

Quaestiones Infinitae

PUBLICATIONS OF THE DEPARTMENT OF
PHILOSOPHY AND RELIGIOUS STUDIES
UTRECHT UNIVERSITY
VOLUME CXIX

Copyright ©2019 by Jeroen Hopster

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without prior permission in writing of the author.

Cover art:

Affe mit Schädel. Replica of statue by Hugo Rheinhold (1892)

ISBN 978-94-6103-072-6

Printed by ProefschriftMaken || www.proefschriftmaken.nl

Moral Objectivity: Origins and Foundations

Morele Objectiviteit: Oorsprong en Grondslagen (met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. H.R.B.M. Kummeling, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op vrijdag 1 februari 2019 des middags te 16.15 uur

door

Jeroen Klaas Goswin Hopster

geboren op 6 maart 1987

te Amsterdam

Promotor: Prof.dr.mr. H. Philipse

This work is part of the research programme 'Evolutionary Ethics? The (Meta-)Ethical Implications of Evolutionary Explanations of Morality' with project number 360-20-340, which is financed by the Netherlands Organisation for Scientific Research (NWO).

'That there is a connection between human morality and our evolutionary past, I have no doubt. That the connection can be stated simply is surely belied by the history of brave, but disastrous, ventures in evolutionary ethics.'

– Philip Kitcher (1998)

Acknowledgements

This PhD research is part of the research project *Evolutionary Ethics: The (Meta-)Ethical Implications of Evolutionary Explanations of Morality* (2014–2019), funded by the Netherlands Organization for Scientific Research (NWO). Special thanks go to my supervisor, Professor Herman Philipse, who initiated the project and generously supported me during all stages of my research. Herman provided detailed comments on all the chapters included in this dissertation, as well as various drafts which did not make the final text. He gave me the freedom to find my own way into the research and helped me to improve both my philosophical writing and my thinking. Above all, discussing and collaborating with Herman was great fun; it was a delight to be a member of his research group.

Thanks, as well, to my colleagues in the research group: Joeri Witteveen, Julia Hermann, Michael Klenk and Wouter Kalf (in order of their appointment). We had many interesting discussions, on topics ranging from moral universals to the percentage of vegan sandwiches that should be served during the lunches of the conferences we organized. I laughed a lot and learned a lot; collaborating with every one of you has been a joy. Moreover, I benefited greatly from the excellent advice, friendly feedback and constructive criticism I received on several of the chapters to come. A special mention goes to Michael Klenk, with whom I had the pleasure of sharing a philosophical retreat in Observatory Hill, Cambridge, MA. There we initiated an experiment in moral psychology, which has been included in this dissertation.

I wrote part of this dissertation while I was a visiting fellow at Harvard University; thanks to my supervisor there, Selim Berker. Other occasional readers and commentators on the drafts of this dissertation are the members of the monthly Analytic Philosophy Seminar; thanks to all of you. Additionally, I have had the opportunity to present work in progress at various conferences and workshops; thanks to all the questions and remarks posed by attendees. And thanks to the reviewers of various academic journals, some of whom have been very generous with their feedback on my work; they are the anonymous backbone of the profession.

Devoting several years to academic research would not have been possible without the support of the NWO. I consider it a great privilege to have had the opportunity to spend my time doing what I love to do – philosophical research and writing – and am grateful to the NWO for providing the financial means to do so. Over the course of the project, the research team has organized various academic conferences and symposia, and we had the opportunity to invite to Utrecht several leading academics whose work plays a role in this dissertation, such as Sharon Street, Allan Gibbard, Kim Sterelny, Justin Clarke-Doane and Michael Huemer, as well as many others. The opportunity to discuss philosophical issues directly with these scholars has been enriching and one of the privileges of being a member of this research project.

In the summer of 2015, while briefly visiting Paris during a layover when travelling to the bi-annual ISHPSSB (International Society for the History, Philosophy and Social Studies of Biology) conference in Montreal, I met and fell in love with Giuseppe Guagliardo. I thank Peppe for his patience in putting up with me during holidays with laptop and during research hours that occasionally colonized evenings and weekend days; you have always supported and inspired me.

Whether genealogies can be ‘debunking’ is a question addressed in the upcoming pages. But genealogies can also be a source of encouragement and pride. I thank my parents, Paul Hopster and Renée Arendsen de Wolff, for planting and nurturing the seeds of my academic curiosity. And I thank my 100-year-old grandmother, Els Arendsen de Wolff-Exalto, who was instrumental in triggering my interest in science and biology. Shortly after the Second World War, after having defended her *doctoraal examen* to Niko Tinbergen, she was on the verge of pursuing PhD research in the field of biology, but personal circumstances ultimately prevented her from doing so. Nonetheless, her interest in biology has been retained in the family line and is partly reflected in the contents of the writings at hand; I dedicate this dissertation to her.

Table of Contents

Acknowledgements	6
Table of Contents	8
Chapter 1. Introduction	15
1.1 Aims and Claims.....	17
1.2 Morality: Evolutionary Origins.....	19
1.2.1 Evolutionary Explanations of Morality	20
1.2.2 Explaining Moral Capacities.....	21
1.2.3 Explaining Morality’s Contents.....	22
1.3 Moral Objectivity: Metaethical Foundations	23
1.3.1 Moral Realism versus Moral Antirealism.....	24
1.3.2 Realism and Stance-Independence	24
1.3.2.1 The Realist’s Scientific Hypothesis.....	25
1.3.3 Moral Antirealism and Stance-Invariance	26
1.3.3.1 The Antirealist’s Scientific Hypothesis	27
1.3.4 Why Vindicate Moral Objectivity <i>at all</i> ?	28
1.4 Method	29
1.4.1 External and Internal Accommodation	29
1.4.2 Beyond the External–Internal Dichotomy	29
1.4.3 Holistic Evaluation.....	30
1.4.4 Debunking and Vindicating	30
1.4.5 Empirical Metaethics	31
1.5 Preview of Chapters	31
References.....	34

Chapter 2. Striking Coincidences:

How Realists Should Reason about Them	41
2.1 Introduction.....	43
2.2 How Should We Reason about Coincidences?.....	44
2.2.1 Defining Coincidences: Causal versus Psychological Characteristics	44
2.2.2 Coincidence Observation versus Coincidence Explanation	44
2.2.3 Steps in Reasoning about Coincidences	45

2.2.4	Why do Coincidences Often Strike Us as Surprising?	45
2.3	The Practical/Theoretical Puzzle	46
2.3.1	Clarifying the Puzzle and its Target.....	46
2.3.2	Reframing the Practical/Theoretical Puzzle in terms of Grounding	47
2.3.3	Solving the Puzzle? The Realist’s Coincidence Explanation and Street’s Criticism.....	48
2.4	Pursuing a Coincidence Explanation	49
2.5	Questioning Street’s Reference Class	51
2.5.1	Street’s Reference Class without Darwinian Dilemma	52
2.5.2	FitzPatrick’s Counterproposal	53
2.5.3	Cuneo and Shafer-Landau’s Counterproposal	53
2.5.4	Lott’s Counterproposal	53
2.5.5	Conclusion: Are Coincidence Explanations Successful?	54
2.6	The Tenability of Alternative Explanations.....	54
2.6.1	Are Third-Factor Accounts Trivially Question-Begging?.....	56
2.6.2	Does Enoch’s Explanation Establish that Our Normative Judgments Are Generally Reliable?.....	57
2.7	Conclusion	58
	References.....	58

Chapter 3. Evolutionary Arguments against Moral Realism:

	Why the Empirical Details Matter (and Which Ones Do)	63
3.1	Introduction.....	65
3.2	Street’s ‘Darwinian Dilemma’	66
3.2.1	The Reliability Challenge	68
3.3	Ruse’s Contingency Challenge	70
3.4	Entangled between Street’s and Ruse’s EDAs	72
3.4.1	Countering Third-Factor Objections.....	73
3.4.2	Countering Objections regarding Evolutionary Constraints.....	74
3.5	Denying Influence: A Solution for Realists?	75
3.5.1	The Truth-Tracking Hypothesis.....	75
3.5.2	Divide-and-Conquer Strategy	77
3.6	Best Historical Explanation: A Comeback for Debunkers	78
3.6.1	Historical Debunking Argument.....	78
3.6.2	Genealogies without Stance-Independent Truths	79
3.6.3	Making Up the Balance	80
3.7	Conclusion	80
	References.....	80

Chapter 4. Explaining Historical Moral Convergence:

The Empirical Case against Realist Intuitionism 85

4.1	Introduction.....	87
4.2	Realism and Diffusing Explanations	88
4.3	The Historical <i>Explanandum</i>	89
4.4	Why Huemer takes the Realist <i>Explanans</i> to be Superior	90
4.4.1	Evolutionary Explanations and the ‘Scope Challenge’	90
4.4.2	Cultural Explanations and the ‘Coincidence Challenge’	91
4.4.3	Rational Explanations and the ‘Inference Challenge’	92
4.5	How Antirealists can Counter Huemer’s Challenges	92
4.5.1	Answering the ‘Scope Challenge’	92
4.5.2	Answering the ‘Coincidence Challenge’	93
4.5.2.1	First Step: Direct Cultural Influence.....	93
4.5.2.2	Second Step: Non-Moral Convergence	94
4.5.2.3	Third Step: Functional Explanation	95
4.5.3	Answering the ‘Inference Challenge’	96
4.6	Why the Antirealist’s <i>Explanans</i> is Superior.....	97
4.6.1	Antirealist Explanations Are More Robust.....	98
4.6.2	Antirealist Explanations Have a Better Fit with the Data.....	98
4.6.3	Antirealist Explanations Are More Predictive	99
4.6.4	Realist Explanations Are Insufficiently Detailed	99
4.6.5	Realist Explanations Are Incomplete.....	100
4.6.6	Realist Explanations Are Unclear.....	100
4.6.7	Realist Explanations Contain Redundancies	101
4.6.8	Realist Explanations Have a Poor Fit with Other Disciplines	102
4.7	Conclusion	102
	References.....	102

Chapter 5. Two Accounts of Moral Objectivity:

From Attitude-Independence to Standpoint-Invariance..... 105

5.1	Introduction.....	107
5.2	The Metaethical Relevance of Ordinary Discourse	110
5.3	Refining the <i>Explanandum</i>	111
5.4	Humean Constructivism.....	113
5.4.1	Factual Misinformation.....	114
5.4.2	Incoherence	114
5.5	Objectivity as Standpoint-Invariance.....	115

5.6	The Relativity Objection.....	117
5.7	The Alignment Objection	119
5.8	The Revision Objection	121
5.9	Conclusion	122
	References.....	123

Chapter 6. The Metaethical Significance of Experiments

about Folk Moral Objectivism..... 127

6.1	Introduction.....	129
6.2	Experimental Findings: An Overview	130
6.2.1	Interpersonal Variation	131
6.2.2	Intrapersonal Variation	131
6.3	Metaethical Implications? Pitfalls and Promises	132
6.4	The Concept of Moral Judgment	135
6.5	The Metaphysics of Moral Judgment	137
6.5.1	Criticism of <i>Premise 1</i>	139
6.5.2	Criticism of <i>Premise 2</i>	140
6.5.3	Criticism of <i>Premise 3</i>	140
6.6	The Psychology of Metaethics.....	141
6.7	Conclusion	143
	References.....	144

Chapter 7. Evolution and the Perceived Objectivity

of Moral Judgments..... 149

7.1	Introduction.....	151
7.1.1	Perceived Moral Objectivity: Its Relevance for EDAs.....	151
7.1.2	The Tenability of Ruse’s EDA: Philosophical and Empirical Issues	152
7.1.3	Outline of the Chapter.....	153
7.2	Determinants of Perceived Moral Objectivity: Our Experimental Hypothesis	153
7.2.1	Social versus Moral Norms.....	154
7.2.2	Self-Serving Bias	154
7.3	Experiment.....	155
7.3.1	Study	155
7.3.2	Results.....	156
7.3.3	Methodological Implications	156
7.3.4	Limitations of the Study.....	157

7.3.5	Do Our Findings Provide Support for Specific Metaethical Views?.....	157
7.4	General Discussion: The Psychosocial and Evolutionary Functions of Perceived Moral Objectivity.....	158
7.4.1	What Is the Function of Externalizing Moral Demands?	159
7.4.2	Moral Objectification: A Signalling Hypothesis	159
7.4.3	Choosing Sides	160
7.5	A Psychological Challenge for Moral Realism?.....	161
7.6	Conclusion	162
	References.....	162
	Appendix.....	165
 Chapter 8. Conclusion		167
8.1	Synthesis of the Research Results.....	168
8.1.1	The Genealogical Challenge for Moral Realism	168
8.1.2	The Psychological Challenge for Moral Realism	170
8.1.3	Explicating Moral Objectivity in terms of Stance-Invariance	170
8.2	Summary of Chapters	171
8.3	Realism versus Antirealism: The Overall Balance of Plausibility.....	174
	References.....	174
 Bibliography		176
Samenvatting in het Nederlands.....		187
Curriculum Vitae.....		191
Additional Scholarly Publications (2014–2018)		192

1

Chapter 1

Introduction

Keywords

Metaethics; moral realism; moral antirealism; evolutionary debunking arguments; explanation; justification

Abstract

What implications do recent empirical findings from the fields of biology, primatology, anthropology, psychology and history have for metaethical theories about moral objectivity? I defend the thesis that these findings detract from the plausibility of a realist account of moral objectivity but leave room for a more moderate, antirealist account of objectivity, framed in terms of stance-invariance. The dissertation consists of six articles written for publication in academic journals. In these articles I present two novel, empirically informed challenges to moral realism, and point out some shortcomings of existing challenges. One of the novel challenges builds on the second horn of Sharon Street's 'Darwinian Dilemma', according to which moral realists are committed to an implausible evolutionary hypothesis, and extends this criticism to the historical realm. The other novel challenge is fuelled by theoretical and experimental work in moral psychology, and takes issue with the presumed advantages of moral realism in explaining the qualities of our moral experience. Apart from criticizing moral realism, I also develop and defend an alternative account of moral objectivity in antirealist terms and argue that it is as least as successful as a realist account in capturing the objectivist commitments of ordinary moral discourse. In this introductory chapter I outline the aims and methods of the dissertation and demonstrate the coherence of the treatise.

‘[W]e must first introduce the concept of moral objectivity – a notion that gets used differently by different philosophers, and one so slippery that some have recommended its elimination.’

– Richard Joyce (2016)

1.1 Aims and Claims

Can we procure a sufficiently strong account of moral objectivity in the face of morality’s evolutionary origins? The aim of this dissertation is to argue, on the one hand, that we can, while criticizing, on the other hand, a realist understanding of moral objectivity, as defined below.

Let’s start with the negative proposal. Several leading contemporary metaethicists, including Thomas Nagel (1986), David Brink (1989), Russ Shafer-Landau (2003), Michael Huemer (2005), William FitzPatrick (2008) and David Enoch (2010), defend the realist thesis that moral judgments typically purport to state objective truths and that the truth-makers of these judgments are mind-independent moral properties or facts. To claim that a moral judgment is objectively true, on this view, is to claim that the contents of this judgment coincide with a mind-independent moral fact. Moral realism has been challenged on different grounds, some of them scientific – for instance by Sharon Street (2006) in her Darwinian Dilemma. Street’s so-called evolutionary debunking argument (EDA) has fuelled a recent industry of scholarship, much of which has been rather critical of her EDA. Upon reviewing the current literature, one might easily be left with the impression that Street’s argument has been discredited – either because the scientific considerations which it invokes are dispensable and do not add anything new to existing epistemological challenges to realism, or because realists are able to overcome Street’s dilemma (e.g. Copp 2008; Enoch 2010; Wielenberg 2010; Brosnan 2011; Skarsaune 2011; Shafer-Landau 2012; Berker 2014; FitzPatrick 2014, 2015; Tropman 2014; Vavova 2014; Artiga 2015; Das 2015; Mogensen 2015; Clarke-Doane 2016; Deem 2016; Hanson 2016; Huemer 2016; see Wielenberg 2016 for an overview). My general aim in this dissertation is to let the pendulum swing the other way and to provide new arguments that rehabilitate Street’s basic insight: in the light of our best science a realist understanding of moral objectivity is highly problematic.¹ Moreover, scientifically driven arguments against realism are not simply restatements of existing epistemological worries: they add a new and important component to the case against realism.

That said, the criticisms of Street’s original challenge deserve to be taken seriously; some of them are quite persuasive. Work needs to be done to demonstrate that moral realism can indeed be challenged on scientific grounds, and in a way that differs from existing metaphysical (e.g. Mackie 1977) and epistemological challenges (see Schechter 2018 for a

¹ This is not to deny that other, non-scientific arguments may successfully challenge realism as well; to evaluate the success of all philosophical challenges against realism is beyond the scope of this dissertation. I will, however, discuss one purely epistemological challenge – Sharon Street’s (2016a) practical/theoretical puzzle – and argue that her argument is unsuccessful in challenging realism *tout court*, unless it is supported by scientific considerations (**chapter 2**).

discussion of recent challenges). I take up such work in the present dissertation by advancing two novel, empirically informed arguments against moral realism.²

My first argument takes off where Street's EDA falls short: in providing a sufficiently exhaustive genealogy of our moral endorsements. Several critics of Street have argued that her Darwinian Dilemma does not get off the ground, since the contents of our moral endorsements cannot be adequately explained in evolutionary terms (Shafer-Landau 2012; FitzPatrick 2014, 2015; Huemer 2016). I argue that this objection can be overcome by widening the scope of the genealogies of our moral endorsements, and specifically by extending them to the *historical* realm – i.e. the domain studied and explained by historians.³ Although historical explanations of our moral endorsements do not serve to reinstate Street's original dilemma, they do establish the following: the best, empirically informed explanations of the origins of our moral endorsements discredit the thesis that over the course of evolution and human history, we have tracked mind-independent moral truths. To posit this 'tracking thesis' – to which realists are committed if they want to procure a 'success epistemology' – is scientifically problematic in various respects. Therefore, if metaethicists take evolutionary and historical science seriously, moral realism is in trouble.

My second challenge builds on recent findings from experimental moral psychology (e.g. Sarkissian et al. 2011; Goodwin and Darley 2012; Wright et al. 2013; Beebe 2014), as well as leading theories from evolutionary psychology (DeScioli and Kurzban 2013; Stanford 2018). One of the main arguments that realists have traditionally provided in support of their view comes from moral phenomenology, and it appeals to the *seeming* objectivity of moral demands (e.g. Brink 1989; Enoch 2014). However, recent experimental findings do not straightforwardly vindicate the assumption that moral demands are conceived as objective in the realist's sense. Moreover, to the extent that moral demands *are* conceived as strongly objective, there are reasons for thinking that, at least in part, this is an artefact of our evolved psychology (Ruse 1995; Joyce 2006; DeScioli and Kurzban 2013; Stanford 2018). This does not establish that the objectivity of moral demands is merely an illusion, as some debunkers have argued (Ruse 1995): after all, that we have evolved to perceive moral demands as objective does not imply that their objectivity cannot be justified on metaethical grounds. But recent psychological findings and theories do serve to undercut the 'presumptive support' that our moral phenomenology bestows upon moral realism – which, as mentioned, is one of the main arguments that realists have traditionally provided in support of their view.

² I develop these challenges throughout the chapters of the dissertation. In **chapter 3** and especially **chapter 4** I advance my historical argument against moral realism. In **chapter 6** and especially **chapter 7** I advance my psychological argument against moral realism. I state succinct versions of these arguments in the conclusion.

³ Of course, evolutionary explanations typically have a historical component as well. Indeed, in the context of our present inquiry into morality's origins, evolutionary explanations are entangled with the historical details of human evolution. But it is also common to delineate the historical domain in a narrower way – as the domain studied by historians, which spans roughly the period from 10kya (thousand years ago) until the present. Here I use the term 'historical explanations' in the latter, narrower sense. When speaking of historical explanations, I mean the diverse set of explanations that historians put forward to elucidate *inter alia* processes of moral change. When referring to 'historical science', what I have in mind is the English equivalent of the Dutch *geschiedwetenschap*, i.e. our body of knowledge about the past that is based on reliable sources and well-validated methods of research.

The positive proposal of this dissertation is to explicate claims about moral objectivity in terms of stance-invariance rather than mind-independence. According to this proposal, to claim that a moral judgment is objective is to claim that this judgment withstands argumentative scrutiny from a wide range of moral standpoints. The greater the diversity of standpoints from which it withstands such scrutiny, the more objective the judgment may be taken to be.⁴ This conception of moral objectivity is not entirely new; it builds, for instance, on accounts put forward by Rawls (1971), Sen (1993), Habermas (1995), Nozick (2001) and Street (2008). The main contribution that I seek to make is to establish that defenders of this alternative account of objectivity are at least as well placed as realists to capture the commitments of ordinary moral thought and practice. Contrary to the often-made claim that moral realism is superior to antirealism in capturing morality's 'internal' commitments (e.g. Finlay 2007; Street 2016b; Enoch 2018), then, there may not be any decisive reason to prefer either view for its distinctive ability to satisfy the objectivist *desiderata* internal to moral discourse. But there *is* a decisive reason to favour the antirealist's view on 'external', scientific grounds: antirealists are not committed to the realist's controversial tracking thesis and can advance a hypothesis about the origins of objective moral judgments that is clearly superior in scientific terms.

This introduction proceeds as follows. In **section 1.2** I overview how evolutionary considerations might help us to explain our moral capacities and the contents of our moral endorsements. In **section 1.3** I introduce some central questions of metaethical inquiry and delineate the topic of the present dissertation by outlining the two rival metaethical positions that will be central to my discussion: moral realism and moral antirealism. In **section 1.4** I outline the methodology of the dissertation, highlighting some of the main *desiderata* relevant for metaethical analysis. I emphasize that although the scientific endeavour of explaining morality's origins and the metaethical endeavour of justifying morality's objective foundations are logically independent, the former is nonetheless relevant for the latter, since it is a *desideratum* of metaethical theories to be compatible with our best scientific understanding of morality's origins. Indeed, I will argue that this *desideratum* plays a decisive role in tilting the overall balance of plausibility in favour of an antirealist account of objectivity. In **section 1.5** I provide a preview of the chapters to come.

