to the Best Explanation*. Routledge/Taylor and Francis Group.

Lord, Errol (2013). From Independence to Conciliationism: An Obituary. *Australasian Journal of Philosophy* (2): 1-13.

- Lord, Errol & Maguire, Barry (eds.) (2016). *Weighing Reasons*. Oxford University Press USA.
- Luper, Steven (ed.) (1987). *The Possibility of Knowledge: Nozick and His Critics*. Rowman & Littlefield.
- Mackie, John L. (1977). *Ethics: Inventing Right and Wrong*. Penguin Books.
- Majors, Brad (2007). Moral explanation. *Philosophy Compass* 2 (1): 1–15.
- Matheson, Jonathan (2015). Disagreement and Epistemic Peers. *Oxford Handbooks Online*.
- McPherson (2012), Tristram. Ethical Non-Naturalism and the Metaphysics of Supervenience. In: *Oxford Studies in Metaethics Vol 7*.
- Miller, Alex (2003). *An Introduction to Contemporary Metaethics*. Polity.
- Miščević, Nenad (2007). Armchair luck: Apriority, intellection and epistemic luck. *Acta Analytica* 22 (1): 48-73.
- Mogensen, Andreas (ms). *Evolutionary Debunking Arguments*. DPhil Thesis. Oxford. URL = < <http://andreamogensen.com/wp-content/uploads/2014/10/COMPLETE-Evolutionary-debunking-arguments-in-ethics.pdf> > (14/8/2018).
- Mogensen, Andreas L. (2015). Evolutionary debunking arguments and the proximate/ultimate distinction. *Analysis* 75 (2): 196-203.
- Mogensen, Andreas L. (2016). Do evolutionary debunking arguments rest on a mistake about evolutionary explanations? *Philosophical Studies* 173 (7): 1799-1817.
- Mogensen, Andreas L. (2017). Contingency Anxiety and the Epistemology of Disagreement. *Pacific Philosophical Quarterly* 98 (1).
- Moon, Andrew (2017). Debunking Morality: Lessons from the EAAN Literature. *Pacific Philosophical Quarterly* 98 (S1): 208-226.

- Nagel, Thomas (1986). *The View From Nowhere*. Oxford University Press.
- Nozick, Robert (1981). *Philosophical Explanations*. Harvard University Press.
- Oddie, Graham (2005). *Value, Reality, and Desire*. Clarendon Press.
- Orozco, Joshue (2011). Epistemic Luck. *Philosophy Compass* 6 (1): 11-21.
- Pappas, George, (2017). Internalist vs. Externalist Conceptions of Epistemic Justification, *The Stanford Encyclopedia of Philosophy* (Fall 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2017/entries/justep-intext/> (14/8/2018).
- Parfit, Derek (1984). *Reasons and Persons*. Oxford University Press.
- Parfit, Derek (2013). *On What Matters: Volume Two*. Oxford University Press: Oxford.
- Peacocke, Christopher (2003). *The Realm of Reason*. Oxford University Press.
- Plantinga, Alvin (1993). *Warrant and Proper Function*. Oxford University Press.
- Plantinga, A. (2002). Reply to Beilby's Cohorts. in J. Beilby (ed.) *Naturalism Defeated?* Ithaca, NY, Cornell University Press: 204–75.
- Plunkett, David & McPherson, Tristram (2015). Deliberative Indispensability and Epistemic Justification. In: *Oxford Studies in Metaethics, vol. 10*. New York, USA: Oxford University Press: 104-133.
- Price, Richard (1758/1969). *A Review of the Principle Questions in Morals*, in: *The British Moralists 1650–1800, II*, D. D. Raphael (ed.), Oxford: Clarendon Press: 131–198.
- Pritchard, Duncan (2005a). The Structure of Sceptical Arguments. *Philosophical Quarterly* 55 (218): 37-52.
- Pritchard, Duncan (2005b). *Epistemic Luck*. Clarendon Press.
- Pritchard, Duncan (2012). Anti-Luck Virtue Epistemology. *Journal of Philosophy* 109 (3): 247-279.

Rachels, Stuart (1998). Counterexamples to the transitivity of better than. *Australasian Journal of Philosophy* 76 (1): 71-83.