1.2 Morality: Evolutionary Origins

The American anthropologist and primatologist Sarah Hrdy begins her book *Mothers and Others* (2009) by recounting her experience upon entering an overloaded airplane:

With nods and resigned smiles, passengers make eye contact and then yield to latecomers pushing past. When a young man wearing a backpack hits me with it as he reaches up to cram his excess paraphernalia into an overhead compartment, instead of grimacing or baring my teeth, I smile (weakly), disguising my irritation. Most people on board ignore the crying baby, or pretend to. (Hrdy 2009, p. 1)

⁴ As this formulation reveals, my proposal allows for a graded understanding of moral objectivity. In **chapter 5** I will argue that this explication accords well with many of our pre-theoretical intuitions.

A few pages on, Hrdy ponders what would happen if her fellow passengers suddenly morphed into another species of primate:

What if I were traveling with a planeload of chimpanzees? Any one of us would be lucky to disembark with all ten fingers and toes still attached, with the baby still breathing and unmaimed. Bloody earlobes and other appendages would litter the aisles. Compressing so many highly impulsive strangers into a tight space would be a recipe for mayhem. (idem, p. 3)

Hrdy's thought experiment serves to underscore the uniquely social habits of members of our species, compared to other primates. This emphasis contrasts with recent work by other primatologists, such as Frans de Waal, who stress the continuity between the social behaviour of humans and other species. Notably, de Waal has observed that behaviours associated with psychological altruism, as well as a basic sense of fairness (Brosnan and de Waal 2003), are present in many of our primate relatives.⁵ But while the continuities between our motivational impulses and those of other primates are quite striking, so are the discontinuities, as Hrdy's thought experiment reveals. Consider our linguistic abilities, our rule-following capacities, our mind-reading skills and our impulse control, all of which are salient components of human social life. There is only one primate species whose members fasten their seatbelts upon command, leave the food on their neighbour's tray table untouched and might even donate some of their own resources to distant strangers. How should we explain this peculiar behaviour?

1.2.1 Evolutionary Explanations of Morality

It is fairly uncontroversial, nowadays, to explain the emergence of prosocial motivations by appealing to a variety of evolutionary mechanisms, including kin selection (Hamilton 1963), reciprocal altruism (Trivers 1971), indirect reciprocity (Alexander 1987), group selection (Wilson and Sober 1994), gene-culture co-evolution (Richerson and Boyd 2005) and mutualism (Baumard et al. 2013). More controversially, some scientists, such as de Waal (2006), have argued that evolutionary mechanisms can also explain the 'building blocks' of *morality*, as evinced by the prosocial behaviour of non-human primates. Contemporary philosophers, however, are typically reluctant to submit that such behaviour amounts to *moral* behaviour. For instance, the philosopher Richard Joyce observes that

an explanation of how humans came to have prosocial inclinations and aversions – whether grounded in love and sympathy or anger and disgust – is not an explanation of how humans came to judge things morally right and wrong, and to this extent is no explanation of an innate moral faculty. At best it is the start of an explanation. (Joyce 2016, p. 51)

⁵ In explaining animal behaviour, de Waal's interest is in *psychological altruism* – roughly, a disposition for other-regarding behaviour. This disposition – which de Waal argues is a major precursor of *moral altruism* – should be distinguished from *biological altruism*: behaviour that reduces the fitness of the organism that engages in it, while raising the fitness of others (see, for example, Kitcher 1998).

Following Joyce, there is an ‘explanatory gap’ between prosocial motivations on the one hand, and moral judgments on the other. In addition to elucidating the origins of the former, can evolutionary mechanisms also account for the latter?

To answer this question we first need to clarify what we mean by moral judgments.⁶ I will not attempt to offer any strict definition, in terms of necessary and sufficient conditions. Perhaps no uncontested definition along such lines can be given: morality’s conceptual boundaries may be vague. That said, moral judgments do have certain paradigmatic features, which clearly set them apart from prosocial inclinations and aversions. For instance, they typically make reference to norms whose force is regarded as binding for a large group of people, and perhaps even *categorically* so – that is, irrespective of the specific desires and preferences that members of this group might have. Also, these norms are typically socially enforced and are often codified in legal and religious prescriptions; they typically serve to regulate social interaction.

Can these paradigmatic features be explained in evolutionary terms? In recent years, scholars have proposed that morality serves as a ‘social technology’ (Kitcher 2011, pp. 221–241), which evolved in hominin societies to foster social cooperation (Tomasello 2016) – a function it still serves in many contemporary societies (Curry 2016). This seems to be a fruitful general hypothesis, keeping in mind that cooperation comes in different forms – e.g. cooperation in *dyadic* versus *group* settings (cf. Tomasello 2016), or cooperation aimed at *generating* versus *distributing* benefits (Warneken 2018) – and keeping in mind that morality’s functions may be manifold and that they may have changed over time. We should not expect that evolutionary explanations of morality will be simple and uniform: moral practices are a complex mosaic with different evolutionary origins and they respond to different selective pressures (Fraser and Sterelny 2017). Moreover, the *scope* of evolutionary explanations should not be overestimated (cf. Buchanan and Powell 2015): while it has become clear, in recent decades, that many facets of morality can be explained at least partly in evolutionary terms, it might well be the case that other facets, at least to some extent, cannot (see **chapter 3**).⁷

1.2.2 Explaining Moral Capacities

A key conceptual distinction is whether explanations of moral judgments concern the *capacity* for moral judgment or its *contents*. There is accumulating evidence that several major components of our *capacity* for moral judgment originated during the Pleistocene epoch (roughly 2.6mya until 12kya), the geological period which spans most of the era of hominin evolution. Such components include, for instance, the capacity to distance ourselves from our immediate impulses and to reflect on them, and the capacity to express ourselves linguistically. Can these – and other – constitutive components of a capacity for moral judgment be explained

⁶ Morality is a mixed bag; possible *explananda* include moral behaviours, norms, principles, values, judgments, *etc.* In this dissertation my discussion will typically centre on explaining the contents of moral judgments – or, slightly more broadly, the contents of our moral endorsements.

⁷ We should also keep in mind, however, that the scope of evolutionary explanations is itself subject to change. For instance, in recent decades gene-culture models have become increasingly important in explaining hominin evolution; the explanatory scope of these models can be very encompassing.

as evolutionary adaptations? Several leading scientists think that many of them can. Consider the following:

- *A cognitive capacity to follow norms*

Our hominin ancestors seem to have evolved a dedicated ‘norm-psychology’: a suite of cognitive heuristics, biases, motivations and dispositions for perceiving, learning, adhering to and punishing violations of the shared behavioural standards of one’s community (Chudek and Henrich 2011). Henrich (2015) hypothesizes that this norm-psychology is the result of selective pressures which emerged in a process of gene-culture co-evolution and that it was fuelled by the growing importance of social learning in the lifestyle of early hominins.

- *An emotional sensitivity to the well-being of others*

From early childhood onwards, humans show a clear willingness to help others and to share with them (Hamlin 2014; Warneken and Tomasello 2011). Our capacity to empathize with other people, to read their minds and to form shared intentions is very unlike other primates. Hrdy (2009) hypothesizes that these capacities are due to selective pressures that accompanied a transition in the life history of our hominin ancestors: hominins evolved an elongated childhood, with intensive (allo-)parental care.

- *An emotional sensitivity to one’s own reputation*

We do not only care about the well-being of others, but also about what others think of us. This sensitivity to reputation is an important precondition for moral behaviour; it may lead people to seek praise by acting virtuously, or to avoid the shame of being regarded as a free-rider. Alexander (1988) hypothesizes that our sensitivity to reputation can be evolutionarily explained in terms of indirect reciprocity.

1.2.3 Explaining Morality’s Contents

Can evolutionary considerations also help us to explain morality’s *contents*? At least with regard to *some* paradigmatic judgments, norms and principles, this does seem to be the case. Consider, for instance, the judgment that ‘stealing is wrong’: endorsements of this judgment are likely to be adaptive in social contexts, and may have evolved as a mechanism to prevent free-riding. Or consider the so-called golden rule – treat others as you would like others to treat you – which may be a universal rule of moral systems (e.g. Curry 2016).⁸ The contents of this rule resemble the mechanism of reciprocal altruism: reducing your own fitness to the evolutionary benefit of the person you are interacting with, with the future prospect that the service will be repaid – ‘I’ll scratch your back, you’ll scratch mine.’⁹ Of course, we should keep in mind the many complications in determining whether a trait is indeed an adaptation (Pözlner 2017) and should not just tell ‘just-so stories’ about their origins (Gould and Lewontin 1979). But at least regarding the basic contents of many rudimentary moral judgments, there are good grounds for thinking that they can be explained in evolutionary terms.¹⁰

⁸ As Nowak (2012, p. 56) observes, versions of this rule are omnipresent; he lists examples found in Greek philosophy, Buddhism, Taoism, Hinduism, Christianity, Judaism and Islam.

⁹ Here we are discussing the *biological* notion of altruism – altruism understood in terms of fitness costs and benefits (cf. footnote 5).

¹⁰ As Schloss (2014) observes, strictly speaking the golden rule is not an instantiation of ‘tit-for-tat reciprocity’, like the rule ‘I’ll scratch your back, you’ll scratch mine.’ Instead, in game theory the golden

On the other hand, the contents of some paradigmatic *full-fledged* moral judgments, norms and principles – i.e. judgments, norms and principles that are typically linguistically expressed, consciously reflected upon and endorsed after a process of moral deliberation – cannot obviously be evolutionarily explained. Consider judgments that counsel impartial benevolence, compassion for vulnerable strangers and concern for future generations (Shafer-Landau 2012, p. 7): *prima facie* such judgments run counter to the evolutionary prediction, on the basis of kin selection, that moral judgments should typically be *favouritist* and give preference, first and foremost, to one’s own kin. Or consider the prevalence of deontological intuitions in ordinary moral judgments (Greene 2014), which suggest that some actions are morally wrong, regardless of the consequences. *Prima facie*, the fact that many archetypical moral demands are non-consequentialist and purport to state categorical demands seems difficult to explain from an evolutionary point of view: we would expect natural selection to favour judgments that achieve better fitness consequences and to respond flexibly to contextual factors (DeScioli and Kurzban 2009, 2013).¹¹ Of course, it is possible that *ultima facie* these phenomena *can* be explained in evolutionary terms. These preliminary observations merely serve to point out that the adequacy of an evolutionary framework in explaining morality’s contents, and the insight this explanatory framework provides, cannot be taken for granted but needs to be evaluated on a detailed, case-by-case basis.

In sum, there is evidence that the contents of many prosocial motivations, as well as the basic contents of some moral judgments, can be explained as adaptive solutions to ‘cooperation problems’ in the lives of ancestral hominins. To what extent this evolutionary influence also permeates our full-fledged moral judgments, however, is a more controversial issue, to which we return in the ensuing chapters.

1.3 Moral Objectivity: Metaethical Foundations

Does the finding that there is a connection between the evolutionary origins of our basic moral motivations and the contents of the moral rules and theories that we typically regard as justified have any implication for *moral philosophy*? As philosophers have long pointed out, we should not fall into the trap of thinking that empirical findings concerning the origins of our moral sentiments have direct implications for the tenability of our first-order moral judgments. For instance, if an anthropologist observes that incest avoidance is a human universal (e.g. Westermarck 1906), and explains its universal endorsement in terms of its beneficial fitness consequences, this does not mean that morally we should avoid incest. Considered in isolation, empirical findings cannot tell us what we ought to do morally; to think otherwise is to commit a version of the naturalistic fallacy.¹²

rule conforms to a strategy called ‘forgiving tit-for-tat’, which typically emerges under conditions of imperfect communication.

¹¹ Though one’s expectations may of course vary depending on the details of the evolutionary explanation one favours. Note that DeScioli and Kurzban (2013) themselves maintain that the contents of this judgment *can* in fact be explained in evolutionary terms, but that an adequate evolutionary explanation for this (and many other) moral judgments makes no appeal to kin selection or direct reciprocity.

¹² Here I use the label ‘naturalistic fallacy’ in a broad sense, which includes two rather different statements by David Hume and G.E. Moore. Hume’s (1733) insight is that in moral reasoning ought-statements do not logically follow from is-statements: purely factual information does not suffice to

On a *metaethical* level, on the other hand, empirical findings may, at least indirectly, provide evidence that is relevant for evaluating the status of theories about morality. For instance, evolutionary findings may shed light on what best explains the basic contents of our moral judgments, norms and principles, and such explanations may lower the credentials of metaethical theories which are committed to a very different kind of explanation. In this way, explanations about morality's origins can be relevant for theories about morality's foundations, as I shall illustrate in this dissertation.

1.3.1 Moral Realism versus Moral Antirealism

Metaethical theories purport to elucidate the nature of morality – specifically the semantic, epistemological, metaphysical and psychological properties of moral thought, talk and practice. For instance, metaethicists may inquire whether or not moral judgments purport to be truth-apt, whether or not they purport to be objective and how this claim to objectivity is best understood. The latter issue – morality's claim to objectivity – is central to the present inquiry. In what follows, when referring to 'metaethical theories' I typically mean to refer to theories that make claims about the nature of moral objectivity.

Among objectivist views, the metaethical landscape can be divided in different ways. One division is commonly made between naturalist and non-naturalist views. Both naturalists and non-naturalists hold that moral judgments are truth-apt and (at least sometimes) purport to be objective. Their difference concerns the question whether the truth-makers of moral judgments are natural or non-natural properties.

In this dissertation, I shall restrict my discussion by focussing on a different distinction: the distinction between realist and antirealist views of moral objectivity. What sets apart these two positions is a different question: whether or not moral truths are *stance-independent*. In the remainder of this section, I clarify how this claim to stance-independence should be understood and highlight the characteristic features of both positions.

1.3.2 Realism and Stance-Independence

The designation 'moral realism' is used in different ways by different metaethicists. On a fairly inclusive definition, moral realists subscribe to the following claims about the nature of moral discourse (cf. Leibowitz and Sinclair 2016, pp. 4-5):

- 1) *Semantic descriptivism and success*: Assertoric moral claims, when interpreted literally, offer descriptions of a corresponding moral domain. Some of these descriptions are accurate (they state facts) and are therefore true.
- 2) *Psychological cognitivism*: Sincere assertoric moral claims express moral beliefs.
- 3) *Epistemological optimism*: Some of our current moral beliefs are epistemically justified (or count as knowledge), and our current methods of forming such beliefs are reasonably good.

establish a normative conclusion. According to Moore (1903), who coined the term naturalistic fallacy, 'good' is an unanalysable concept. Providing an analysis of 'goodness' in terms of other properties, either natural or metaphysical, amounts to committing a fallacy.

- 4) *Metaphysical commitment*: There exist properties corresponding to some of the terms and claims of moral discourse.

On a narrower definition, the distinctive clause of realist theories is a further metaphysical claim, namely that moral properties are mind-independent: they do not constitutively depend on our attitudes towards and thoughts about them. The version of realism committed to this claim is sometimes called ‘robust realism’ (e.g. Enoch 2007; FitzPatrick 2008); sometimes the claim about mind-independence is taken to be the defining clause of realism *as such* (e.g. Street 2006). In this dissertation I shall ordinarily refer to realism in the latter sense: as a metaethical theory committed to the claim that mind-independent moral truths, facts or properties exist.

According to realists, then, the purported objectivity of true moral statements is best understood as a claim about the mind-independence of these truths. This is not to say that realists do not grant that *at least some* moral truths depend on people’s mental states. Instead, realists mean to assert the positive thesis that there is a moral reality beyond what people think, feel and agree upon. Russ Shafer-Landau has given this thesis its canonical expression, reframed as a claim about the *stance-independence* of moral truths:

Realists believe that there are moral truths that obtain independently of any preferred perspective, in the sense that *the moral standards that fix the moral facts are not made true by virtue of their ratification from within any given actual or hypothetical perspective*. That a person takes a particular attitude toward a putative moral standard is not what makes that standard correct. (...) [T]he truth of any first-order normative standard is not a function of what anyone happens to think of it. Such standards, if true, are not *made true*, and, in particular, are not correct in virtue of being vindicated by some process of (inter)personal election or approbation. (Shafer-Landau 2003, pp. 15-16, his italics)

One implication of this claim to stance-independence is that some moral standards might be correct, even if no actual person takes them to be; indeed, even if everyone renounces them (*idem*, p. 17). Hence, claims about the stance-independence of moral truths can be reframed as claims about the counterfactual condition under which such truths are preserved. For instance, on a realist view, some moral judgments, such as ‘torturing innocent people is morally wrong’ might be true, even if no one thinks they are. After all, there are moral truths which hold independently of the attitudes of *all* moral agents.

1.3.2.1 The Realist’s Scientific Hypothesis

Moral realism is a metaethical theory. First and foremost, it seeks to *justify* the objective foundations of our moral endorsements rather than to *explain* their origins. Nonetheless, in virtue of their metaethical commitments, several realists are also committed to a scientific hypothesis: over the course of human evolution and history, we have been able to track these stance-independent moral truths. After all, if realists want to retain their claim to epistemological optimism, and want to avoid positing that our judgments simply *happened to coincide* with stance-independent truths, then they are forced to commit themselves to the thesis that we have *tracked* these truths. This tracking thesis is a scientific hypothesis: its

adequacy at least partly depends on its fit with our knowledge from domains such as evolutionary biology and human history.¹³ Hence, realists should seek to establish that the tracking thesis is underwritten by the best scientific explanations of our moral endorsements – or, at the very least, that our best scientific explanations do not conflict with it. I will argue, however, that this scientific hypothesis is problematic and that this constitutes a potent reason to favour an antirealist view of moral objectivity (see especially **chapter 4**).

1.3.3 Moral Antirealism and Stance-Invariance

The claim of moral antirealists, on a maximally inclusive definition, is that there are no stance-independent moral truths or facts.¹⁴ Thus understood, a wide range of metaethical doctrines count as antirealist, such as expressivism (which rests on the denial that moral discourse is factive) and error theory (which rests on the denial that stance-independent moral truths exist, but asserts that our prototypical moral judgments typically do refer to such truths). But like its realist counterpart, moral antirealism can also be characterized in a more constrained manner, as the view that moral discourse is factive and that some moral judgments are true, but that the standards that legislate their correctness are stance-*dependent*. On this more constrained definition, the hallmark of antirealism is that, contra realism, *all* moral truths are *at least partially* grounded in our attitudes (cf. Berker 2014). Henceforth, when speaking of antirealism, this is the view I shall be referring to.

On this version of antirealism, we can legitimately speak about moral truths, facts or properties. It should be kept in mind that these assumptions are not uncontroversial in metaethics. As previously mentioned, several metaethicists deny that moral judgments are truth-apt (e.g. Horgan and Timmons 2006); others grant that moral judgments are apt for consideration of truth and falsehood, but deny that true moral judgments state moral facts (e.g. Putnam 2004); still others grant that moral judgments are truth-apt and state moral facts, but deny that these moral facts are objective (e.g. Street 2010). The version of antirealism that I will develop (see **chapter 5** of this dissertation) should be delineated from all of these positions.

Indeed, in many general respects, the version of antirealism that I defend agrees with the *broad* version of realism outlined in **section 1.3.2**. It agrees, for instance, that *at least some* moral judgments purport to state truths, that in *some* sense these truths purport to be objective and that *at least sometimes* they succeed in doing so. The disagreement concerns the question of how this pre-theoretical claim to moral objectivity is best explicated. While realists understand it as a claim about the *stance-independence* of moral truths, on the antirealist view that I will be defending, it is a claim about their *stance-invariance*.

How should this claim be understood? First, antirealists submit – like moral realists – that agents may be mistaken about their moral judgments. But they disagree with realists that moral mistakes are due to a failure to grasp mind-independent moral facts. Instead, according

¹³ Some realists may want to deny that they are committed to this tracking thesis, understood as a hypothesis that is apt for scientific evaluation. However, denying that we have tracked moral truths leads to a different problem for realists: it seems difficult for them to explain the reliability of our moral judgments. This problem is pressed by Sharon Street in her ‘practical/theoretical puzzle’, which I discuss in **chapter 2**.

¹⁴ Or more fully: no stance-independent truths, facts or properties. In what follows, I will typically leave out the latter additions and will centre the discussion on moral truths.

to the version of antirealism I shall be defending, they are ultimately mistakes by agents' own lights: if an agent had been more coherent or better informed, she would have judged differently. In other words, moral mistakes transpire in virtue of the agent's own idealized evaluative standpoint (cf. Street 2008).

This does not yet provide us with an account of moral objectivity, however. In order to hold objectively, a moral judgment should not only be able to withstand argumentative scrutiny from the agent's own idealized standpoint, but also from the evaluative standpoints of other agents – and the more diverse these other standpoints are, the more objective the agent's judgment may be taken to be. This is what it means for claims to hold objectively on the antirealist view advanced in this dissertation: such claims are able to withstand a process of *intersubjective scrutiny* in a diverse community of moral agents. Hence, if an ideally coherent agent makes the moral judgment that *p*, and her judgment withstands scrutiny from diverse angles, then *p* may be regarded as an objective moral truth. Thus understood, objectivity does not boil down to intersubjective consensus: what makes a moral claim objective is not whether different agents actually agree about its truth, but whether this claim can withstand scrutiny from the standpoints of other coherent and well-informed agents.

Contrary to the realist view outlined above, on this account, objective moral truths transpire *in virtue* of the attitudes of moral agents and can never be fully independent of them. They have their binding authority irrespectively of the *particularities* of agents' evaluative commitments, but are not completely independent of their evaluative commitments *as such*. How the commitments entailed by these realist and antirealist views of objectivity differ, then, comes out most clearly when we ask whether there might be any moral facts that exist independently of the attitudes of all moral agents. According to realists, this constitutes a genuine possibility; according to antirealists, it does not.

It is sometimes claimed that moral realists maintain that we 'discover' moral truths, whereas antirealists maintain that we 'invent' or 'construct' them. On the antirealist view outlined here, the metaphor of 'discovery' can be misleading, if it is meant to suggest that the object of discovery is fully external to the attitudes of moral agents. However, the metaphors of 'invention' and 'construction' can be somewhat misleading as well, at least if construction is taken to denote an arbitrary process that might have easily led to very different results. Instead, antirealists might argue that to the extent that moral standards are 'invented', such inventions are akin to technological inventions: they solve the design problems of social life and often serve to enhance cooperative interactions (Kitcher 2011). Perhaps a more suitable metaphor, on the antirealist view outlined above, is that moral deliberation is a process of *exploration* – not of a pre-existing moral reality, but of what norms contribute to the flourishing of human individuals and societies in a given sociohistorical context. Such a process of exploration involves an open-ended social discussion about what we morally ought to do, but it does not involve the discovery of a moral reality fully external to ourselves.

1.3.3.1 The Antirealist's Scientific Hypothesis

Like many realists, antirealists are committed to a general scientific hypothesis. The distinguishing feature of this hypothesis is not a claim to the existence of stance-independence moral truths, but rather the absence of such a claim. The antirealist's general hypothesis – which allows for different specifications – is that starting from our evolved moral impulses,

which typically orient us towards cooperative behaviour, we have modified our moral judgments in increasingly complex cultural and technological settings, with the help of moral philosophers and lessons learned from history, in an open-ended process of social deliberation. Realists may, of course, agree with this. The difference with the realist's hypothesis is that antirealists are additionally committed to the following claim: our best, empirically informed genealogies of moral judgments involve no appeal to tracking stance-independent moral truths.

1.3.4 Why Vindicate Moral Objectivity *at all*?

The endeavour to vindicate a strong kind of objectivism is regarded, by some, as the 'Holy Grail in secular meta-ethics' (Street 2016b, p. 165). Why so? The arguments traditionally advanced in support of this *desideratum* typically have to do with the nature of moral experience and of moral deliberation. Phenomenologically, it seems that the demands of morality are objective in a way that matters of taste or etiquette are not (cf. Enoch 2014). For instance, contrary to many conventional norms, moral norms are often strongly enforced; transgression typically leads to punishment. What is more, even in the absence of external enforcement, moral demands are sometimes experienced as being externally imposed. For some, their binding force appears to be 'authority-independent' (Brink 1989, p. 24). Additionally, archetypical moral norms purport to be impartial: they apply categorically to all members of a given group, irrespective of existing loyalties and relations (DeScioli and Kurzban 2013). These considerations suggest that there is more to moral objectivity than mere intersubjective consensus, and have led many metaethicists to favour a realist view.

Of course, we should be aware of the fact that pre-theoretical intuitions informed by moral phenomenology may be rather vague, and perhaps more diverse than philosophers have traditionally assumed (see **chapter 6** for an overview of state-of-the-art research on folk moral objectivism). Moreover, our moral experience might be mistaken; we might be deluded and think that moral demands are objective, whereas, in fact, they are not.

That said, if a claim to objectivity is deeply entwined in moral discourse, then it is reasonable to start our inquiry by analysing whether this claim can be justified. Moreover, given that a majority of professional philosophers are self-acclaimed moral realists (Bourget and Chalmers 2014), this view merits scholarly attention.¹⁵ But we should keep in mind that there is an alternative to the realist's view of objectivity. Is it necessary, in order to regard moral demands as objective, to think of their truth as being mind-independent? Might a somewhat weaker, antirealist account of objectivity not satisfy the relevant metaethical *desiderata* equally well – i.e. vindicate (among others) the suggestion that there is more to moral objectivity than intersubjective consensus, without positing the existence of a realm of moral facts that is fully independent of human attitudes? With regard to the issue of moral objectivity, the philosophical task we set ourselves in the ensuing chapters will be twofold: both to inquire how strong an account of objectivity is needed to vindicate the commitments

¹⁵ Surveying the beliefs of approximately 2000 professional philosophers, Bourget and Chalmers (2014) found that approximately 56 percent accepts or leans towards moral realism, whereas approximately 28 percent accepts or leans towards moral antirealism. Since different philosophers draw the boundary between moral realism and antirealism somewhat differently, this number should not be over-interpreted, but it does underline the scholarly relevance of discussing the merits of a realist metaethics.

internal to moral discourse and to scrutinize which account of objectivity is most plausible on external, scientific grounds.

1.4 Method

The realist and antirealist positions outlined in the previous section both make a substantive proposal about how claims to moral objectivity should be understood. Which of them is most plausible? In this section I outline the general methodology I follow to evaluate which view best satisfies the relevant metaethical *desiderata*.

1.4.1 External and Internal Accommodation

As Finlay (2007) observes, debates between realism and antirealism are typically shaped by the general dialectical pressures of *internal* and *external* accommodation. *External accommodation* is the pressure on metaethicists to articulate a view whose contents are compatible with knowledge that is external to the moral domain – for instance, with our scientific background knowledge. It is an uncontroversial *desideratum* of metaethical theories to accord with findings from evolutionary theory, psychology, anthropology and history, among other disciplines. Traditionally, it has been held that antirealists have an advantage over realists in terms of external accommodation: they are better able to reconcile their theory about the nature of morality with our best science.

Internal accommodation is the pressure on metaethicists to articulate a view best able to shed light on the relevant moral *explananda*. Consider our ordinary moral language, moral phenomenology, and people’s intuitions about moral discourse. As highlighted in the previous section, it is typically held that these commit us to the view that morality purports to be objective. One task for metaethicists, then, will be to shed light on these phenomena and explain why morality typically has these objective-seeming features. Traditionally, it has been held that realists have an advantage over antirealists in this regard. Indeed, as I discuss in **chapters 5, 6 and 7** of this dissertation, some metaethicists have advanced ‘presumptive arguments’ in support of moral realism on this basis: arguments to the effect that realism is the ‘view to beat’ in metaethics (Enoch 2018, p. 30).

1.4.2 Beyond the External–Internal Dichotomy

The distinction between external and internal accommodation has its limitations. First, what initially appears to be a weakness of a metaethical view in terms of its external validity may actually turn out to hinge on a trade-off between its external and internal validity. Secondly, what is internal and external to the moral domain is itself subject to change.

Consider the trade-off between external and internal validity. Many arguments that have been advanced against moral realism turn on the problematic features of a realist moral epistemology. One notorious problem is the metaethical analogue of the ‘Benacerraf-Field challenge’ for mathematical Platonism: if moral truths are presumed to be mind-independent and causally inert, then there seems to be no reason to suppose that the contents of our judgments coincide with these truths. This challenge can be overcome by stipulating that moral truths have causal powers and that perceiving or intuiting moral properties causes our judgments. This commits realists to the earlier-mentioned ‘truth-tracking’ thesis: in making moral judgments, we track mind-independent truths. But arguably, this tracking thesis is

untenable on scientific grounds (cf. Street 2006). If so, realists are faced with a dilemma: either their view leads to epistemological problems, or it is incompatible with our best science. In other words, realists are faced with a trade-off between vindicating the general reliability of our moral judgments (internal accommodation) and vindicating the scientific plausibility of their theory (external accommodation).¹⁶

Next, consider how the external/internal boundary might shift. One of the main internal *desiderata* for metaethicists is to illuminate the nature of moral discourse and of people's moral intuitions. But what intuitions people have is, to a certain extent, an empirical question. Conceptual analysis done from the armchair may shed some light on the answer, but so does rigorous experimental research about the nature of people's intuitions (see **chapter 6**). What phenomena metaethicists should illuminate, then, is partly beholden to scientific findings. Hence, the dichotomy between internal and external accommodation is not robust: scientific evidence can influence which internal *explananda* metaethicists should capture.

1.4.3 Holistic Evaluation

How should we evaluate which of the two metaethical views under consideration – realism or antirealism – is most plausible, all things considered? Following Carnap's (1947) recommendation, all the relevant evidence should be taken into account to estimate their respective probability. It is possible, however, that the evidence will be indecisive: perhaps realists are better at satisfying the metaethical *desideratum* of procuring morality's seeming objectivity, whereas antirealists are better at satisfying the *desideratum* of procuring an account that is scientifically plausible (Finlay 2007). This would leave the metaethical debate in a rather unsatisfying state: neither theory can capture the main metaethical *desiderata* fully. The question, then, becomes what is the greater sacrifice upon holistic evaluation: advancing a theory that conflicts with our best science or advancing a theory that cannot procure morality's internal commitments.

However, I do not think that the traditional assessment that antirealism wins in terms of external accommodation but realism wins in terms of internal accommodation is quite correct. In this dissertation I do vindicate the former claim: the best *scientific explanation* of the origins of our moral judgments proceeds on the assumption of moral antirealism rather than moral realism. But I also question the latter claim and argue that it is not at all clear that a realist view is superior in vindicating morality's internal commitments. In fact, matters are much less decisive here; arguably antirealism even does *better* than realism in this respect.

1.4.4 Debunking and Vindicating

As the preceding reflections reveal, it need not follow from the greater *scientific* plausibility of moral antirealism that antirealism is also *metaethically* superior. Scientific plausibility is not the only *desideratum* relevant for evaluating metaethical views. What matters is a total balance of evidence; if moral realism is only slightly less plausible in scientific terms but vastly superior in satisfying other important *desiderata*, then it might still be the superior metaethical view.

¹⁶ I spell out this argument in further detail in **chapter 3** when discussing Street's Darwinian Dilemma.

It is for this reason that the term ‘debunking’, in which contemporary discussions about the tenability of realism are often framed, can be somewhat misleading.¹⁷ Only rarely can a philosophical thesis be definitively disproven or shown to be fallacious. The same goes for the arguments against realism that I present here: they do not strictly show that realism is *false*, but rather that the position is *implausible*. The aim of my analysis, then, is not strictly speaking to *debunk* moral realism, at least not if this is understood as the proposition that all versions of realism are definitively rebutted. Nor is it to *vindicate* an account of objectivity in terms of stance-invariance, understood as the claim that this account must ultimately be correct. I do, however, claim that of the two accounts of objectivity discussed here – objectivity understood in terms of stance-independence versus objectivity understood in terms of stance-invariance – the latter is metaethically superior.

1.4.5 Empirical Metaethics

Charles Darwin realized that empirical findings can yield profound philosophical insights. As he scribbled in one of his *Notebooks* (1838): ‘He who understands baboon would do more towards metaphysics than Locke.’ This dissertation is written in the same spirit, and seeks to show that empirical evidence can be relevant for evaluating the adequacy of extant philosophical theses – while acknowledging that the metaethical implications of empirical findings are typically by no means straightforward. The majority of the empirical findings that will be discussed are experiments and observations of other authors. However, I have also conducted a survey experiment in moral psychology myself, together with my colleague Michael Klenk, the results of which are incorporated in **chapter 7** of the dissertation.

1.5 Preview of Chapters

Dissertations can be written in different ways – as one long argument seeking to establish a conclusion, or as collections of articles grouped around a coherent theme. The present dissertation is of the second kind: apart from this introduction and the conclusion, all of its chapters have been written as independent articles, aimed for publication in scholarly journals or volumes.¹⁸ Nonetheless, they constitute a coherent collection. **Chapters 2–5** roughly follow the dialectic sketched in the first section of this introduction: from challenging realism on the basis of scientific considerations to a positive defence of an antirealist view of objectivity. **Chapter 5**, as well as **chapter 6–7** engage in detail with findings from experimental moral psychology, and argue *inter alia* that these findings undermine so-called *presumptive arguments* in support of realism.

¹⁷ As far as I can tell, Lillehammer (2003) has been the first to use this term in the current debate.

¹⁸ As of 5 November 2018, the publication status of the dissertation chapters is as follows. **Chapter 2** is forthcoming in the journal *Ratio*, for a special issue on ‘Evolution and Moral Epistemology’. **Chapter 3** is forthcoming in the journal *Biology & Philosophy*. **Chapter 4** is under review by the journal *Philosophical Studies*, and will be accepted for publication pending minor revisions. **Chapter 5** has been published in the journal *Ethical Theory and Moral Practice*, for a special issue on ‘Objectivity in Ethics’ (2017). **Chapter 6** is forthcoming in the journal *Philosophical Psychology*. **Chapter 7** has been invited to be considered for publication in a forthcoming volume based on the conference ‘Evolutionary Debunking Arguments: The Nuts and Bolts Approach’ (Oxford Brookes 2018).

I proceed in **chapter 2** by discussing a recent epistemological challenge for moral realism: Sharon Street's (2016a) so-called practical/theoretical puzzle. This challenge is akin to the aforementioned Benacerraf-Field challenge for mathematical Platonism – the challenge of explaining how our mathematical judgments can be reliable if their truth-makers are mind-independent and causally inert facts. On Street's version, metanormative theorists are challenged to explain the presumed coincidence between the set of normative judgments that causal forces have led us to make and the set of true normative judgments. I argue that Street succeeds in discrediting specific realist views, but highlight that her puzzle does not pose a *general* challenge to normative realism. In order to present such a general challenge, debunkers also have to invoke the scientific claim embedded in the second horn of Street's Darwinian Dilemma (2006): the best explanation of our moral endorsements does not involve an appeal to mind-independent moral truths.

I discuss the Darwinian Dilemma in **chapter 3**, where I analyse what empirical considerations are relevant for EDAs against realism. I start by outlining what I take to be the strongest version of an EDA against moral realism. This EDA incorporates both Street's Darwinian Dilemma and Michael Ruse's (1995) claim that our moral norms are evolutionarily contingent. I argue that these arguments – especially a combination thereof – pose a serious challenge for some realist views, such as Enoch's (2010) third-factor theory. But other realists can escape the evolutionary challenge by denying the empirical premise which generates it – i.e. that our moral endorsements have been saturated with evolutionary influence. I argue that there are indeed good grounds for questioning the adequacy of Street's evolutionary claim, and the 'adaptive-link' account on which it rests. But I argue that realists who pursue this strategy may still be challenged on genealogical grounds, since they are committed to the thesis that over the course of history we have tracked stance-independent moral truths. This thesis is apt for consideration of historical truth or falsehood: in order to vindicate the realist's genealogical claim, our best historical explanations should indeed appeal to stance-independent moral truths.

This brings us to **chapter 4**, where I present a new 'historical debunking argument' against moral realism. I do so by taking issue with Michael Huemer's (2016) recent historical argument in support of moral realism. Huemer argues, on historical grounds, that realism is better placed than rival theories to explain the apparent moral convergence that has taken place over the course of human history, especially during the last few centuries. I argue that, even if we grant Huemer's historical *datum*, antirealists have underappreciated resources to elucidate the dynamics of historical moral change and convergence. These include an appeal to socialization, to lessons learned from history, to convergences in terms of technology and welfare, to changes induced by 'consistency reasoning' and to the social function of moral norms in overcoming some of the cooperation problems that globalizing societies face. I point out that an explanation that proceeds against the background of antirealism has several theoretical virtues and that an explanation that proceeds against a background of realism has several flaws; I conclude that on historical grounds the former is superior.

In **chapter 5** I outline and defend this antirealist account of moral objectivity in further philosophical detail, departing from Street's (2008) Humean constructivism. On Street's view, what makes a moral judgment true is that it withstands scrutiny from the evaluative standpoint of the agent who makes the judgment. Street's constructivist approach has merit, but should be expanded with an account of moral objectivity. With regard to moral judgments, I propose to

explicate the pre-theoretical notion of moral objectivity as follows: objective moral judgments have a capacity to withstand a process of *intersubjective scrutiny*. I further develop this account – objectivity as stance-invariance – and subsequently defend it against three objections that critics have raised: the ‘relativity objection’, the ‘revision objection’, and the ‘alignment objection’. I also present reasons for thinking that this account is able to capture the essential objectivist commitments of ordinary moral discourse.

The latter has to do with recent experimental findings about the nature of people’s objectivist intuitions, and brings us to **chapter 6**, in which I evaluate the metaethical relevance of experiments about folk objectivism more generally. Part of this chapter serves to show that such experiments are troubled by methodological difficulties, and underlines just how difficult it is to draw metaethical implications from empirical research. Nonetheless, I will argue that current findings do provide the beginnings of a new empirical framework about the psychosocial functions of metaethical claims, and that, at least indirectly, some research findings *are* metaethically relevant. *Inter alia* they cast doubt on presumptive arguments in support of moral realism.

In **chapter 7** I add to this work in experimental moral psychology by presenting the results of an experiment that I have conducted in collaboration with Michael Klenk. We start this chapter by reflecting on the question of whether findings and theories from evolutionary moral psychology debunk moral realism, as Ruse (1995) has suggested. We find Ruse’s EDA wanting, both on philosophical and on empirical grounds. We do, however, find recent scientific support for Ruse’s insight that regarding moral statements as objective might serve an evolutionary function, perhaps in ways more intricate than Ruse envisioned. Our experiment is designed to test the following hypothesis: the objectivity attributed to moral statements is at least partly modulated by an individual’s personal interests. We find support for this hypothesis, and discuss our results in the context of DeScioli and Kurzban’s (2013) recent conjecture that moral judgments serve the function of allowing third parties to strategically choose sides in cases of conflict. We expand on their hypothesis by proposing that *objectifying* moral judgments might serve to enhance the side-choosing strategy. We conclude that while our hypothesis, if correct, does not debunk moral realism, it does cast further doubt on presumptive arguments in support of realism.

Can we procure a sufficiently strong account of moral objectivity in the face of morality’s evolutionary origins? In the concluding **chapter 8** I highlight three original arguments, presented over the course of the preceding chapters, which shed light on this question. Firstly, I recapitulate my historical debunking argument and emphasize that, apart from appealing to evolutionary factors, debunkers also need to invoke historical considerations in order to advance a general empirical challenge against moral realism. Secondly, I recapitulate the importance of findings from experimental and evolutionary psychology in challenging presumptive arguments in support of realism. Thirdly, I return to the positive claim of this dissertation, namely that an antirealist view of moral objectivity is sufficiently strong in the light of the relevant metaethical *desiderata*. While moral realism may be untenable on scientific and philosophical grounds, this does not imply that morality’s objectivity cannot be vindicated. We don’t need to be realists in order to be objectivists; in the face of morality’s evolutionary and historical origins, moral objectivity is better understood in antirealist terms.

References

- Alexander, R.D. 1987. *The Biology of Moral Systems*. New York: Aldine de Gruyter.
- Artiga, M. 2015. Rescuing Tracking Theories of Morality. *Philosophical Studies* 172: 3357-3374.
- Baumard, N., André, J.B., Sperber, D. 2013. A Mutualistic Approach to Morality: The Evolution of Fairness by Partner Choice. *Behavioral and Brain Sciences* 36: 59-122.
- Beebe, J.R. 2014. How Different Kinds of Disagreement Impact Folk Metaethical Judgments. In: J.C. Wright and H. Sarkissian, eds. *Advances in Experimental Moral Psychology: Affect, Character, and Commitments*. London: Continuum, 167-187.
- Berker, S. 2014. Does Evolutionary Psychology Show that Normativity is Mind-Dependent? In: J. D'Arms and D. Jacobson, eds. *Moral Psychology and Human Agency: Essays on the Science of Ethics*. New York: Oxford University Press, 215-252.
- Bourget, D. and Chalmers, D.J. 2014. What Do Philosophers Believe? *Philosophical Studies* 170: 465-500.
- Brink, D. 1989. *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- Brosnan, K. 2011. Do the Evolutionary Origins of our Moral Beliefs undermine Moral Knowledge? *Biology & Philosophy* 26: 51-64.
- Brosnan, S.F and de Waal, F. 2003. Monkeys Reject Unequal Pay. *Nature* 425: 297-299.
- Buchanan, A. and Powell, R. 2015. The Limits of Evolutionary Explanations of Morality and Their Implications for Moral Progress. *Ethics* 126: 37-67.
- Carnap, R. 1947. On the Application of Inductive Logic. *Philosophy and Phenomenological Research* 8: 133-148.
- Chudek, M. and Henrich, J. 2011. Culture–Gene Coevolution, Normpsychology and the Emergence of Human Prosociality. *Trends in Cognitive Sciences* 15: 218-226.
- Clarke-Doane, J. 2016. Debunking and Dispensability. In: U.D. Leibowitz and N. Sinclair, eds. *Explanation in Ethics and Mathematics: Debunking and Dispensability*. Oxford: Oxford University Press, 23-36.
- Copp, D. 2008. Darwinian Skepticism about Moral Realism. *Philosophical Issues* 18: 186-206.
- Cuneo, T. and Shafer-Landau, R. 2014. The Moral Fixed Points. New Directions for Moral Nonnaturalism. *Philosophical Studies* 171: 399-443.
- Curry, O. 2016. *Morality as Cooperation: A Problem-Centred Approach*. In T. K. Shackelford and R. D. Hansen eds. *The Evolution of Morality*. Heidelberg: Springer International Publishing, 27-51.
- Das, R. 2015. Evolutionary Debunking of Morality: Epistemological or Metaphysical? *Philosophical Studies* 173: 417-435.
- Deem, M.J. 2016. Dehorning the Darwinian Dilemma for Normative Realism. *Biology & Philosophy* 31: 727–746.
- DeScioli, P. and Kurzban, R. 2009. Mysteries of Morality. *Cognition* 112: 281-299.