Riaz, Amber (2015). Moral understanding and knowledge. *Philosophical Studies* 172 (1): 113-128.

Ridge, Michael (2018). Moral Non-Naturalism. *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2018/entries/moral-non-naturalism/> (August 26, 2018).

Riggs, Wayne (2007). Why epistemologists are so down on their luck. *Synthese* 158 (3): 329 - 344.

Roberts, Debbie (2016). 'Explanatory Indispensability Arguments in Metaethics and Philosophy of Mathematics'. In Uri D. Leibowitz & Neil Sinclair (eds.), *Explanation in Ethics and Mathematics: Debunking and Dispensability*. Oxford University Press.

Rosen, Gideon (2001). Nominalism, Naturalism, Epistemic Relativism. *Noûs* 35 (s15): 69 - 91.

Rosen, Gideon (ms). What is Normative Necessity? Unpublished Manuscript. URL = https://www.academia.edu/9159728/Normative_Necessity (August 28, 2018).

Roush, Sherrilyn (2005). *Tracking Truth: Knowledge, Evidence, and Science*. Oxford University Press.

Rowland, Richard (2013). Moral Error Theory and the Argument from Epistemic Reasons. *Journal of Ethics and Social Philosophy* 7 (1): 1-24.

Rowland, Richard (forthcoming). Rescuing Companions in Guilt Arguments. *Philosophical Quarterly*.

Ruse, Michael (1986). *Taking Darwin Seriously: A Naturalistic Approach to Philosophy*. Prometheus Books.

Sayre-McCord, Geoffrey (1988). Moral Theory and Explanatory Impotence. *Midwest Studies in Philosophy* 12 (1): 433-457.

- Scanlon, T. M. (2014). *Being Realistic About Reasons*. Oxford University Press.
- Schafer, Karl (2010). Evolution and Normative Scepticism. *Australasian Journal of Philosophy* 88 (3): 471-488.
- Schafer, Karl (2014). Knowledge and Two Forms of Non-Accidental Truth. *Philosophy and Phenomenological Research* 89 (2): 373-393.
- Schechter, Joshua (unpublished manuscript). Luck, Rationality, and Explanation. URL = <
<http://www.brown.edu/Departments/Philosophy/onlinepapers/schechter/LuckRationalityExplanation.pdf> > (26.2.2017).
- Schechter, Joshua (2010). The reliability challenge and the epistemology of logic. *Philosophical Perspectives* 24 (1): 437-464.
- Schechter, Joshua (2013). Could Evolution Explain Our Reliability about Logic? In: Tamar Szabo Gendler & John Hawthorne (eds.), *Oxford Studies in Epistemology* 4.
- Schoenfield, Miriam (2013). Permission to Believe: Why Permissivism Is True and What It Tells Us About Irrelevant Influences on Belief. *Noûs* 47 (1): 193-218.
- Searle, John R. (2001). *Rationality in Action*. MIT Press.
- Setiya, Kieran (2013). *Knowing Right from Wrong*. Oxford University Press.
- Severini, Eleonora (2016). Evolutionary Debunking Arguments and the Moral Niche. *Philosophia* 44 (3): 865-875.
- Shafer-Landau, Russ (2003). *Moral Realism: A Defence*. Oxford University Press.
- Shafer-Landau, Russ (2012). Evolutionary Debunking, Moral Realism and Moral Knowledge. *Journal of Ethics and Social Philosophy* 7 (2): 64-78.
- Sher, George (2001). But I Could Be Wrong. *Social Philosophy and Policy* 18 (02): 64-78.
- Sinclair, Neil (2013). Moral Explanations. In Hugh LaFollette (ed.), *International Encyclopedia of Ethics*. Blackwell.

Sinclair, Neil (2018). Belief pills and the possibility of moral epistemology. In: *Oxford Studies in Metaethics*. Oxford, U.K.: Oxford University Press.

Singer, Peter (1972). Famine, affluence, and morality. *Philosophy and Public Affairs* 1 (3): 229-243.

Skarsaune, Knut Olav (2011). Darwin and moral realism: Survival of the fittest. *Philosophical Studies* 152 (2): 229-243.