- DeScioli, P. and Kurzban, R. 2013. A Solution to the Mysteries of Morality. *Psychological Bulletin* 139: 477-496.
- Enoch, D. 2007. An Outline of an Argument for Robust Metanormative Realism. In: R. Shafer-Landau, ed. *Oxford Studies in Metaethics 2*. New York: Oxford University Press, 21-50.
- Enoch, D. 2010. The Epistemological Challenge to Metanormative Realism: How Best to Understand it, and How to Cope with It. *Philosophical Studies* 148: 413-438.
- Enoch, D. 2014. Why I Am an Objectivist about Ethics (And Why You Are, Too). In: R. Shafer-Landau, ed. *The Ethical Life*. Oxford: Oxford University Press, 192-205.
- Enoch, D. 2016. Dispensability Arguments in Metaethics: Even better than Mathematics? In: U.D. Leibowitz and N. Sinclair, eds. *Explanation in Ethics and Mathematics: Debunking and Dispensability*. Oxford: Oxford University Press, 236-254.
- Enoch, D. 2018. Non-Naturalistic Realism in Meta-Ethics. In: T. McPherson and D. Plunkett, eds. *The Routledge Handbook of Metaethics*. New York: Routledge, 29-42.
- Finlay, S. 2007. Four Faces of Moral Realism. *Philosophy Compass* 6: 820-849.
- FitzPatrick, W. 2008. Robust Ethical Realism, Non-Naturalism, and Normativity. In: R. Shafer-Landau, ed. *Oxford Studies in Metaethics, vol. 3*. Oxford: Clarendon Press, 159-205.
- FitzPatrick, W. 2014. Why There Is no Darwinian Dilemma for Ethical Naturalism. In: M. Bergmann and P. Kain, eds. *Challenges to Moral and Religious Belief: Disagreement and Evolution*. Oxford: Oxford University Press, 237-255.
- FitzPatrick, W. 2015. Debunking Evolutionary Debunking of Ethical Realism. *Philosophical Studies* 172: 883-904.
- Fraser, B. and Sterelny, K. 2017. Evolution and Moral Realism. *The British Journal for the Philosophy of Science* 68: 981-1006.
- Goodwin, G.P. and Darley, J.M. 2012. Why Are some Moral Beliefs Perceived to be More Objective than Others? *Journal of Experimental Social Psychology* 48: 250-256.
- Gould, S.J. and Lewontin, R.C. 1979. The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proceedings of the Royal Society of London B* 205: 581-598.
- Greene, J. 2014. Beyond Point-and-Shoot Morality: Why Cognitive (Neuro)Science Matters for Ethics. *Ethics* 124: 695-726.
- Habermas, J. 1995. Reconciliation through the Public Use of Reason: Remarks on John Rawls's Political Liberalism. *The Journal of Philosophy* 92: 109-131.
- Hamilton, W.D. 1963. The Evolution of Altruistic Behavior. *American Naturalist* 97: 354-356.
- Hamlin, K. 2014. The Origins of Human Morality: Complex Socio-Moral Evaluations by Preverbal Infants. *New Frontiers Social Neuroscience* 21: 165-188.
- Hanson, L. 2016. The Real Problem with Evolutionary Debunking Arguments. *The Philosophical Quarterly* 67: 508-533.
- Henrich, J. 2015. *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton (NJ): Princeton University Press.
- Horgan, T. and Timmons, M. 2006. Expressivism, yes! Relativism, no! In: R. Shafer-Landau, ed. *Oxford Studies in Metaethics, vol. 1*. Oxford: Clarendon Press, 73-98.

- Hrdy S.B. 2009. *Mothers and Others: The Evolutionary Origins of Mutual Understanding*. Cambridge: Harvard University Press.
- Huemer, M. 2005. *Ethical Intuitionism*. New York: Palgrave Macmillan.
- Huemer, M. 2016. A Liberal Realist Answer to Debunking Skeptics: The Empirical Case for Realism. *Philosophical Studies* 173: 1983-2010.
- Hume, D. 1738. *A Treatise of Human Nature*, Book III, part I, section I.
- Joyce, R. 2006. *The Evolution of Morality*. Cambridge: MIT Press.
- Joyce, R. 2016. *Essays in Moral Skepticism*. Oxford: Oxford University Press.
- Kahane, G. 2011. Evolutionary Debunking Arguments. *Noûs* 45: 103-125.
- Kitcher. P. 1998. Psychological Altruism, Evolutionary Origins, and Moral Rules. *Philosophical Studies* 89: 283-316.
- Kitcher, P. 2011. *The Ethical Project*. Cambridge: Harvard University Press.
- Leibowitz, U.D. and Sinclair, N. eds. 2016. *Explanation in Ethics and Mathematics: Debunking and Dispensability*. Oxford: Oxford University Press.
- Lillehammer, H. 2003. Debunking Morality: Evolutionary Naturalism and Moral Error Theory. *Biology & Philosophy* 18: 567-581.
- Mackie, J.L. 1977. *Ethics: Inventing Right and Wrong*. New York: Penguin.
- Mogensen, A.L. 2015. Evolutionary Debunking Arguments and the Proximate/Ulimate Distinction. *Analysis* 75: 196-203.
- Moore, G.E. 1993 [1903]. *Principia Ethica*. Cambridge: Cambridge University Press.
- Nagel, T. 1986. *The View from Nowhere*. New York: Oxford University Press.
- Nagel, T. 2012. *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature is almost certainly False*. Oxford: Oxford University Press.
- Nowak, M. 2012. *Supercooperators: Altruism, Evolution, and Why We Need Each Other to Succeed*. New York: Free Press.
- Nozick, R. 2001. *Invariances: The Structure of the Objective World*. Cambridge: Belknap Press of Harvard University Press.
- Pölzler, T. 2017. Are Moral Judgements Adaptations? Three Reasons Why It Is So Difficult to Tell. *South African Journal of Philosophy* 36: 425-439.
- Putnam, H. 2004. *Ethics without Ontology*. Cambridge: Harvard University Press.
- Rawls, J. 1971. *A Theory of Justice*. Cambridge: Harvard University Press.
- Richerson, P.J. and Boyd, R. 2005. *Not by Genes Alone: How Culture Transformed Human Evolution*. Chicago: The University of Chicago Press.
- Ruse, M. 1995. Evolutionary Ethics: A Phoenix Arisen. In: P. Thompson, ed. *Issues in Evolutionary Ethics*. Buffalo (NY): SUNY Press, 225-247.
- Sarkissian, H., Parks, J., Tien, D., Wright, J.C. and Knobe, J. 2011. Folk Moral Relativism. *Mind & Language* 26: 428-505.
- Schechter, J. 2018. Explanatory Challenges in Metaethics. In: T. McPherson and D. Plunkett, eds. *The Routledge Handbook of Metaethics*. New York: Routledge, 443-459.

- Schloss, J. 2014. Evolutionary Accounts of Morality: Accounting for the Normal but not the Normative. In: H. Putnam, S. Neiman and J. Schloss, eds. *Understanding Moral Sentiments: Darwinian Perspectives*. New Brunswick: Transaction Publishers, 81-121.
- Sen, A.K. 1993. Positional Objectivity. *Philosophy & Public Affairs* 22: 126-145.
- Shafer-Landau, R. 2003. *Moral Realism: A Defence*. Oxford: Oxford University Press.
- Shafer-Landau, R. 2012. Evolutionary Debunking, Moral Realism and Moral Knowledge. *Journal of Ethics and Social Philosophy*, www.jesp.org, 7(1): 1-37.
- Skarsaune, K.O. 2011. Darwin and Moral Realism: Survival of the Iffiest. *Philosophical Studies* 152: 229-243.
- Stanford, P.K. 2018. The Difference Between Ice Cream and Nazis: Moral Externalization and the Evolution of Human Cooperation. *Behavioral and Brain Sciences* 41: 1-57.
- Street, S. 2006. A Darwinian Dilemma for Realist Theories of Value. *Philosophical Studies* 127: 109-166.
- Street, S. 2008. Constructivism about Reasons. *Oxford Studies in Metaethics* 3: 207-45.
- Street, S. 2010. What is Constructivism in Ethics and Metaethics? *Philosophy Compass* 5: 363-384.
- Street, S. 2016a. Objectivity and Truth: You'd Better Rethink It. In: R. Shafer-Landau, ed. *Oxford Studies in Metaethics*, vol. 11. Oxford: Clarendon Press, 293-334.
- Street, S. 2016b. Constructivism in Ethics and the Problem of Attachment and Loss. *Aristotelian Society Supplementary Volume xc*: 161-189.
- Tomasello, M. 2016. *A Natural History of Human Morality*. Cambridge: Harvard University Press.
- Trivers, R.L. 1971. The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology* 46: 35-57.
- Tropman, E. 2014. Evolutionary Debunking Arguments: Moral Realism, Constructivism, and Explaining Moral knowledge. *Philosophical Explorations* 17: 126-140.
- Vavova, K. 2014. Debunking Evolutionary Debunking. In: R. Shafer-Landau, ed. *Oxford Studies in Metaethics*, vol. 9: 76-101.
- Waal de, F. 2006. *Primates and Philosophers: How Morality Evolved*. Princeton: Princeton University Press.
- Warneken F. and Tomasello, M. 2011. The Roots of Human Altruism. *British Journal of Psychology* 100: 455-471.
- Warneken, F. 2018. How Children Solve the Two Challenges of Cooperation. *Annual Review of Psychology* 69: 205-229.
- Westermarck, E.A. 1906. *The Origin and Development of the Moral Ideas*. London: Macmillan.
- Wielenberg, E. 2010. On the Evolutionary Debunking of Morality. *Ethics* 120: 441-464.
- Wielenberg, E. 2016. Ethics and Evolutionary Theory. *Analysis* 76: 502-515.
- Wilson, D.S. and Sober, E. 1994. Reintroducing Group Selection to the Human Behavioural Sciences. *Behavioural and Brain Sciences* 17: 585-654.
- Wright, J.C., Grandjean, P.T. and McWhite, C.B. 2013. The Meta-Ethical Grounding of Our

Chapter 1

Moral Beliefs: Evidence for Metaethical Pluralism. *Philosophical Psychology* 26: 336-361.

2

Chapter 2

Striking Coincidences: How Realists Should Reason about Them

Forthcoming in *Ratio*, special issue on ‘Evolution and Moral Epistemology’.

Keywords

*Practical/theoretical puzzle; coincidence challenge; normative epistemology;
reference class problem; third-factor accounts*

Abstract

Many metaethicists assume that our normative judgments are both by and large true, and the product of causal forces. In other words, many metaethicists assume that the set of normative judgments that causal forces have led us to make largely coincides with the set of true normative judgments. How should we explain this coincidence? This is what Sharon Street (2016) calls the practical/theoretical puzzle. Some metaethicists can easily solve this puzzle, but not all of them can, Street argues; she takes the puzzle to constitute a specific challenge for normative realism. In this chapter I elucidate Street's puzzle and outline possible solutions to it, framed in terms of a general strategy for reasoning about coincidences. I argue that the success of Street's challenge crucially depends on how we set the 'reference class' of normative judgments that we *could* have endorsed, assuming realism. I conclude that while the practical/theoretical puzzle falls short of posing a general challenge for normative realism, it can be successful as a selective challenge for specific realist views.

'Singular coincidence, Holmes. Very smart of you to notice it, but rather uncharitable to suggest that it was cause and effect.'

– Sir Arthur Conan Doyle (1913)

2.1 Introduction

On the evening of 11 February 2013, shortly after Joseph Ratzinger – alias Pope Benedict XVI – had taken many by surprise by announcing his decision to abdicate from his papal office, lightning struck St. Peter's Basilica in Vatican City. The image was captured on camera and rapidly spread via news agencies and social media. Several Twitter-users responded incredulously, claiming that the picture was either a hoax or a sign of divine intervention. This could not be *just a coincidence*; something else had to be going on.

To express incredulity is a common response when faced with a striking coincidence. Striking coincidences trigger epistemic anxiety: they appear to be too unlikely, too ingeniously set up or too miraculous to be accepted as *mere coincidence*. These coincidences call for further context and explanation, which may serve either to vindicate the intuition that an event is no mere coincidence or to debunk this intuition. In the former case, some alternative hypothesis is devised to account for the coincidental event. In the latter case, support is given to the null hypothesis that the coincidence is due to chance. In either case, the aim of providing further context and explanation is to alleviate the epistemic anxiety triggered by the initial observation. Coincidences may be explained or explained away, but they should not be ignored – otherwise we are left in an epistemically unsettling position.

What holds for coincidences in Vatican City also holds for coincidences in metaethics: they require some kind of context or explanation in order to generate epistemic ease. This explanatory demand is at the root of a metaethical challenge that Sharon Street (2016) has recently raised, which she calls the practical/theoretical puzzle. While this puzzle – outlined in **section 2.3** – is faced by many metaethicists, it is specifically troubling for normative realists, Street argues: she holds that realists are incapable of generating the desired epistemic ease.

The aim of the present chapter is to elucidate Street's puzzle in terms of a general strategy for reasoning about coincidences. I point out that the ambition she formulates at the beginning of her paper – to provide 'a completely general argument against normative realism, whether of a naturalist or non-naturalist variety' (idem, p. 300) – cannot be delivered by her puzzle alone. However, I argue that the puzzle can be successful as a *selective* challenge for a specific version of normative non-naturalism, which relies on a problematic 'third-factor account'.

The chapter proceeds as follows. In **section 2.2** I give a characterization of coincidences, distinguish between some of their defining features and outline a general framework for reasoning about them. In **section 2.3** I introduce Street's practical/theoretical puzzle, and argue that this general framework for reasoning about coincidences is also applicable to her puzzle. In **sections 2.4** and **2.5** I evaluate whether realists can solve the puzzle by pursuing the null hypothesis that the coincidence between the contents of our normative judgments and the mind-independent normative truths is due to chance. I argue that they can do so by attempting to narrow down the reference class of possible normative truths, assuming

realism. This move requires realists to provide some ‘alternative explanation’ of the coincidence. I discuss different alternative explanations in **section 2.6**, and highlight that at least one of them – Enoch’s (2010) third-factor account – does not solve the puzzle adequately. But I conclude in **section 2.7** that in order to provide a *general* challenge to normative realism, Street’s puzzle requires the support of further arguments, such as her Darwinian Dilemma (2006).

2.2 How Should We Reason about Coincidences?

Before I turn to Street’s metaethical puzzle, I begin this section by disambiguating some characteristic features of coincidences. In doing so, I will work towards outlining the general steps involved in reasoning about them.

2.2.1 Defining Coincidences: Causal versus Psychological Characteristics

On Diaconis and Mosteller’s (1989, p. 853) influential definition, a coincidence is ‘a surprising concurrence of events, perceived as meaningfully related, with no apparent causal connection.’ Consider a famous example from Aristotle’s *Physics*: a man goes to the market in order to see a spectacle and runs into a debtor of his, who has just sold his cattle, which allows him to instantly repay his debt. There is no causal connection explaining why the man and his debtor go to the market simultaneously and encounter each other under just the right circumstances for the debt to be repaid. Given the absence of this connection, the meaningful concurrence is likely to strike us (and particularly them) as surprising: it is a prototypical coincidence.¹⁹

The definition by Diaconis and Mosteller captures two key characteristics of coincidences: the absence of a *causal connection* between coinciding phenomena and the *psychological* response coincidences engender. Prototypical coincidences often have both of these characteristics, although the psychological element is dispensable. Borrowing an example from Street (2016), consider the concurrent presence of McDonald’s franchises and M-shaped golden arches on poles. This coincidence is entirely mundane and does not generate any surprise, but there is a colloquial sense in which it is a coincidence still: it constitutes the concurrent presence of two items that appear to be meaningfully related.

2.2.2 Coincidence Observation versus Coincidence Explanation

On a narrower definition, then, coincidences can be characterized in purely causal or explanatory terms. For instance, Owens (1992, p. 13) defines a coincidence as ‘an event which can be divided into components separately produced by independent causal factors.’ Along similar lines, Mogensen (forthcoming, p. 9) defines a coincidence as ‘a conjunction of facts whose conjuncts are explanatorily independent of one another: neither fact figures in the explanation of the other, and there is no relevant explanatory factor shared by the members of the conjunction.’ Both of these definitions characterize coincidences in terms of causal or explanatory independence. Mogensen’s proposal additionally captures the intuition that once we can point out some explanatory connection between the concurrent events, we typically no longer regard their concurrence as coincidental.

¹⁹ Psychological research suggests that coincidences that occur in one’s own life are typically regarded as more striking than coincidences that occur in the lives of others (Falk 1989).

As this last intuition suggests, whether we regard a concurrence of events as coincidental is subject to change: what we initially regard as a coincidence, we may no longer regard as such once we have explained it. To make this intuition more precise, it is useful to follow Sober (2012) in distinguishing between ‘coincidence observations’ and ‘coincidence explanations’. Street’s example of the concurrent presence of McDonald’s franchises and M-shaped golden arches on poles is a coincidence observation: the observation of two items that appear to be meaningfully related. But this is not to say that their concurrent presence is best explained as *mere coincidence*. In fact, it clearly is not: what explains their concurrent presence is the fact that the presence of McDonald’s franchises and M-shaped golden arches on poles have a common cause. Yet for some concurrences no such causal explanation is available; our best explanation is that two items just happened to coincide. The coincidence is just that: mere coincidence. This is what Sober calls a coincidence explanation, which, on his account, is characterized by the absence of a causal explanation for the concurrence of events (Sober, 2012 p. 362). Each event was produced via an independent causal process; their concurrence is due to chance.

2.2.3 Steps in Reasoning about Coincidences

Building on Sober’s distinction, we can now identify two steps involved in reasoning about coincidences. First, we can observe that two (or more) events or items appear to be related. Next, we can assess which of two hypotheses is the most plausible: the coincidence is due to chance (a coincidence explanation) or there is some alternative explanation for it.

1. Coincidence observation: Two (or more) concurrent events appear to be related.

2a. Coincidence explanation:

- There is no causal relation between the observed events;
- There is no defeater for the null hypothesis that the coincidence is due to chance;
- Further context may be provided to support the hypothesis that this is ‘mere coincidence.’

2b. Alternative explanation:

- There is some causal relation between the observed events;
- This serves as a defeater for the hypothesis that the coincidence is due to chance;
- The causal explanation serves to debunk the hypothesis that this is ‘mere coincidence.’

Figure 2.1: Steps in Reasoning about Coincidences

2.2.4 Why do Coincidences Often Strike Us as Surprising?

These steps comprise our general strategy for reasoning about coincidences; in the following sections I will apply it in a metaethical context. But before we turn to metaethics, let’s revisit the psychological aspect that we disentangled from Diaconis and Mosteller’s definition. Coincidences often strike us as surprising. Why so? Typically, coincidence observations lend themselves to explanations that conflict with our prevailing models about how the world works. They provide *prima facie* support for some alternative to a currently favoured theory, but this alternative has a lower prior probability than the theory currently favoured (cf. Griffiths and

Tenenbaum 2007). As a result, the observations steer us into an epistemic twilight zone. Consider that coincidences can be a source of superstition and conspiracies, but also a source of scientific discoveries. Is the apparent meaningfulness we observe noise that should be discarded, or a signal that should be picked up? If it turns out to be noise, then the null hypothesis that a coincidence observation is due to chance (i.e. the coincidence explanation) is correct; if it is a signal that should be picked up, then the hypothesis that the concurrence is no mere coincidence (i.e. the alternative explanation) is correct.

Following Griffiths and Tenenbaum, this understanding of coincidences can be stated in Bayesian terms. The likelihood of a hypothesis h with respect to evidence e is defined as the probability that h confers on e , expressed as $P(e|h)$. By contrast, the prior probability of a hypothesis h is defined as the probability that h is true on the basis of our background knowledge k before evidence e is taken into account, expressed as $P(h|k)$. Consider the hypothesis that papal abdications and thunderstorms in Vatican City are causally related. The prior probability $P(h|k)$ of this hypothesis is rather low, given our background knowledge about causal relations, meteorology, *etc.* However, the hypothesis has a high likelihood $P(e|h)$: if papal abdications and thunderstorms in Vatican City are indeed causally related, then we might expect lightning to have struck St. Peter's Basilica on 11 February 2013. Hence, although the hypothesis has a low prior probability, it does have a high likelihood – and this contrast is precisely what makes the observed coincidence particularly striking.

2.3 The Practical/Theoretical Puzzle

In this section I outline Street's practical/theoretical puzzle and show how we may approach it using the same two strategies for reasoning about coincidences. Borrowing Kantian terminology, Street observes that from a 'practical standpoint' – i.e. when we reflect on what courses of action are good, valuable and worthwhile – we think of ourselves as beings capable of making normative judgments that are by and large true. From a 'theoretical standpoint' – i.e. when we understand ourselves as beings who are part of the world of cause and effect – we understand these judgments as products of causal forces and subject to scientific explanation, broadly understood (e.g. explicable in terms of our cultural upbringing, psychological inclinations, social pressures, natural selection, *etc.*). Combining these two standpoints, we arrive at the following: we typically assume that the set of practical normative judgments that causal forces have led us to make roughly coincides with the set of true practical normative judgments.²⁰ How should we account for this coincidence? This is what Street calls the practical/theoretical puzzle.

2.3.1 Clarifying the Puzzle and its Target

²⁰ Here I follow Street (2016) in restricting the scope of the puzzle to *practical* normativity and excluding *epistemic* normativity (although Street (*idem*, fn. 2, pp. 293-294) herself suggests that the challenge might carry over to the epistemic domain as well). In what follows, then, the terms 'normativity' and 'normative reasons' are meant to be shorthand for 'practical normativity' and 'practical normative reasons'. In the ensuing chapters of this dissertation I will restrict the scope of my treatment even further and discuss evolutionary debunking arguments specifically in the context of *moral* normativity.

Let's start by clarifying and elaborating on the puzzle in four respects. First, Street states that this is a puzzle that *all* metaethicists, including herself, face. This is, in fact, a hasty generalization: it is a puzzle only faced by metaethicists who think that our normative judgments are products of causal forces and that these judgments are largely true. While this includes Street's Humean constructivism, it excludes error theorists, who maintain that our normative judgments are (typically) false. Street does not consider this option; I will follow her and set it aside for the remainder of this chapter.²¹

Second, Street does not mean to suggest that the puzzle cannot be solved. In fact, she maintains that this *can* be done, albeit not by metaethicists of all stripes. Specifically, she argues that non-naturalist realists, according to whom the truth-makers of normative or evaluative judgments are mind-independent and causally inert properties or facts, are ill-equipped to solve the puzzle.²² As Street herself notes, this makes the practical/theoretical puzzle similar to the Benacerraf-Field challenge for mathematical Platonism.

Third, it might be thought that Street's puzzle is not only problematic for realists regarding practical normativity but also for realists regarding sense perception judgments. After all, we also think that our judgments about manifest surroundings are by and large true and products of causal forces. In other words, we assume that the set of empirical judgments that causal forces have led us to make roughly coincides with the set of true empirical judgments. Street, however, argues that this coincidence (i.e. the coincidence observation) can be explained (i.e. can be given an alternative explanation): the objects in our manifest surroundings are things with causal powers, and the evolutionary fitness of our ancestors has been promoted by the formation of veridical perceptions of such objects. Normative realists who deny that normative properties have causal powers cannot appeal to an analogous explanation. Therefore, Street thinks she can advance her puzzle as a challenge specific to realism about the *normative* domain.

Fourth, as I will illustrate in the upcoming sections, Street's puzzle – and strategies for solving it – are better appreciated once reframed in terms of the general strategy for reasoning about coincidences outlined in **section 2.2**. Street's challenge starts with a presumed coincidence observation: presumably, the set of true normative judgments and the set of normative judgments that causal forces have led us to make roughly coincide. We should keep in mind that this coincidence is not directly *observed*, but *conjectured*. For present purposes, however, this disanalogy is inconsequential. Since (*ex hypothesi*) all metaethicists want to preserve the idea that there is a coincidence between these two sets, they have to pursue one of the two aforementioned explanatory strategies: to provide a coincidence explanation or an alternative explanation for the conjectured coincidence.

2.3.2 Reframing the Practical/Theoretical Puzzle in terms of Grounding

²¹ Street (*idem*, p. 295) seems to think that it is a *desideratum* to preserve the idea that our normative judgments are typically true: she calls both our practical and theoretical perspectives 'inescapable' and expresses the 'hope' that they can be reconciled.

²² Street uses the terms 'normative' and 'evaluative' judgments interchangeably; in this dissertation I will do so as well.

Before we go into strategies for solving the puzzle, it may be helpful to reframe it in the language of grounding.²³ ‘Grounding’ may be understood as a relation of metaphysical priority in which the *relata* are properties or facts. The grounding relation is often expressed using other notions, such as ‘because of’ or ‘in virtue of’. Saying that A *grounds* N is typically another way of saying that A *makes it the case* that N obtains, N obtains *in virtue of* A or N obtains *because* A obtains. While grounds talk has only recently become en vogue in philosophy, the grounding relation to which such talk refers is by no means new. Following Berker (2014, 2018, forthcoming), grounding is not a technical term; philosophers have discussed the metaphysical dependence relation it picks out for centuries – indeed, since the dawn of philosophy. Consider the question that Socrates asks in Plato’s dialogue *Eutyphro*: ‘Is the pious loved by the gods because it is pious, or is it pious because it is loved by the gods?’ Socrates is asking what depends on what, and the dependence here is not causal, but metaphysical in nature; in other words, he is asking about their grounding relation.

Confusingly, in ordinary language the conjunction ‘because’ is commonly used for explanations in terms of causes as well as grounds. Yet causal and grounding explanations are completely independent. Scientists tend to be occupied with the former, whereas philosophers – especially metaphysicians – have a specific interest in the latter. Indeed, grounding can be regarded as the philosophical counterpart of causation, as grounding pioneer Kit Fine remarks: ‘Ground, if you like, stands to philosophy as cause stands to science’ (Fine 2012, p. 40). This division is not absolute, however; for instance, both kinds of explanation play a role in metaethics. This brings us back to the practical/theoretical puzzle, which we can understand as the challenge of reconciling these two kinds of explanation with regard to the contents of our normative judgments. From the theoretical standpoint we ask what *causes* the contents of our judgments. From the practical standpoint we ask what *grounds* the truth of our judgments – i.e. what are their truth-makers. Hence, what Street calls the practical/theoretical puzzle might also be called the ‘grounding/causal puzzle’: the puzzle of coming up with a harmonious explanation of the causes and grounds of true normative judgments.

2.3.3 Solving the Puzzle? The Realist’s Coincidence Explanation and Street’s Criticism

Street (2016) discusses Ronald Dworkin (1996) as the protagonist of realists who rely on a coincidence explanation. According to Street,

Dworkin’s answer to the practical/theoretical puzzle is that the coincidence in question is just that – a coincidence; one ought to understand it as a ‘piece of luck’ (Dworkin 1996, p. 125) that the true normative judgments, on the one hand, and the normative judgments that causes led one to make, on the other, line up as much as they do. (Street 2016, pp. 309-310)

²³ Grounding is a recent and much discussed topic in current analytic metaphysics; unsurprisingly, controversies on the topic abound. Following Berker (2018a), here I assume that grounding is best understood as a *relation*, that the grounding relation is *monistic* and that it is *explanatory* – yet none of these assumptions is uncontroversial. The controversies are not directly relevant to our current treatment, however. See Berker’s lucid treatment for further detail.

Dworkin's position is epistemically unsatisfactory, Street argues: it is akin to a person who claims to have won a low-odds lottery merely on the basis of having entered it. After all, there are countless sets of normative judgments that we could have endorsed. Assuming that there is a unique set of mind-independent normative truths, there are countless ways in which our normative judgments could be 'off-track'. Given this predicament, the realist should do more to explain how our normative judgments have (roughly) been able to track the independent normative truths. In the absence of such an explanation, realists should conclude that they probably didn't win the normative lottery: they are probably mistaken about the contents of the mind-independent normative truths.

Is Street's dismissal of the realist's coincidence explanation successful? How, in fact, can we determine the odds of this 'metaethical lottery'? In the next two sections I will explore the options realists have to challenge Street's contention that the odds of this lottery are very unfavourable for realism. I will argue that realists have the resources to do so, but only if they provide the outlines of a positive epistemology. In other words, a 'pure coincidence explanation' will not suffice to deflect Street's challenge; additionally, realists should provide some alternative explanation.

2.4 Pursuing a Coincidence Explanation

The general strategy when offering a coincidence explanation is to point to contextual factors which make the prior coincidence observation more likely to have occurred. Consider once more the observation that lightning struck St. Peter's Basilica on the same day that the pope announced his abdication. Suppose that one is presented with two rival hypotheses: either this concurrence was mere coincidence (the coincidence explanation) or it was the result of divine intervention (the alternative explanation). The following considerations are relevant to weighing these hypotheses and to raising the probability of the former with respect to the latter:

A1) High likelihood of the observation given the null hypothesis. Satellite research suggests that, around the globe, lightning flashes around 40 times per second; even though not all flashes hit the Earth, this engenders an enormous potential source of apparently meaningful impacts. The observed coincidence may be regarded as an instance of what Diaconis and Mosteller (1989) call 'the law of truly large numbers': with a large enough sample, surprising concurrences are bound to be plentiful.

A2) Low likelihood of the observation given the alternative hypothesis. Additionally, one might argue that given the alternative hypothesis, the coincidence was in fact not that likely to occur. Consider the apparent 'design' of the situation, which raises the initial suspicion of a divine set-up: this design is not nearly as perfect as it first appears. Lightning only struck in the evening, many hours after the pope had voiced his impending abdication. What caused this delay in the divine response? Wouldn't the coincidence have made much more impact if God had struck immediately? And why didn't He express His fury in a more efficient manner, say by electrocuting the pope, rather than by directing his rage at a basilica? *Post facto* auxiliary assumptions that purport to increase the probability of the hypothesis (e.g. by rationalizing why God aimed His fury at the basilica rather than the pope) are bound to be *ad hoc*.

A3) *Low prior probability of the alternative hypothesis.* Indeed, one might argue that the alternative hypothesis is rather implausible to begin with. I single out one of many relevant considerations: the alternative hypothesis exhibits a combination of characteristics that is typical of conspiracy theories, such as hindsight bias, selective use of evidence and *ad hoc* hypothesizing. With hindsight, the concurrence of the pope's announcement and the lightning strike appears remarkable, yet beforehand no one had predicted this concurrence of events in any detail. Any hypothesis which displays these theoretical vices should be held in low epistemic regard.

These three considerations mutually support each other: while none of them may be conclusive, in combination they give substantial support to the hypothesis that the concurrence of events on 13 February 2013 was due to chance.

Now return to the practical/theoretical puzzle. In roughly analogous fashion, realists might push back against Street's challenge by arguing as follows:

B1) *High likelihood of the coincidence given realism.* If there is only a limited number of ways in which causal forces might have led us to endorse normative judgments that do not coincide with the normative truth, then the suggestion that such a coincidence *has* in fact occurred becomes much more palatable. Accordingly, realists might argue that there is only a limited number of ways our normative judgments might have been constituted. Given these limited possibilities, it is not much of a miracle that causal forces roughly made them coincide with the set of normative truths.

B2) *Low likelihood of the coincidence given the alternative hypothesis.* The success of Street's alternative explanation hinges on the assumption that if normative truth is mind-dependent, then we *can* be reliable truth-trackers. But it might be argued that assuming Street's version of antirealism – Humean constructivism (see **chapter 5**) – it is not at all clear that our normative judgments will roughly coincide with the mind-dependent normative truths.²⁴

B3) *Low prior probability of the alternative hypothesis.* It might be argued that Humean constructivism is a rather implausible theory to begin with, for example because it fails to vindicate the intuition that normative truths are robustly objective: the view commits us to an implausible degree of relativism.

Of course, these propositions need to be supported with further argument. Here, I merely mention **B2** and **B3** to illustrate that the plausibility of a coincidence explanation partly hinges on the merits of the alternative explanation with which it contends. We are engaged in a holistic explanatory game: the success of any given hypothesis is relative to the failure of its rivals.

For present purposes, let's set Street's alternative explanation aside and focus

²⁴ Berker (2014, p. 234) argues for this claim. I do not find Berker's considerations persuasive, however; given the shared evolutionary origins of our basic moral judgments, antirealists can make a good case for arguing that we are fairly reliable at tracking each other's moral reasons. I argue against Berker's criticism in **chapter 5.7** of this dissertation, and against the most widely endorsed type of argument for thinking that realism is in a better shape overall (so-called presumptive or experiential arguments in support of realism) in **chapter 6.5**.

specifically on **B1** – the likelihood of the hypothesis that the contents of our normative judgments have accidentally coincided with the mind-independent normative truths, *assuming realism*. Following Enoch (2010, p. 427), we should note that the conjectured coincidence is imperfect: no realist wishes to argue that *all* normative judgments *anybody* makes are true. People regularly make normative mistakes and depending on the class of people whose normative judgments we are considering (in which culture, which historical era, *etc.*), realists may submit that such mistakes can be plentiful. On the other hand, realists want to resist the contention that, assuming realism, the normative judgments that we typically endorse are generally unreliable. The presumed coincidence that requires explanation, then, is that the contents of our normative judgments *roughly* overlap with the normative truths – at least sufficiently to warrant their general reliability.

Even so, if the set of normative judgements that we *could* have made is incredibly large compared to this subset of true normative judgments that we *actually* endorse, then Enoch's coincidence explanation provides little solace for the realist. This is the view that Street takes: she maintains that realists are still committed to a coincidence that is 'incredible' (Street 2006, p. 125), 'astonishing' (Street 2008, p. 208), 'striking', or 'puzzling' (Street 2016, pp. 305-308). Others have followed Street's terminology and described the presumed coincidence as 'suspicious' (Wielenberg 2010, p. 464), 'extraordinary' (Shafer-Landau 2012, p. 10), 'remarkable' (Cuneo and Shafer-Landau 2014, p. 424), 'miraculous' (Tropman 2014, p. 129), 'startling' (Mogensen 2015, p. 201) and 'implausibly lucky' (FitzPatrick 2015, p. 890). Are these characterizations apt, or misleading? As we saw in **section 2.2**, how much one is struck by a coincidence observation depends *inter alia* on the likelihood of this observation, given the hypothesis one favours. Therefore, to assess whether a coincidence between our normative judgments and the normative truths would indeed be highly unlikely, assuming normative realism, we have to assess to which possibilities realists are committed.

2.5 Questioning Street's Reference Class

Street's reason for thinking that the presumed coincidence between our normative judgments and the mind-independent normative truths is highly unlikely rests on the suggestion that there are numerous ways in which our normative judgments could have gone off-track with respect to these truths, assuming realism. Street argues that there are countless evaluative systems, 'including, for example, systems that place above all else the value of grass-counting, or hand-clasping, or not having one's finger scratched, or counting to the number 78 and back again, and so on' (Street 2016, p. 314). The eccentric agents who endorse such systems are merely creatures of the imagination; we do not encounter them in our actual world. Nonetheless, they are possible creatures that are relevant from a realist perspective, Street argues. Indeed, realists *make* such creatures relevant, since they insist that 'if they existed, they would be mistaken' (*idem*, p. 317). This insistence seems to imply that, assuming realism, all of these rather unorthodox normative judgments are at least *possible*.

What kinds of possibilities, then, are realists committed to? Street (*idem*) seems to think that they are theoretically committed to the claim that all *imaginable* normative judgments are possible normative judgments. It follows that assuming realism, *all imaginable evaluative outlooks* are relevant when assessing the likelihood of the hypothesis that our actual normative judgments happened to coincide (at least roughly) with the mind-independent normative truths.

In other articles Street (2006, 2008a, 2011) advances slightly different proposals, arguing that the relevant ‘possibility space’ consists of the full range of judgments that are *logically* or *conceptually* possible. For instance, she claims that

as a conceptual matter, the independent normative truths could be *anything*. (...) [W]hat are the odds that our values will have hit, as a matter of sheer coincidence, on those things which are independently really worth pursuing? That the odds seem low is an understatement. (Street 2011, p. 114)

Naturally, if this possibility space is infinitely large, then the odds of a coincidence will be infinitely low. And since Street contends that assuming realism, this possibility space is indeed vast, the occurrence of a coincidence is highly unlikely. In order to counter Street’s proposal, realists have to challenge her contention that all of the possibilities she imagines are relevant from a realist point of view. If the relevant possibility space turns out to be much smaller, then the odds will come out more favourably for the realist.

We can reframe the task of establishing what is the relevant possibility space as the task of establishing the proper reference class for assessing the likelihood that a coincidence has occurred. Such a task is notoriously contentious; delineating the relevant possibility space leads to the infamous ‘reference class problem’. As such, we should not be surprised if Street’s proposals are somewhat controversial. Nonetheless, there may be good grounds for rejecting reference classes that are clearly too broad or too small. Consider Street’s proposals: are all imaginable, conceptually or logically possible value systems relevant, assuming realism?

2.5.1 Street’s Reference Class without Darwinian Dilemma

To understand Street’s suggested reference class, it is relevant to point out that Street originally proposed it in the context of her Darwinian Dilemma (2006). This dilemma, which I discuss in further detail in **chapter 3**, departs from the assumption that the contents of our evaluative endorsements are saturated with evolutionary influence, and asks realists whether or not there is a relation between this evolutionary influence and the contents of the mind-independent evaluative truths that realists posit. For realists who go the route of *denying* that there is such a relation, a good case can be made in support of the reference class that Street attributes to them. After all, realists who grant that evolutionary influences *do not* constrain the relevant possibility space of candidate mind-independent evaluative truths, do in fact seem to be committed to the view that these truths might have been *anything*, or at least that the relevant possibility space is extremely broad.

However, not all realists pursue this strategy in response to Street’s dilemma. Some realists assert that there *is* a relation between the evolutionary influences on our evaluative endorsements and the contents of mind-independent evaluative truths (e.g. Enoch 2010; Lott 2018); other realists deny the evolutionary assumption that underlies Street’s dilemma to begin with (e.g. Shafer-Landau 2012; FitzPatrick 2015). In the context of Street’s (2016) practical/theoretical puzzle, then, it cannot be taken for granted that realists are theoretically committed to the claim that the evaluative truths might have been *anything*. But counterproposals to limit Street’s proposed reference class require theoretical justification as well. To see this, consider briefly three of the counterproposals that realists have advanced.

2.5.2 FitzPatrick's Counterproposal

According to William FitzPatrick, many of the possible evaluative judgments (e.g. 'clasping one's hands is more valuable than saving the life of a drowning child') contained in Street's reference class do not only strike us as farfetched but also as unintelligible. He argues that the normative possibilities Street presents are not 'viable candidates for a true system of values because they are divorced from any background framework within which talk of value is intelligible' (FitzPatrick 2014, p. 253).²⁵ It follows that the reference class proposed by Street is uncharitable to realists: she ascribes possibilities to them which realists themselves do not envision.

Importantly, for this criticism to be effective, it does not suffice merely to *assert* that the possibilities Street envisions aren't intelligible, assuming realism. Instead, the task for realists is to argue that it *follows from their theory* that they aren't. Hence, FitzPatrick should come up with a theory-driven proposal of why the set of intelligible normative judgments is limited, assuming realism. Moreover, he should be able to explain why this set roughly aligns with the set of normative judgments that causal forces have led us to make. In other words, he has to provide the outlines of an alternative explanation for the coincidence.

2.5.3 Cuneo and Shafer-Landau's Counterproposal

Alternatively, non-naturalists might argue that the space of conceptual possibilities is much more restricted than Street envisions. Indeed, some non-naturalists maintain that some normative truths – viz. moral truths – are *conceptually necessary*. For instance, Terence Cuneo and Russ Shafer-Landau (2014) argue that our basic moral truths, or what they call the 'moral fixed-points', are conceptual truths 'for beings like us in worlds like our own' (idem, *passim*). These truths are necessary; they could not have been any different. Hence, if we restrict the realist's preferred reference class to conceptual possibilities, it follows trivially that our moral judgments, if conceptually coherent, coincide with moral truths.

As it stands, it is questionable whether this line of reply adequately answers Street's explanatory challenge. After all, insisting that moral truths are conceptually necessary for conceptually coherent moral agents does little to explain why we are conceptually coherent moral agents. Such an explanation may be forthcoming, however: conceptual realists may argue, for instance, that these mind-independent conceptual truths causally influence the normative judgments of moral agents (cf. Shafer-Landau 2012). This would solve Street's practical/theoretical puzzle, but note that these non-naturalists are thereby committed to the controversial view that moral properties have causal powers – a commitment that many non-naturalists regard as problematic, and that is arguably untenable on scientific grounds (see **chapter 3**).

2.5.4 Lott's Counterproposal

²⁵ Simon Blackburn (Unpublished Manuscript) advances a similar consideration in defending quasi-realism, concluding, with Aristotle, that 'the right method for epistemology is not to wipe the slate clean, and then ask what we know in abstraction from anything that could count as a way of knowing. It is to trust the *endoxa* sufficiently to gain a picture of who and what we are that we have managed to achieve beliefs that have every chance of being true.'

Another theory-driven proposal for narrowing Street's reference class is to argue that normative truths are species-dependent or anthropocentric truths. Micah Lott (2018) advances a response along these lines, defending a version of Aristotelian naturalism. In thinking about possible normative truths, Lott argues, we should only consider those possibilities which we take seriously in deliberative contexts, when thinking about what to do and how to live. Practical reasons serve to make our actions intelligible, and there are limits to what might intelligibly count as worthwhile pursuits for creatures like us (Lott 2018, pp. 92-94). On this version of naturalist realism, Street's reference class is certainly inapt; normative systems that place most value on grass-counting, hand-clasping and the like are not genuine candidate systems of human values. Hence, by tying normative truths to our species-typical characteristics, naturalist realists can solve Street's puzzle.

2.5.5 Conclusion: Are Coincidence Explanations Successful?

I have briefly discussed three proposals for limiting the reference class of possible normative judgments, assuming realism. In the process of outlining these proposals, realists have to explicate their own theoretical commitments and provide some alternative explanation for the coincidence. Rather than a pure coincidence explanation, then, these realists explain the coincidence by using a mixed strategy – just like our earlier argument in support of the null hypothesis for the papal thunderstorm.

In the light of these alternative explanations, Street's proposed reference class may come out as uncharitable: there is room for normative realists, on the basis of substantive theoretical considerations, to push back against Street's claim that a coincidence would be astonishing, assuming realism. However, we should keep in mind that each of these alternative explanations rests on an hypothesis about how causal forces enabled us to track normative truths – and the tenability of these hypotheses may be criticized (**chapter 3**). Moreover, as I will now argue in closing this chapter, not all alternative explanations succeed in sufficiently downsizing the reference class to certify the general reliability of our normative judgments.

2.6 The Tenability of Alternative Explanations

The general strategy of providing an alternative explanation for a coincidence between A and B is to argue that they are related, either because A causes B, or because B causes A, or because they share a common cause. The same three general strategies are available when explaining the presumed coincidence between our normative beliefs and the normative facts, albeit with a slight modification: the explanation need not be given in causal terms. Street is aware of this:

(...) although I claim that *some* kind of explanation of the coincidence is demanded, I am presupposing nothing about what *kind* of explanation is needed – whether it be causal, normative, philosophical, or of any other particular kind. (Street 2016, p. 305)

More specifically, since the *explanandum* of the practical/theoretical puzzle is to account for the presumed relation between the causes of our normative judgments and the grounds of normative truths, the *explanans* will have to be some sort of '*causal-cum-grounding* account': an account that reconciles the grounds of normative truths with a causal explanation of the

contents of our normative judgments. The three general strategies for providing such an account are the following:

- 1) to argue that all normative facts are at least partly grounded in our attitudes
- 2) to argue that normative facts have causal powers
- 3) to argue that there is a third factor which indirectly shapes our normative judgments and grounds their truth.

The first solution to the puzzle, which Street (2016, section 6) herself endorses, is to argue that ‘normative truth just is a (rather complex) function of the normative judgments that causal forces led us to make’ (Street 2016, p. 307). More specifically, our own attitudes – which have been shaped by natural selection, our cultural history, the social environment and various other causal forces – are ultimately the truth-makers of normative judgments (see **chapter 5** of this dissertation for further elaboration). Therefore, it is only to be expected that the set of judgments that causal forces have led us to make roughly coincides with the set of true normative judgments. Hence, what explains the purported coincidence between our evolutionarily shaped normative judgments and the normative truths is that the former ground the latter.