Sosa, Ernest (1999). How to defeat opposition to Moore. *Philosophical Perspectives* 13 (s13): 137-49.

Stalnaker, Robert (1968) A Theory of Conditionals. *Studies in Logical Theory, American Philosophical Quarterly* (2): 98-112.

Sterpetti, Fabio (2015). Are Evolutionary Debunking Arguments Really Self-Defeating? *Philosophia* 43 (3): 877-889.

Stich, Stephen (1990). *The Fragmentation of Reason*. Cambridge, MA: MIT Press.

Stratton-Lake, Philip (2014). Intuitionism in Ethics. *The Stanford Encyclopedia of Philosophy* (Winter 2014 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2014/entries/intuitionism-ethics/> (August 18, 2018).

Stratton-Lake, Philip (2016). Intuition, Self-Evidence, and Understanding. In Russ Shafer Landau (ed.), *Oxford Studies in Meta Ethics*. Oxford. Oxford University Press: 28-44.

Street, Sharon (2006). A Darwinian dilemma for realist theories of value. *Philosophical Studies* 127 (1): 109-166.

Street, Sharon (2008). Reply to Copp: Naturalism, normativity, and the varieties of realism worth worrying about. *Philosophical Issues* 18 (1): 207-228.

Street, Sharon (2009). Evolution and the normativity of epistemic reasons. *Canadian Journal of Philosophy* 39 (S1): 213-248.

Street, Sharon (2011). Mind-Independence Without the Mystery: Why Quasi-Realists Can't Have It Both Ways. *Oxford Studies in Metaethics* 6: 1-32.

Unger, Peter (1968). An analysis of factual knowledge. *Journal of Philosophy* 65 (6): 157-170.

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

Vavova, Katia (2014). Debunking Evolutionary Debunking. *Oxford Studies in Metaethics* 9:7 6-101.

Vavova, Katia (2015). Evolutionary Debunking of Moral Realism. *Philosophy Compass* 10 (2): 104-116.

Vavova, Katia (2018). Irrelevant Influences. *Philosophy and Phenomenological Research*: 134-152.

Väyrynen, Pekka (2017). The Supervenience Challenge to Non-Naturalism. In: Tristram McPherson & David Plunkett (eds.), *The Routledge Handbook of Metaethics*. Routledge: 170-84.

Wedgwood, Ralph (2007). *The Nature of Normativity*. Oxford University Press.

Wedgwood, Ralph (2010). The moral evil demons. In: Richard Feldman & Ted Warfield (eds.), *Disagreement*. Oxford University Press.

Werner, Preston J. (2018). Why conceptual competence won't help the non-naturalist epistemologist. *Canadian Journal of Philosophy* 48 (3-4): 616-637.

White, Roger (2010). You just believe that because.... *Philosophical Perspectives* 24 (1): 573-615.

Wielenberg, Erik J. (2010). On the evolutionary debunking of morality. *Ethics* 120 (3): 441-464.

Wielenberg, Erik J. (2016). Ethics and Evolutionary Theory. *Analysis* 76 (4): 502-515.

Williams, Bernard A. O. (1973). *Problems of the Self*. Cambridge University Press.

Williamson, Timothy (2000). *Knowledge and its Limits*. Oxford University Press.

Williamson, Timothy (2013). Gettier cases in epistemic logic. *Inquiry* 56 (1): 1-14.

Woods, Jack (2016). Mathematics, Morality, and Self-Effacement. *Noûs*.

Worsnip, Alex (2016). Explanatory Indispensability and Deliberative Indispensability: Against Enoch's Analogy. *Thought: A Journal of Philosophy* 5 (4): 226-235.

Yamada, Masahiro (2011). Getting It Right By Accident. *Philosophy and Phenomenological Research* 83 (1): 72-105.

Zagzebski, Linda (1994). The inescapability of Gettier problems. *Philosophical Quarterly* 44 (174): 65-73.

Zangwill, Nick (2006). Moral epistemology and the Because Constraint. In James Lawrence Dreier (ed.), *Contemporary Debates in Moral Theory*. Blackwell: 263-281.

Zimmerman, Aaron (2010). *Moral Epistemology*. Routledge.