The second alternative explanation lets the arrow of explanation run the other way around: the normative truths cause our normative judgments. Many naturalists are sympathetic to this explanation, and perhaps some non-naturalists are too. According to these realists, our normative judgments can be causally influenced by mind-independent normative truths – hence, they give up the assumption that normative properties are causally inert. Street (2016, section 7) regards this alternative as implausible. Indeed, she argues against this explanation in her *Darwinian Dilemma* (2006), where she suggests that causal versions of realism are scientifically untenable, since they are inferior to explanations of our evaluative endorsements in terms of the causal influence of natural selection. However, we should note that Street’s (2016) practical/theoretical puzzle contains no independent argument to this effect. Hence, unless Street’s puzzle is supplemented with her original Darwinian Dilemma (and this dilemma succeeds in challenging realists), the thesis that the normative truths cause our normative judgments can be offered as a legitimate solution to the puzzle.

The third type of alternative explanation is to argue that instead of a direct dependency relation between our normative judgments and the normative facts, these are *indirectly* connected. Indirect *causal-cum-grounding* explanations are often labelled ‘third-factor explanations’, as they point at some ‘third factor’ which has shaped our normative judgments in such a way that they roughly align with the normative truths.²⁶ For instance, Enoch (2010, p. 430) conjectures that our normative beliefs are indirectly caused by facts about what promotes survival and that ‘survival is at least somewhat good’ – i.e. on average, what promotes survival correlates with what is normatively good. This third factor – survival – establishes a connection between our normative beliefs and the normative facts, which goes at least some way towards explaining the coincidence.

²⁶ These accounts have mostly been advanced in response to Street’s *Darwinian Dilemma* (2006).

2.6.1 Are Third-Factor Accounts Trivially Question-Begging?

Enoch's third-factor account explicitly relies on a substantive normative assumption, and so do other third-factor accounts. They assume, for instance, that it is morally right to conform to one's society's authoritative moral codes (Copp 2008), that well-being is morally good (Brosnan 2011) or that pleasure is usually good and pain is usually bad (Skarsaune 2011). This reliance on normative assumptions may seem unwarranted, given that Street has cast doubt on the reliability of our moral intuitions, assuming realism. Indeed, Street (2016, p. 319) calls explanations along these lines 'trivially question-begging'.

But others have defended the move of third-factor theorists. As Berker (2014) and Vavova (2014) point out, it is impossible to explain how our normative judgments could ever track normative truths without assuming that some of them are indeed true (see also Lutz forthcoming; Tersman 2017). By way of comparison, consider once more our perceptual beliefs. We assume that these beliefs are generally reliable, and there is an evolutionary explanation for thinking so: the objects in our manifest surroundings have causal powers, and it would be detrimental to our reproductive success not to discern these objects (cf. Street 2016). But in appealing to evolution, we are appealing to the truth of empirical claims whose justification ultimately depends on the reliability of our perceptual beliefs – hence, such an appeal is circular as well. If we could not rely on these beliefs in establishing their reliability, we would be saddled with global scepticism. This threat of global scepticism, in turn, may be real, but it is not a threat that is particular to normative realism.

Street (2016, p. 321) anticipates this line of response and criticizes it: according to her, the realist's reliance on normative truths in presenting a third-factor account is 'trivially question-begging', whereas the reliance on the reliability of our sense data in the manifest surroundings case is only 'ultimately question-begging' and therefore more acceptable. Even if appeals to the reliability of our perceptual beliefs are *ultimately* circular, they are at least supported by a convincing epistemology, which provides good internal reason to think that our perceptual beliefs are generally reliable. But according to Street, non-naturalist realists like Enoch lack a clearly formulated epistemology – i.e. a positive account that specifies how we can track the contents of normative truths, as she highlights towards the end of her exposition of the practical/theoretical puzzle:

I'm not assuming that there must be any *particular type* of epistemology for the normative domain, much less that it be of the same type that's appropriate for objects with causal powers; I'm merely assuming that there must be *an* epistemology, and that's what the non-naturalist normative realist has so far utterly failed to provide. (idem, pp. 323-324)

Let's assume, for the moment, that Street is correct and non-naturalist realists have failed to provide a positive normative epistemology. In the absence of such an epistemology, a third-factor response to the challenge is indeed problematic: Enoch's (2010, p. 430) assumption that 'survival is at least somewhat good' seems no more than an auxiliary assumption devised *ad hoc* to account for the 'coincidence conjecture' – not unlike an *ad hoc* explanation for why God directed His fury at a basilica, rather than at the pope. Indeed, if realists lack a positive epistemology – a theory that specifies how we gain knowledge of normative truths – then *any*

explanation of the presumed coincidence will appear to be remarkably *ad hoc*. After all, if there is no positive epistemic account to rely on, then realists can only *posit* what is necessary to provide a quick fix to the challenge. Moreover, if realists lack an epistemology altogether, then that fact itself renders realism highly problematic – sufficiently problematic, perhaps, to make the practical/theoretical puzzle a redundant challenge.

That said, it is unlikely that realists will grant Street that they don't have any positive epistemology (see, for instance, Dworkin's 2011 (fn. 9, p. 447) response to Street), and this certainly cannot be assumed without further argument. Even if the details of a realist epistemology turn out to be vague and ultimately unsatisfying, their failure cannot simply be assumed but should be argued for by engaging with the details of their actual contents. Since Street (2016) does not engage with these contents, her contention that a realist epistemology is 'trivially question-begging' is uncharitable.

2.6.2 Does Enoch's Explanation Establish that Our Normative Judgments Are Generally Reliable?

As previously stated, the most attractive strategy for solving the practical/theoretical puzzle is to combine a coincidence explanation with an alternative explanation. The former serves to moderate the epistemic unease triggered by the coincidence; the latter serves to explain what is left of the coincidence after such moderation. Enoch (2010, pp. 427-435) explicitly pursues this strategy. First he argues that the coincidence is not all that striking; next he explains what remains of the coincidence by arguing that our normative beliefs are indirectly caused by facts about what promotes survival, and that 'survival is at least somewhat good.' This suffices, Enoch thinks, to vindicate the claim that assuming realism 'we are at least somewhat more likely to get things right rather than wrong with regard to normative issues' (idem, p. 427) – i.e. to think that our normative judgments are generally reliable.

However, it is questionable whether Enoch's third factor can truly secure a somewhat positive correlation between our normative judgments and the normative truths. After all, the reference class of survival-promoting judgments is still vast; presumably there is only a small set of normative judgments within this enormous reference class that tracks normative truths. Enoch's reference class allows for survival-promoting strategies such as killing the offspring of rivals, abusing others for one's own benefit and cheating to the detriment of the group. Clearly, the presumed set of mind-independent normative truths is much more constrained than the set of normative imperatives that these – and many other – survival-enhancing behaviours would allow for. As a result, there is still a coincidence that needs an explanation: of all the survival-enhancing evaluative truths, why do the contents of our judgments largely coincide with the 'right' ones – the mind-independent normative truths?

Prima facie, a line of reply that may be open to Enoch is to argue that apart from the goodness of survival, there are several other factors which have simultaneously caused and grounded our judgments. However, such a reply may appear to be very much *ad hoc* and to provide just the kind of vicious explanation outlined above: unless these further third factors are posited on the basis of theory-driven considerations, their assumed presence is trivially question-begging. Indeed, Enoch (2010) does not indicate that his third-factor account is driven by theoretical considerations, other than answering Street's challenge. It seems, then, that there

are serious worries about the adequacy of Enoch's third-factor account as a solution to the practical/theoretical puzzle.

2.7 Conclusion

How should we solve the practical/theoretical puzzle? I have outlined two general strategies for solving it – providing a coincidence explanation or providing an alternative explanation – which should best be pursued in tandem. With regard to the former, the task for realists is to come up with a theoretically motivated proposal for diminishing Street's proposed reference class of normative judgments that we *could* endorse, assuming realism. With regard to the latter, the task for realists is to argue that their alternative explanation is more plausible, all things considered, than the alternative explanation Street favours.

The prospect of normative realists completing the former task may be brighter than Street (2016) suggests. Realists can come up with various theoretically driven proposals which suggest that, assuming realism, making true normative judgments is not like winning a lottery with extremely low prior odds. But not all realist proposals succeed in establishing the general reliability of our normative judgments. Specifically, I have argued that Enoch's (2010) third-factor account does not sufficiently constrain the relevant reference class: his account still commits him to a leap of faith in assuming that we have been among the normatively lucky ones.

To complete the latter task, and explain the coincidence between our normative judgments and the mind-independent normative truths, realists may want to argue that moral properties have causal powers (**section 2.5**). Realists who pursue this strategy may be able to avoid Street's coincidence challenge, but in doing so they are likely to run into a different challenge – namely, that their view is incompatible with our best science, as Street (2006) argues in her Darwinian Dilemma. In order to pose a general challenge for normative realism, then, debunkers are best advised to follow Street in supplementing her philosophical puzzle with other, scientifically driven arguments. As I will demonstrate in the ensuing two chapters of this dissertation, once supplemented with scientific considerations the practical/theoretical puzzle *does* pose a general challenge for the tenability of moral realism.

References

- Berker, S. 2014. Does Evolutionary Psychology Show that Normativity is Mind-Dependent? In: J. D'Arms and D. Jacobson, eds. *Moral Psychology and Human Agency: Essays on the Science of Ethics*. New York: Oxford University Press, 215-252.
- Berker, S. 2018. The Unity of Grounding. *Mind* 127: 729-777.
- Berker, S. Forthcoming. The Explanatory Character of Moral Principles. *Noûs*.
- Blackburn, S. Unpublished Manuscript. Sharon Street on the Independent Normative Truth *as such*. www2.phil.cam.ac.uk/~swb24/PAPERS/Meanstreet.htm
- Brosnan, K. 2011. Do the Evolutionary Origins of Our Moral Beliefs Undermine Moral Knowledge? *Biology & Philosophy* 26: 51-64.
- Copp, D. 2008. Darwinian Skepticism about Moral Realism. *Philosophical Issues* 18: 186-206.

- Cuneo, T. and Shafer-Landau, R. 2014. The Moral Fixed Points. *New Directions for Moral Nonnaturalism. Philosophical Studies 171*: 399-443.
- Diaconis, P. and Mosteller, F. 1989. Methods for Studying Coincidences. *Journal of the American Statistical Association 84*: 853-861.
- Doyle, A.C. 2011 [1913]. *The Adventure of the Dying Detective*. Portland (OR): The Floating Press.
- Dworkin, R. 1996. Objectivity and Truth: You'd Better Believe It. *Philosophy & Public Affairs 25*: 87-139.
- Dworkin, R. 2011. *Justice for Hedgehogs*. Cambridge: Harvard University Press.
- Enoch, D. 2010. The Epistemological Challenge to Metanormative Realism: How Best to Understand It, and How to Cope with It. *Philosophical Studies 148*: 413-438.
- Falk, R. 1989. Judgment of Coincidences: Mine versus Yours. *The American Journal of Psychology 102*: 477-493.
- Fine, K. 2012. Guide to Ground. In: F. Correia and B. Schnieder eds. *Metaphysical Grounding: Understanding the Structure of Reality*. Cambridge: Cambridge University Press, 37-80.
- FitzPatrick, W. 2014. Why There Is no Darwinian Dilemma for Ethical Naturalism. In: M. Bergmann and P. Kain, eds. *Challenges to Moral and Religious Belief: Disagreement and Evolution*. Oxford: Oxford University Press, 237-255.
- FitzPatrick, W. 2015. Debunking Evolutionary Debunking of Ethical Realism. *Philosophical Studies 172*: 883-904.
- Griffiths, T.L and Tenenbaum, J.B. 2007. From Mere Coincidences to Meaningful Discoveries. *Cognition 103*: 180-226.
- Hopster, J.K.G. 2017. Two Accounts of Moral Objectivity: From Attitude-Independence to Standpoint-Invariance. *Ethical Theory & Moral Practice 20*: 763-780.
- Lott, M. 2018. Must Realists be Skeptics? An Aristotelian Reply to a Darwinian Dilemma. *Philosophical Studies 175*: 71-96.
- Lutz, M. 2017. Forthcoming. What Makes Evolution a Defeater? *Erkenntnis*.
- Mogensen, A. 2015. Evolutionary Debunking Arguments and the Proximate/Ulimate Distinction. *Analysis 75*: 196-203.
- Mogensen, A. Forthcoming. Ethics, Evolution, and the Coincidence Problem: A Skeptical Appraisal. <http://andreamogensen.com/wp-content/uploads/2014/10/Coincidence-Problem-v5.pdf>
- Owens, D.J. 1992. *Causes and Coincidences*. Cambridge: Cambridge University Press.
- Shafer-Landau, R. 2012. Evolutionary Debunking, Moral Realism and Moral Knowledge. *Journal of Ethics and Social Philosophy, www.jesp.org, 7(1)*: 1-37.
- Skarsaune, K.O. 2011. Darwin and Moral Realism: Survival of the Iffiest. *Philosophical Studies 152*: 229-243.
- Street, S. 2006. A Darwinian Dilemma for Realist Theories of Value. *Philosophical Studies 127*: 109-166.
- Street, S. 2008. Reply to Copp: Naturalism, Normativity, and the Varieties of Realism worth worrying about. *Philosophical Issues 18*: 207-228.

Street, S. 2011. Mind-Independence without the Mystery: Why Quasi-Realists Can't Have It Both Ways. In: R. Shafer-Landau, ed. *Oxford Studies in Metaethics*, vol. 6. Oxford: Clarendon Press, 1-32.

Street, S. 2016. Objectivity and Truth: You'd Better Rethink It. In: R. Shafer-Landau ed. *Oxford Studies in Metaethics*, vol. 11. Oxford: Clarendon Press, 293-334.

Sober, E. 2012. Coincidences and How to Reason about Them. In: H.W. de Regt et al., eds. *EPSA Philosophy of Science, Amsterdam 2009*. Springer, 355-374.

Tersman, F. 2017. Moral Skepticism and the Benacerraf Challenge. In: D. Machuca, ed. *Moral Skepticism: New Essays*. New York: Routledge, 183-198.

Tropman, E. 2014. Evolutionary Debunking Arguments: Moral Realism, Constructivism, and Explaining Moral Knowledge. *Philosophical Explorations* 17: 126-140.

Vavova, K. 2014. Debunking Evolutionary Debunking. In: R. Shafer-Landau, ed. *Oxford Studies in Metatethics*, vol. 9: 76-101.

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

3

Chapter 3

Evolutionary Arguments against Moral Realism: Why the Empirical Details Matter (and Which Ones Do)

Forthcoming in *Biology & Philosophy*.

Keywords

Darwinian Dilemma; reliability; evolutionary contingency; historical debunking argument; best explanation

Abstract

The aim of this chapter is to identify the strongest evolutionary debunking argument (EDA) against moral realism and to assess on which empirical assumptions it relies. In the recent metaethical literature, several authors have de-emphasized the evolutionary component of EDAs against moral realism: presumably, the success or failure of these arguments is largely orthogonal to empirical issues. I argue that this claim is mistaken. First, I point out that Sharon Street's and Michael Ruse's EDAs both involve substantive claims about the evolution of our moral judgments. Next, I argue that combining their respective evolutionary claims can help debunkers to make the best empirical case against moral realism. Some realists have argued that the very attempt to explain the contents of our endorsed moral judgments in evolutionary terms is misguided, and have sought to escape EDAs by denying their evolutionary premise. But realists who pursue this reply can still be challenged on empirical grounds: debunkers may argue that the best, scientifically informed historical explanations of our moral endorsements do not involve an appeal to mind-independent truths. I conclude, therefore, that the empirical considerations relevant for the strongest empirically driven argument against moral realism go beyond the strictly evolutionary realm; debunkers are best advised to draw upon other sources of genealogical knowledge as well.

'[T]he burden is on the evolutionary debunker to substantiate every one of the premises of the Darwinian Argument.'

– Russ Shafer-Landau (2012)

3.1 Introduction

In the recent metaethical debate about evolutionary debunking arguments (EDAs), there has been a tendency to de-emphasize the relevance of genuinely evolutionary considerations for debunkers' purposes. For instance, Raymond Das describes it as

a deep irony that the import of the specifically *evolutionary* aspect of EDAs is itself fairly minimal. Apart from being (for the most part) a 'how possibly' story about the origins of our moral judgments (...), the evolutionary component of such arguments is, as Joyce (2016, p. 125) has noted, 'strictly, dispensable.' Any equally plausible causal explanation of our moral judgments that does not presuppose the truth of such judgments would serve the evolutionary debunker's purposes just as well – or as poorly. (Das 2016, p. 419)

Similarly, referring to the most renowned EDA against moral realism – Sharon Street's (2006) Darwinian Dilemma – David Enoch submits that

there is nothing essentially Darwinian about the Darwinian Dilemma. Replace any other (non-tracking) causal explanation of why we make the normative judgments that we do in fact make, and the realist will again find herself up against the problem of explaining strong correlations analogous to the ones Street draws attention to. (Enoch 2010, p. 426)

Das, Joyce and Enoch have not been the first ones to question the relevance of the strictly evolutionary aspect of EDAs. In fact, in the concluding section of her article, Street herself makes it explicit that

I have focused on the case of Darwinian influences on our evaluative judgments because I think it raises the problem for realism in a particularly acute form. In principle, however, an analogous dilemma could be constructed using any kind of causal influence on the content of our evaluative judgments. (...) At the end of the day, then, the dilemma at hand is not distinctly Darwinian. (Street 2006, p. 155)

If not as an evolutionary challenge, how should Street's EDA be understood? Several commentators (Enoch 2010; Clarke-Doane 2012; Crow 2016; Klenk 2017; Tersman 2017; Schechter 2018) have interpreted her argument as the moral analogue of the Benacerraf-Field challenge for mathematical Platonism. Roughly, in its generalized form, this is the challenge of explaining how we can reliably track facts or truths about some specified domain, if we assume that these facts or truths are stance-independent and causally inert. Reframed in the

context of moral realism, the challenge is to explain how we have managed to arrive at moral beliefs that are mostly true (hence reliable), assuming – as many moral realists do – that the truth-makers of these beliefs are stance-independent and causally inert facts. Evolutionary considerations are inessential to this challenge; any causal explanation for why we endorse the moral judgments we do suffices to generate it.

I submit that these commentators are right to point out that the Benacerraf-Field challenge is one of the most potent philosophical challenges for moral realism and that part of Street's work – especially her (2016) practical/theoretical puzzle – relies on a challenge along much of the same lines. But it would be a mistake to identify the Benacerraf-Field challenge with the challenge embodied by the Darwinian Dilemma. One obvious difference between them is that evolutionary considerations are irrelevant to the former, whereas they are crucial to the latter, as I will demonstrate in this chapter.

These commentators are also right to point out that many EDAs in metaethics, including Street's Darwinian Dilemma, partly rely on metaphysical and epistemological assumptions. But even if evolutionary considerations – or empirical considerations more generally – don't do *all* the work in challenges against realism, it would be a mistake to think that they barely do *any* work in them. I will argue that with regard to Street's Darwinian Dilemma, empirical considerations are central to evaluating its ultimate success. Moreover, I will argue that in order to make the best empirical case against realism, debunkers following Street's tracks should move beyond the strictly evolutionary realm and invoke historical considerations as well, in order to fashion an adequate response to realists who deny the dilemma's evolutionary premise (e.g. Shafer-Landau 2012; FitzPatrick 2015; Huemer 2016).

The chapter proceeds as follows. In **section 3.2** I recapitulate Street's evolutionary premise and the Darwinian Dilemma it gives rise to. In **section 3.3** I compare and contrast Street's argument with Michael Ruse's claim about the evolutionary contingency of our moral attitudes and show how Ruse's contingency claim gives rise to a slightly different EDA against realism. In **section 3.4** I argue that taken individually, Ruse's and Street's respective EDAs each face a specific objection, which they can salvage by drawing upon each other's arguments. In **section 3.5** I outline a strategy for realists to block these EDAs: they can deny that evolutionary forces have deeply influenced the contents of our endorsed moral judgments. I submit that realists may be right to insist that the contents of some – perhaps many – endorsed judgments cannot be adequately explained in evolutionary terms. However, in **section 3.6** I proceed to argue that this does not safeguard realism from empirically driven debunking arguments: debunkers may still argue that the *best historical explanation* of our moral endorsements involves no appeal to mind-independent moral truths.

3.2 Street's 'Darwinian Dilemma'

The target of Street's Darwinian Dilemma is 'evaluative realism', the defining claim of which is that evaluative facts or truths are stance-independent: their truth is not determined by endorsements from any actual or hypothetical perspective.²⁷ In what follows I shall discuss a

²⁷ Street's most obvious target among realist views is evaluative non-naturalism – roughly, the view that evaluative facts or truths are not reducible to natural facts and do not figure in causal explanations. Street (2006, section 7) also claims to target some naturalist views, but grants that other naturalists

version of her argument with a more constrained scope that is specifically geared towards the *moral* domain.²⁸

Street's dilemma is generated by the empirical hypothesis that evolutionary forces have been a major causal influence shaping our evaluative attitudes (which includes our *moral* attitudes). More specifically, Street (2006, p. 119) claims that our 'basic evaluative tendencies' have been shaped by a process of natural selection. These tendencies are our intuitive inclinations – shared with many other species – to perceive certain behaviours as 'called for' or 'demanded by' the circumstances; for example, a natural urge to protect one's offspring, or a tendency to seek pleasure and to avoid pain. This is Street's so-called adaptive-link account of the origins of our evaluative tendencies: these tendencies originated as adaptive responses to environmental circumstances. While the contents of our evaluative tendencies do not strictly determine the contents of our 'full-fledged evaluative judgments', Street hypothesizes that they are related nonetheless: '[H]ad the general content of our basic evaluative tendencies been very different, then the general content of our full-fledged evaluative judgments would also have been very different, and in loosely corresponding ways' (Street 2016, p. 120). In a first approximation, then, Street's empirical premise is that natural selection has had a substantial – albeit indirect – influence on our endorsed evaluative judgments.

Why would this empirical premise, if true, be troublesome for moral realists? Realists typically assume that the set of moral judgments that people endorse largely coincides with the set of stance-independent moral truths. But Street argues that this assumption cannot be maintained in the light of the evolutionary influences that permeate the contents of our moral judgments. This is what the first horn of her Darwinian Dilemma is meant to establish: assuming that there are stance-independent moral truths, we probably fail to track them.

Why so? Street confronts realists with the dilemma of specifying whether there is a relation between stance-independent moral truths and the evolutionary influences on our moral judgments:

- either the realist holds that there is *no* relation between the evolutionary forces that have influenced the contents of our moral judgments and the independent moral facts or truths
- or the realist holds that there *is* such a relation.

Both horns of the dilemma leave realists in an unappealing position, Street argues. On the first horn of the dilemma, realists must either explain how the contents of our moral judgments happened to coincide with the contents of the stance-independent truths or regard the

escape her challenge. On her taxonomy, the latter views do not count as genuinely realist, since they do not subscribe to the view that moral facts or truths are stance-independent. The term 'stance-independence' comes from Shafer-Landau (2003, p. 15), who characterizes it as follows: 'The moral standards that fix the moral facts are not made true by virtue of their ratification from within any given actual or hypothetical perspective.'

²⁸ I submit that the Darwinian Dilemma, understood as an argument that relies on truly evolutionary considerations, becomes more interesting when it is restricted to the moral domain. This is because much of the evidence in support of the hypothesis that many of our moral judgments can (at least indirectly) be evolutionarily explained is not – or not obviously – available in explaining many of our non-moral evaluative judgments.

evolutionary influences on these judgments as ‘distorting’ – i.e. grant that these influences have resulted in judgments that are probably *false*, assuming realism. Simply *positing* that they coincide is question-begging, Street maintains: the realist needs to *explain* this coincidence. This ‘demand for an explanation’ is better understood as a demand for theoretical justification: realists need to show how *it follows from their theory* that our evolutionarily influenced moral judgments probably coincide with the stance-independent moral truths. But if realists deny that there is a relation between the evolutionary forces that influence our judgments and these independent truths, then it seems very difficult to provide such justification.

Realists who accept the second horn of the dilemma and grant that there *is* a relation between the evolutionary forces that have shaped our judgments and mind-independent moral truths are committed to a truth-tracking explanation: presumably, the correlation between our moral judgments and these independent moral truths is best explained by the hypothesis that it tended to be reproductively advantageous for our ancestors to make true moral judgments. But Street argues that this truth-tracking explanation is scientifically inferior to her adaptive-link account, according to which moral truths are ultimately a construction of our evolved attitudes. Our moral judgments don’t tend to be fitness-enhancing because we have tracked mind-independent truths that correlate with reproductive success. Instead, we tend to regard fitness-enhancing judgments as true, because this tendency increased the reproductive success of our ancestors. Hence, as an inference to the best explanation, moral realism is untenable.

3.2.1 The Reliability Challenge

The first horn of the Darwinian Dilemma can be reframed as a ‘reliability challenge’ for moral realism that is meant to establish the following conclusion: assuming realism, we have no guarantee that our moral judgments are generally reliable. To elucidate this challenge, we may frame it in terms of sets. Realists assume that the contents of the moral judgments that we endorse roughly coincide with the contents of stance-independent moral truths. But Street argues that given the vast number of alternative judgments that we *could* have endorsed, such a coincidence would be very implausible. Consider the various moral judgments that we do not endorse – from the judgment that ‘infanticide is laudable’ to the judgment that ‘plants are more valuable than human beings’ to the judgment that ‘the fact that something is purple is a reason to scream at it’ (Street 2006, p. 133). If the realist denies that there is a relation between the influence of natural selection on our judgments and the independent moral truths, then given the vast range of logical possibilities, it would be very remarkable if natural selection accidentally shaped our judgments in concordance with these truths. In all likelihood it has not, Street concludes.

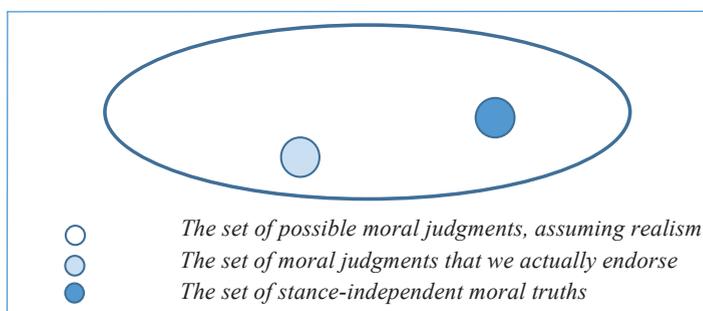


Figure 3.1: Street's Reliability Challenge²⁹

How might realists respond to this challenge? Some have done so by challenging the relevance of the possibilities that Street invokes (see **chapter 2** of this dissertation). Since Street is presenting an *internal* challenge to realism – assuming realism, we have no guarantee that our judgments are reliable – the judgments relevant to her challenge are the judgments we could make *assuming realism*. But not all realists grant that all logical possibilities are relevant, assuming realism (e.g. Berker 2014, p. 246; FitzPatrick 2014, p. 253; Wielenberg 2010, pp. 458-459).³⁰ If the relevant possibility space can be strongly diminished, for example by showing that the possibilities Street envisions are unintelligible (FitzPatrick 2014) or conceptually deficient (cf. Cuneo and Shafer-Landau 2014), then Street's claim that our judgments are likely to be off-track loses much of its force. Perhaps the rough alignment of our moral judgments with the independent moral truths will still be *somewhat* of a coincidence, but not nearly as unlikely as Street takes it to be.

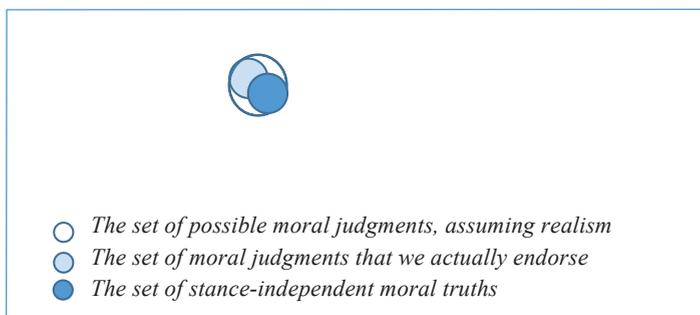


Figure 3.2: Realist Strategy to Restore Reliability

²⁹ The details of this diagram may vary depending on the version of realism under consideration. For instance, some realists might argue that the semi-transparent circle should be bigger than the solid circle: after all, we endorse many evaluative judgments that are false. Other realists might argue that the solid circle should be bigger than the semi-transparent circle: after all, there might well be moral truths which have not (yet) been tracked by our moral judgments. The set of all possible moral judgments (transparent circle) is (nearly) infinite: following Street (2006) it contains all logically possible judgments.

³⁰ In later writings Street speaks of *conceptual* possibilities (Street 2008) or conjures all *conceivable* possibilities (Street 2016). *Mutatis mutandis*, the same considerations apply to them.

The *desideratum* for realists who take on the first horn of Street's dilemma, then, is to argue that a coincidence is sufficiently likely to justify the assumption that our moral judgments are generally reliable. Realists can do so by arguing that the set of possible moral judgments we *could* make, assuming realism (the transparent circle), is much smaller than Street suggests. However, it won't help the realist to merely *posit* that the set of possible judgments we could endorse is limited and clearly correlated with the set of judgments that are evolutionarily beneficial. Street's explanatory demand is to justify *why* this is the case – i.e. to offer a theory-driven consideration, assuming realism, for limiting the relevant possibility space. To offer such a consideration *after* having granted that there is no relation between the evolutionary influences on our moral endorsements and the contents of mind-independent moral truths seems very difficult. In effect, by granting that evolutionary influences do not constrain the possibility space of candidate mind-independent moral truths, realists *do* commit themselves to the thesis that these truths might have been *anything*. The more promising option for realists, then, is to argue that there *is* a relation between the mind-independent moral truths and our evolutionarily shaped judgments. But this brings us back to the second horn of Street's dilemma: realists can only frame this as a truth-tracking relation, which should be discarded on scientific grounds, or so Street argues.

3.3 Ruse's Contingency Challenge

In Ruse's (1995) work we find two metaethical challenges for moral realism. In this section I briefly discuss both of them and zoom in on the latter challenge, which relies on an assumption about the evolutionary contingency of moral norms.

According to Ruse, morality can be regarded as a biological mechanism for fostering cooperation. Our moral sense is an adaptation, which is all the more effective in fulfilling its evolutionary function because it creates the illusion that morality derives from objective foundations external to ourselves. This is what makes morality work: we obey moral norms precisely because we take them to be objective. But the suggestion that morality has a stance-independent foundation is in fact illusory, as its evolutionary origins make clear.

How does Ruse's evolutionary hypothesis challenge realism? We can distinguish between two strands in Ruse's work. Sometimes, Ruse invokes the metaethical premise that given the evolutionary explanation of our moral beliefs the objective foundation of morality has to be judged redundant:

You would believe what you do about right and wrong, irrespective of whether or not a 'true' right and wrong existed! The Darwinian claims that his/her theory gives an entire analysis of our moral sentiments. Nothing more is needed. Given two worlds, identical except that one has an objective morality and the other does not, the humans therein would think and act in exactly the same ways. (1995, p. 254)

In this passage, what does the work in Ruse's EDA against realism is the assumption that a Darwinian explanation of our beliefs in objective moral truths makes the existence of their truth-makers metaphysically redundant: we no longer need to posit them. If our best evolutionary explanation nowhere supposes that our moral beliefs track objective truths, Ruse suggests, then we should erase these objective truths from our ontology.

Elsewhere Ruse advances a different evolutionary argument against realism. If we assume that ‘objective morality’ corresponds to some reality external to human beings, Ruse argues, then

we might have evolved in such a way as to miss completely its real essence. We might have developed so that we think we should hate our neighbors, when really we should love them. Worse than this even, perhaps we really should be hating our neighbors, even though we think we should love them! (Ruse 1995, p. 242)

Here Ruse relies on the idea that our value-judgments are evolutionary contingent: we might have evolved making very different value-judgments. Given this possibility, how do we know that our actual value-judgments are correct? If the range of possibilities implied by the contingency of our values is sufficiently large, then we should conclude that the reliability of our actual endorsements cannot be guaranteed; they are probably false. Hence, Ruse’s argument can be reframed as a reliability challenge for realism much like that of Street (**figure 1**), but framed in terms of the contingency of our moral judgments.

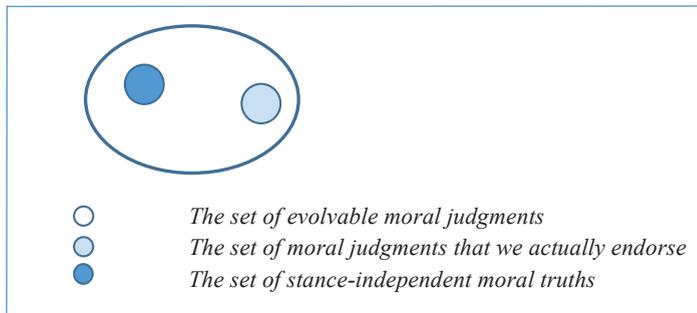


Figure 3.3: Ruse’s Contingency Challenge

Some recent commentators have framed Ruse’s challenge somewhat differently: as targeting the *epistemic safety* of our moral beliefs (e.g. Bogardus 2016, p. 645). Even if we assume that our current moral beliefs reflect the contents of stance-independent truths, the *ease* with which we might have evolved different moral beliefs – beliefs that would be false, assuming realism – seems troubling for realists.³¹ This ease suggests that our moral beliefs lack the property of being epistemically ‘safe’, assuming realism:

Safety: S’s beliefs about a domain D are safe iff there is no nearby world in which S, using similar cognitive means to those in our actual world, arrives at false beliefs about D.

³¹ The notion of ‘ease’ – and the corresponding ‘proximity’ of possible worlds – is difficult to make precise. A potential strategy for doing so in an evolutionary context is to advance a variation on Stephen Jay Gould’s thought experiment *replaying the tape of life*: ‘nearby worlds’ are worlds that might ‘easily evolve’ if we could replay life’s tape.

Safety is generally (though not universally) accepted as a necessary condition for knowledge.³² The underlying idea is that if S's beliefs about a domain are unsafe – they could easily have been false – then their truth is merely a matter of luck. This realization seems to undercut the justification of S's beliefs about D. Or in the context of moral realism: it provides an undercutting defeater for the hypothesis that our moral beliefs coincide with mind-independent moral truths. Justification may be reinstated, but the burden of proof is with realists to show that it can be.

Whether Ruse's EDA is framed as a challenge for the truth of our moral judgments or the safety of our moral beliefs, assuming realism, on both versions the success of his argument hinges on the claim that our moral endorsements are evolutionarily contingent, which is – at least in part – an empirical claim. More precisely, for his EDA to succeed, our moral endorsements should be contingent in a way that is troublesome for realism: the contingency of our endorsements should imply that either they are probably false, or that they might have easily been false, assuming realism.

3.4 Entangled between Street's and Ruse's EDAs

In some respects, Ruse's evolutionary challenge for realism is quite similar to Street's Darwinian Dilemma. Both EDAs challenge, on evolutionary grounds, the assumption that our moral judgments track mind-independent truths, and both of them do so, at least in part, by presenting a reliability challenge for moral realism. But the challenges differ in the details. What motivates Ruse to question the reliability of our moral judgments, assuming realism, is that realism allows for the possibility that we might have evolved in such a way as to completely miss the mind-independent moral truths. This claim presupposes that our moral judgments are evolutionarily contingent; if they are, then either our moral judgments are false or they might easily have been, assuming realism.

By contrast, for the success of Street's argument it is irrelevant whether or not our evolved moral beliefs are contingent in an *evolutionary* sense. Even if they are not – i.e. even if from an evolutionary point of view, our moral beliefs could not have been any different – Street can still raise her challenge for moral realism: to account for the striking correlation (striking, that is, in the light of the vast space of what is logically possible) between the moral judgments that we take to be true and judgments that tend to contribute to our reproductive success. Hence, what motivates Street to question the reliability of our moral judgments, assuming realism, is that realists do not have a proper explanation, informed by their own theory, for the assumed coincidence between our moral judgments and the mind-independent moral truths. What seems to be the only possible realist proposal to explain this coincidence – to advance a truth-tracking account – is untenable on scientific grounds.

While Street's Darwinian Dilemma does not depend on a claim about evolutionary contingency, in this section I will demonstrate that such a claim might nonetheless help her argument. I will argue that some of the most popular responses that realists have offered to Street's dilemma – so-called third-factor accounts – lead to contingency worries like those that Ruse invokes. Moreover, I will argue that one potent reply against Ruse's EDA – an appeal to evolutionary constraints – arguably reinforces the empirical hypothesis underlying Street's

³² Bogardus (2016, p. 647) himself criticizes the safety condition.

Darwinian Dilemma. Hence, by advancing Street's and Ruse's respective EDAs in tandem, debunkers can deflect some common realist replies.

3.4.1 Countering Third-Factor Objections

First, consider some prominent third-factor accounts that realists have advanced in reply to Street's dilemma. Third-factor theorists submit that there is a relation between the evolutionary influences on our moral judgments and the independent moral truths, but argue that this relation is indirect: there is some third-factor involved which indirectly causes our basic evaluations and grounds their truth. In order to come up with a third-factor explanation, realists have to make some modest assumptions about which evaluations are in fact true. For instance, David Enoch (2010, p. 430) bases his third-factor explanation on the assumption that 'survival is at least somewhat good'. Since evolution 'aims' at survival, what evolution aims at is at least somewhat good. Another third-factor account, proposed by Knut Olav Skarsaune (2011, p. 230), starts from the assumption that 'pleasure is usually good and pain is usually bad'. Since natural selection has generally led us to seek pleasure and to avoid pain, it has caused us to value things that tend to be good.

The alleged upshot of third-factor accounts is to explain why it need not be regarded as a striking coincidence that our evolved moral endorsements correlate with the independent moral truths. But it is questionable whether the accounts of Enoch and Skarsaune succeed in showing this. One of the difficulties that their accounts face is that the modest moral assumptions they invoke allow for the evolution of various kinds of moral judgments, many of which realists will presumably regard as false.³³ For instance, there is a wide variety of behaviours that can be survival-enhancing, including cheating, stealing, free-riding and making self-serving moral judgments.³⁴ Likewise, there is a variety of behaviours that can induce pleasure, including making jokes at the expense of helpless others or eating factory-farmed animals. Presumably, realists will regard the moral judgments that accompany such behaviours, or at least many of them, as off-track: they do not track independent truths. However, it is not the third-factor explanation that tells us so; indeed, these are all behaviours that the proposed third-factor explanation might allow for. As a result, third-factor theorists once again face a reliability challenge: of all the moral judgments we could have made in the light of the proposed third-factor explanation, what guarantees that our actual judgments coincide with the mind-independent truths?

This challenge is structurally similar to the first horn of Street's Darwinian Dilemma. The debunker argues that realists need to show that *it follows from their theory* that our evolutionarily influenced moral judgments probably coincide with the stance-independent moral truths, and that third-factor theorists like Enoch and Skarsaune have failed to show this. But in addition to Street's original EDA, the reliability challenge now explicitly relies on considerations about evolutionary contingency. If the realm of moral judgments that we might

³³ Debunkers have criticized third-factor accounts on other grounds as well. Here I restrict my discussion to the question of whether they can restore the reliability of our moral beliefs. Note that my criticism applies specifically to third-factor accounts which rely on highly general moral assumptions.

³⁴ Braddock (2016) lists several examples of such behaviours, which have evolved over the course of human evolution. Presumably, realists maintain that the judgments fuelled by these 'nasty norms' are – and have always been – false.

evolve, assuming realism, is merely constrained by the assumption that survival is somewhat good, or that pleasure is usually good and pain usually bad, then the reliability of our actual judgments cannot be guaranteed. The realist has established a relation between the evolutionary influences on our judgments and the independent truths, but she has not sufficiently constrained the relevant possibility space: replay the tape of life and we might evolve survival-enhancing judgments with very different contents. Enoch's and Skarsaune's accounts still allow for the possibility that we might have evolved to love our neighbours, whereas, really, we should hate them.

3.4.2 Countering Objections regarding Evolutionary Constraints

Now turn to Ruse's contingency challenge. One way in which realists have challenged Ruse's EDA is by questioning the evolutionary contingency thesis on which it relies. Ruse illustrates the supposed contingency of our moral judgments by invoking counterfactuals that involve species with phenotypes very different from our own. For instance, he asks us to imagine that

instead of evolving from savannah-living primates (which we did), we had come from cave dwellers. (...) Or take the termites (to go to an extreme example from a human perspective). They have to eat other's feces, because they lose certain parasites, vital for digestion, when they molt. Had humans come along a similar trail, our highest ethical imperatives would have been very strange indeed. (Ruse 1995, pp. 241-242)

This counterfactual is reminiscent of a passage by Darwin, who speculated in *The Descent of Man* that

if men were reared under precisely the same conditions as hive-bees, there can hardly be a doubt that our unmarried females would, like the worker-bees, think it a sacred duty to kill their brothers, and mothers would strive to kill their fertile daughters; and no one would think of interfering. (Darwin 2013, p. 58)

Such evolutionary counterfactuals may be suggestive, but how plausible are they? Realists might counter that a capacity to make moral judgments is only evolvable for species quite similar to ourselves – e.g. species with a language-infused capacity for making judgments and deliberating about different courses of action, a cognitive capacity for norm-guided behaviour and an emotional sensitivity for the well-being of others. Hence, rather than accepting that our moral judgments are evolutionarily contingent, realists might argue that there are developmental or adaptive constraints which limit the possible contents of such judgments. If this counterargument holds up, then Ruse's contingency argument does not get off the ground: if we could not have easily evolved as a species with a moral capacity while endorsing very different moral norms, then we could not have easily missed morality's 'real essence'. As a result, we would have no reason to question the truth or justification of our moral beliefs, assuming realism.

The biologist Jeffrey Schloss (2014) has recently argued along these lines. Schloss holds that there are strict structural and developmental constraints on the evolution of a moral capacity, which limit the resultant contents of judgments that a species with a moral capacity

might feasibly endorse. These constraints include not only high levels of sociality and intelligence, but also ‘life expectancy, mortality rate, fertility rate, body size, and the relationship between degree of infant dependence, parental care, lifelong pair bonding, group hunting, and even bipedal gait (which modified the pelvis resulting in increased dependence, care, and pair bonds)’, according to Schloss (2014, p. 110). This combination of features does not evolve easily – and neither does a capacity for moral judgment. The fact that on our planet this capacity has evolved only in one species suggests as much.³⁵ If Schloss is correct, then we should expect that the evolution of a moral sense in any species will be closely tied to capacities that are distinctive of human beings. Counterfactuals like those of Ruse, Schloss submits, may be ‘as impossible for biology as a square circle is for geometry’ (idem, p. 109).

Schloss’s argument, if correct, helps realists to block Ruse’s contingency challenge. But it also serves to reinforce the general empirical claim underlying Street’s Darwinian Dilemma – namely, that evolutionary forces have been a major causal influence shaping our moral attitudes. While Street (2006) herself gives substance to this evolutionary claim by advancing an adaptationist explanation of the contents of our moral judgments, her appeal to natural selection is not strictly necessary for the purposes of her EDA: debunkers following Street’s approach merely need to provide evidence that evolutionary forces – however specified – have played a major causal role in shaping our judgments. Since Schloss’s appeal to developmental constraints serves to corroborate this empirical hypothesis, it leads to the Darwinian Dilemma: realists have to explain how the evolutionary forces that have influenced our moral judgments are related with the stance-independent moral truths.

In sum, by borrowing elements from both Street’s and Ruse’s respective EDAs, debunkers can enhance their best argument against moral realism. Street’s dilemma can be extended by an appeal to evolutionary contingency, which serves to refute some of the most popular third-factor accounts. An evolutionary rebuttal of Ruse’s contingency thesis, in turn, serves to corroborate the evolutionary premise of the Darwinian Dilemma.

3.5 Denying Influence: A Solution for Realists?

Might realists be able to resist the debunker’s argument? Perhaps the most promising strategy for realists is to take a more radical stance and to resist the premise that the contents of our moral judgments have been saturated with evolutionary influence. In this section I present the best case for realists who pursue this strategy. I will argue that some realists are indeed well positioned to resist Street’s and Ruse’s empirical claims and to argue that evolution’s causal influence in shaping our moral endorsements has been relatively minor. As I will argue in the next section, however, realists who pursue this strategy are still committed to a controversial empirical claim – namely, that over the course of human history we have been able to track mind-independent moral truths.

3.5.1 The Truth-Tracking Hypothesis

Consider how realists who deny the empirical premise of the Darwinian Dilemma might think of the relation between evolution and our moral judgments. These realists may grant that

³⁵ That morality has evolved only in one species does not imply that it has evolved only once. In fact, a moral capacity may well have evolved in several human populations.

evolutionary forces have had a moderate influence on our moral endorsements, in two respects. First, they may submit that evolutionary influences have thoroughly influenced the contents of *some* of our moral judgments, but submit that this influence does not generalize. Second, they may grant Street's claim that natural selection has shaped the contents of our *basic* moral evaluations, but resist the further claim that this influence also affects our *full-fledged* moral endorsements. Instead, an adequate explanation of the latter should involve an additional factor: the process of grasping stance-independent moral truths.

Indeed, as it stands there may be good grounds for questioning Street's adaptive-link hypothesis: the hypothesis our moral judgments can be explained in evolutionary terms, since our basic evaluative tendencies – e.g. feelings of pleasure and pain, and the accompanying intuitions about what is to-be-pursued and what is to-be-avoided – originated as adaptive responses to environmental circumstances. Even if our basic moral tendencies evolved as adaptive responses, it remains to be shown that they also provide the main source of input for our *full-fledged* judgments – and it is questionable whether they do in fact provide this input. Street acknowledges that our full-fledged judgments can 'stray, perhaps quite far, from alignment with our more basic evaluative tendencies' (Street 2006, p. 120), and stresses that it 'is likely that we were selected above all else to be extremely flexible when it comes to our evaluative judgments' (Street 2006, p. 158, fn. 20). But if our full-fledged moral judgments are rather different from our basic proto-moral tendencies, then Street's adaptive-link account fails to explain, for the most part, the contents of our full-fledged moral endorsements.

Several realists (e.g. Shafer-Landau 2012; FitzPatrick 2015; Huemer 2016) have pursued a reply to EDAs along these lines, arguing that the contents of many archetypical moral judgments cannot obviously be explained in evolutionary terms. Consider moral judgments that foster inclusionary values, for example the judgment that all human beings ought to be treated respectfully, irrespective of their capacities or group membership, or the judgment that non-human animals should be treated as subjects of moral consideration (cf. Buchanan and Powell 2015). These value-judgments have only won widespread adherence in recent human history, which raises some doubt over the suggestion that their contents have been deeply influenced by evolutionary forces. Moreover, it is not obvious that such judgments are fitness-enhancing.

The most important factor shaping our moral endorsements, these realists maintain, is our capacity to grasp stance-independent truths. How did this capacity originate? Some realists have sought to explain it as a by-product of our general emotional and intellectual capacities. These capacities, perhaps stimulated by the emergence of human language and our subsequent capacity to systematically reflect upon and reason about our moral judgments, have enabled us to track truths of a mind-independent reality. FitzPatrick (2015, p. 889), for example, regards this truth-tracking capacity as an 'intelligent extension of evolutionarily influenced evaluative judgment'.

If this account is along the right lines, then realists are well positioned to criticize EDAs on empirical grounds. First, consider Ruse's contingency EDA. If rational moral reflection, fuelled by the recognition of stance-independent moral truths, has been the dominant factor shaping our moral judgments, then the contents of these judgments are likely to be invariant with respect to our species-typical characteristics. As a result, Ruse's suggestion that our value-judgments are evolutionarily contingent loses much of its plausibility. Whatever biological

inclinations Darwin’s hive-bees or Ruse’s termites might have, this need not translate into what they regard as a ‘sacred duty’ or ‘ethical imperative’. If hive-bees would indeed make moral judgments very different from ours, then their rational faculty has probably been ill-calibrated: their judgments are mistaken.

Next, consider Street’s Darwinian Dilemma. If we have indeed evolved a capacity to track stance-independent moral truths, then the first horn of Street’s dilemma is unproblematic for realists: the general reliability of our moral judgments is trivially guaranteed. Moreover, if the influence of natural selection, as well as other evolutionary forces, on the contents of our moral judgments turns out to be modest, then on the second horn of the dilemma, Street’s evolutionary explanation is no longer obviously superior to the realist’s truth-tracking account. After all, there is no longer a strong correlation between our evolved judgments and moral truths, and the weaker this correlation, the easier it will be for realists to explain it.

3.5.2 Divide-and-Conquer Strategy

As previously stated, realists who maintain that the best explanation of our moral endorsements is that we have grasped stance-independent moral truths may still grant that evolution has influenced the contents of *some* of our moral judgments. Whether this influence has been distorting or not – i.e. whether evolutionary forces have pushed our judgments towards or away from the independent moral truths – should be decided on a case-by-case basis. Accordingly, realists may advance a divide-and-conquer strategy against debunkers (cf. Berker 2014) by partitioning our moral endorsements into different subsets and independently considering their justificatory status:

- Some moral judgments (e.g. ‘We ought to foster the well-being of future generations’) have hardly been shaped by evolutionary forces. Instead, they are primarily products of our capacity to reflect on stance-independent moral truths. This is particularly plausible for moral judgments in favour of inclusionary moral commitments, which are typically beyond the scope of evolutionary explanations (Buchanan and Powell, 2015).
- Some moral judgments have been shaped by natural selection but simultaneously reflect stance-independent moral truths. The earlier-mentioned judgment that ‘pleasure is generally good and pain generally bad’ might be such a judgment. Arguably, something similar is true for several moral judgments that are explicable in terms of kin selection (‘If forced to choose, you ought to favour the well-being of your own children over the well-being of strangers’), reciprocal altruism (‘The fact that someone has treated you well is a reason to treat that person well in return’) or indirect reciprocity (‘The fact that someone is an example to society is a reason to admire her’). Note that for this part of the divide-and-conquer strategy to work, realists still have to dismantle the second horn of Street’s dilemma and argue for the superiority of a tracking hypothesis over an adaptive-link account (see Artiga 2015 and Deem 2016 for such attempts).
- A third class consists of moral judgments that are the product of evolutionary forces which have led us away from the moral truths. For instance, ethicists typically argue that judgments fuelled by racial bias or by moralized disgust reactions are morally off-

track. Evolutionary psychologists typically argue that such responses are largely explicable in evolutionary terms. Hence, realists may argue that these judgments belong to a class of evolved responses that should be regarded as generally unreliable.

3.6 Best Historical Explanation: A Comeback for Debunkers

Realists who submit that the act of grasping mind-independent truths might play a legitimate – and important – role in causal explanations of our moral endorsements, and who are not impressed by the evidence that purports to show that these endorsements have been influenced by evolutionary forces, are well positioned to put pressure on EDAs. But as I will argue in this section, debunkers might be able counter this argument by shifting their empirical considerations from the evolutionary to the historical domain. Even if our full-fledged judgments have not been saturated with evolutionary influences, they may still be saturated with other causal influences – causal influences that typically belong to the explanatory domain of historians. If the best historical explanation of our moral endorsements nowhere involves an appeal to the fact that over the course of human history we have tracked mind-independent moral truths, then the realist’s truth-tracking hypothesis should be abandoned on historical grounds, or so debunkers might argue.

3.6.1 Historical Debunking Argument

How does this historical debunking argument relate to Street’s EDA? As noted in the previous section, Street (2006) herself favours an adaptive-link account of the origins of our moral judgments. But she also seems to be aware of the limitations of this account and is sympathetic to the view that our basic attitudes are only weak determinants of our full-fledged judgments. Indeed, she highlights that apart from the indirect influence of natural selection, our full-fledged judgments have also been shaped by various other processes, such as social learning and deliberation, as well as the *sui generis* influence of rational reflection (see Street 2006, esp. sections 4 and 5).

If these various non-evolutionary factors play a major role in shaping the contents of our moral judgments, then it is not obvious that the Darwinian Dilemma can get off the ground. But the ‘best explanation challenge’ embodied in the second horn of the dilemma need not be articulated in evolutionary terms. Debunkers can appeal to our best scientific explanations more broadly – particularly to explanations from the field of history – to argue that antirealist theories are superior on empirical grounds. The crucial condition for such explanations to have any debunking force is that they show that the best, scientifically informed genealogies of our moral endorsements involve no appeal to stance-independent moral truths.

Importantly, this debunking argument will only be problematic for realists who maintain that mind-independent moral truths have causal powers and that appealing to these truths is an integral part of the best causal explanation of our moral endorsements. But this, we have seen, is indeed the view of realists who deny the empirical premise of the Darwinian Dilemma (Shafer-Landau 2012; FitzPatrick 2015; Huemer 2016). These realists are committed to a *scientific hypothesis* – namely that we have been able to track, over the course of human evolution or history, stance-independent moral truths. In other words, by positing the existence of stance-independent moral truths and arguing that these truths cause our judgments, these realists enter the game of scientific explanation. Assuming that compatibility with our best

scientific explanations is regarded as an important metaethical *desideratum*, if it turns out that this hypothesis cannot be corroborated – or is even clearly invalidated – by our best science, this will count as highly damaging to these realist positions.

3.6.2 Genealogies without Stance-Independent Truths

Do the best, scientifically informed genealogies of our moral endorsements indeed involve no appeal to stance-independent moral truths? Answering this question in appropriate detail requires a historical case study of specific moral endorsements, which is beyond the scope of the present discussion. Still, to give an impression of how debunkers might proceed, consider the following brief sketch of the genealogy of inclusionary moral values.

In *The Descent of Man*, Darwin observed that the process of civilization comes along with a broadening of human's moral concern:

As man advances in civilization, and small tribes are united into larger communities, the simplest reason would tell each individual that he ought to extend his social instincts and sympathies to all the members of the same nation, though personally unknown to him. This point being once reached, there is only an artificial barrier to prevent his sympathies extending to the men of all nations and races. (Darwin 2013, p. 76)

Indeed, many contemporary moral judgments promote inclusionary values. How did this extension of moral concern come about, if not by grasping the stance-independent truth that 'men of all nations' deserve equal respect?

Debunkers may argue as follows. First, even if inclusionary moral judgments have not directly been shaped by natural selection, they may still be *indirectly* saturated with evolutionary influence. For instance, there may be an evolved empathy response underlying our inclusionary judgments which has been co-opted in the process of moral reflection and has subsequently been extended beyond kith and kin. Second, debunkers are not committed to the claim that the *only* input in this process of moral reflection comes from our evolved attitudes. The input is likely to be much broader and to include our socioculturally developed attitudes and factual knowledge, as well as the norm of impartiality that typically guides moral reflection – a norm whose origins may plausibly be explained in evolutionary terms (cf. DeScioli and Kurzban 2013). Third, debunkers may point out that it took a notoriously long time before sympathies were extended to 'men of all nations and races', and that even nowadays the boundaries of our moral concern are contested. This observation should not surprise us if moral judgments are largely the products of our evolved inclinations, lessons learned from history and ongoing social dialogue. By contrast, the observation seems difficult to reconcile with the hypothesis that we have grasped the stance-independent moral truth that all people deserve equal respect.³⁶

³⁶ Realists may, of course, reply that it is very difficult to discover moral facts, and they may seek to explain why some of these facts have only been discovered recently (cf. Huemer 2016). The question, then, becomes whether this rival explanation is compatible with our best historical knowledge and superior to the antirealist's account.

3.6.3 Making Up the Balance

A crucial difference between the accounts of realists and debunkers rests on the question of whether the best explanation of our moral beliefs involves an appeal to stance-independent truths. This brings us back to the question with which we began – namely, which empirical considerations are relevant to the success of EDAs. If my argument has been along the right lines, then this empirical input is not limited to evolutionary considerations. In fact, it may well turn out that the most revealing test for the empirical tenability of moral realism is to scrutinize the *historical* genealogy of our moral endorsements in detail and to evaluate whether the relevant data are best explained with or without appealing to stance-independent moral truths. In the next chapter, I will discuss such historical explanations at greater length.

3.7 Conclusion

I have argued that the evolutionary considerations that Street and Ruse invoke can reinforce each other in presenting the evolutionary debunker's strongest case against moral realism. For realists who try to block Street's EDA by proposing a third-factor account based on a general moral assumption, the contingency of the moral judgments that are evolvable in the light of this assumption seems problematic, as it allows debunkers to reinstate a version of the reliability challenge. For realists who try to block Ruse's EDA by arguing that the contents of our moral judgments are evolutionarily convergent or constrained by developmental factors, the strong causal influence of evolutionary forces on our moral judgments thereby implied seems problematic, as it reinforces the general evolutionary claim underlying Street's Darwinian Dilemma.

Perhaps the most promising reply of realists to evolutionary debunkers is to deny that evolutionary forces have saturated the contents of our full-fledged moral judgments. Several realists have insisted that these contents are better explained by our recognition of stance-independent moral facts. In order to support this thesis, realists have to vindicate a scientific hypothesis: over the course of evolution and human history, we have been able to track stance-independent moral truths. Since the thesis that we have been able to track stance-independent truths is key to a success epistemology of moral realism, vindicating this hypothesis is of crucial importance to its overall metaethical standing. Debunkers, in turn, may argue that the best historical explanation of our moral endorsements involves no such appeal. What counts as the best historical explanation of the contents of our moral endorsements, then, will be crucial for settling this debate.

References

- Artiga, M. 2015. Rescuing Tracking Theories of Morality. *Philosophical Studies* 172: 3357-3374.
- Berker, S. 2014. Does Evolutionary Psychology Show that Normativity is Mind-Dependent? In: J. D'Arms and D. Jacobson, eds. *Moral Psychology and Human Agency: Essays on the Science of Ethics*. New York: Oxford University Press, 215-252.
- Bogardus, T. 2016. Only all Naturalists Should Worry about Only One Evolutionary Debunking Argument. *Ethics* 126: 636-661.

- Braddock, M. 2016. Evolutionary Debunking: Can Moral Realists Explain the Reliability of Our Moral Judgments? *Philosophical Psychology* 29: 844-857.
- Buchanan, A. and Powell, R. 2015. The Limits of Evolutionary Explanations of Morality and Their Implications for Moral Progress. *Ethics* 126: 37-67.
- Clarke-Doane, J. 2012. Morality and Mathematics: The Evolutionary Challenge. *Ethics* 122: 313-340.
- Crow, D. 2016. Causal Impotence and Evolutionary Influence: Epistemological Challenges for Non-Naturalism. *Ethical Theory and Moral Practice* 19: 379-395.
- Cuneo, T. and Shafer-Landau, R. 2014. The Moral Fixed Points. New Directions for Moral Nonnaturalism. *Philosophical Studies* 171: 399-443.
- Darwin, C. 2013 [1871]. *The Descent of Man and Selection in Relation to Sex*. Ware: Wordsworth Classics.
- Das, R. 2016. Evolutionary Debunking of Morality: Epistemological or Metaphysical? *Philosophical Studies* 173: 417-435.
- Deem, M.J. 2016. Dehorning the Darwinian Dilemma for Normative Realism. *Biology & Philosophy* 31: 727-746.
- DeScioli, P. and Kurzban, R. 2013. A Solution to the Mysteries of Morality. *Psychological Bulletin* 139: 477-496.
- Enoch, D. 2010. The Epistemological Challenge to Metanormative Realism: How Best to Understand It, and How to Cope with It. *Philosophical Studies* 148: 413-438.
- FitzPatrick, W. 2014. Why There Is no Darwinian Dilemma for Ethical Naturalism. In: M. Bergmann and P. Kain, eds. *Challenges to Moral and Religious Belief: Disagreement and Evolution*. Oxford: Oxford University Press, 237-255.
- FitzPatrick, W. 2015. Debunking Evolutionary Debunking of Ethical Realism. *Philosophical Studies* 172: 883-904.
- Huemer, M. 2016. A Liberal Realist Answer to Debunking Skeptics: The Empirical Case for Realism. *Philosophical Studies* 173: 1983-2010.
- Joyce, R. 2016. Reply: Confessions of a Modest Debunker. In: U.D. Leibowitz and N. Sinclair eds. *Explanation in Ethics and Mathematics*. Oxford: Oxford University Press, 124-148.
- Klenk, M. 2017. Old Wine in New Bottles: Evolutionary Debunking Arguments and the Benacerraf-Field Challenge. *Ethical Theory & Moral Practice* 20: 781-795.
- Ruse, M. 1995 [1986]. Evolutionary Ethics: A Phoenix Arisen. In: P. Thompson, ed. *Issues in Evolutionary Ethics*. Buffalo, NY: SUNY Press, 225-247.
- Schechter, J. 2018. Explanatory Challenges in Metaethics. In: T. McPherson and D. Plunkett, eds. *Routledge Handbook of Metaethics*. New York: Routledge, 443-459.
- Schloss, J. 2014. Evolutionary Accounts of Morality: Accounting for the Normal but not the Normative. In: H. Putnam, S. Neiman and J. Schloss, eds. *Understanding Moral Sentiments: Darwinian Perspectives*. New Brunswick: Transaction Publishers, 81-121.
- Shafer-Landau, R. 2003. *Moral Realism: A Defence*. Oxford: Oxford University Press.
- Shafer-Landau, R. 2012. Evolutionary Debunking, Moral Realism and Moral Knowledge. *Journal of Ethics and Social Philosophy*, www.jesp.org, 7(1): 1-37.

Chapter 3

- Skarsaune, K.O. 2011. Darwin and Moral Realism: Survival of the Fittest. *Philosophical Studies* 152: 229-243.
- Street, S. 2006. A Darwinian Dilemma for Realist Theories of Value. *Philosophical Studies* 127: 109-166.
- Street, S. 2008. Reply to Copp: Naturalism, Normativity, and the Varieties of Realism worth worrying about. *Philosophical Issues* 18: 207-228.
- Street, S. 2016. Objectivity and Truth: You'd Better Rethink It. In: R. Shafer-Landau, ed. *Oxford Studies in Metaethics, vol. 11*. Oxford: Clarendon Press, 293-334.
- Tersman, F. 2017. Moral Skepticism and the Benacerraf Challenge. In: D. Machuca, ed. *Moral Skepticism: New Essays*. New York: Routledge, 183-198.
- Wielenberg, E. 2010. On the Evolutionary Debunking of Morality. *Ethics* 120: 441-464.

4

Chapter 4

Explaining Historical Moral Convergence: The Empirical Case against Realist Intuitionism

Accepted by *Philosophical Studies*, pending minor revisions.

Keywords

*Michael Huemer; moral progress; moral convergence; rationalist intuitionism;
consistency reasoning*

Abstract

Over the course of human history there appears to have been a global shift in moral values towards a broadly ‘liberal’ orientation. Michael Huemer (2016) argues that this shift better accords with a realist than an antirealist metaethics: it is best explained by the discovery of mind-independent truths through intuition. In this chapter I argue, *contra* Huemer, that the historical data are better explained assuming the truth of moral antirealism. Realism does not fit the data as well as Huemer suggests, whereas antirealists have underappreciated resources to explain the relevant historical dynamics. These resources include an appeal to socialization, to technological and economical convergences, to lessons learned from history, to changes induced by consistency reasoning and to the social function of moral norms in overcoming some of the cooperation problems that globalizing societies face. I point out that the realist’s *explanans* has multiple shortcomings, that the antirealist’s *explanans* has several explanatory virtues, and conclude that the latter provides a superior account of the historical shift towards liberal values.

‘The chief objection to this ‘naturalistic ethics’ is that it overlooks the fact that, even from the Darwinian point of view, the human animal is a human animal.’

– John Dewey (1908)

4.1 Introduction

Moral realists maintain that moral truths are underwritten by mind-independent moral properties or facts and that these properties or facts are epistemically accessible to moral agents.³⁷ It has often been argued that moral realists cannot make sense of the fact that between cultures and over the course of history there has been pervasive moral disagreement (e.g. Mackie 1977; Doris and Plakias 2008). But the empirical record suggests that there is also agreement on a substantial number of moral issues. Moreover, attitudes about many moral issues seem to have converged over time. If moral *divergence* provides evidence against realism, does moral *convergence* provide evidence in support of it?

Michael Huemer (2016) has recently argued so. According to Huemer, the historical record exhibits a clear trend of moral convergence: over the course of centuries, many societies worldwide have been moving towards a set of broadly liberal values (I identify these in **section 4.3**). This trend meshes well with the contention that moral inquiry is fuelled by a somewhat reliable process of tracking mind-independent moral truths through intuition, Huemer claims. If we have at least *some* epistemic access to mind-independent moral facts, then over the *longue durée* of human history our moral values should start to converge.

Huemer presents his argument as an inference to the best explanation, concluding that given the good fit that a realist account has with the historical data ‘it is reasonable to endorse moral realism, unless and until a better account appears’ (idem, p. 2008). The aim of the present chapter is to provide a better account. I argue that the relevant historical data, wedded with insights from cognitive science, moral psychology and evolutionary anthropology, are best explained by assuming some version of an antirealist metaethics – that is, by assuming that there are no mind-independent moral truths.³⁸ Not only can antirealists overcome the challenges that Huemer raises for them; in fact, their account has several virtues which makes it superior, in terms of empirical explanation, to Huemer’s realist account.

The chapter proceeds as follows. In **section 4.2** I outline the realist position that Huemer favours, and the view of moral progress that it entails. In **section 4.3** I summarize the empirical data that Huemer cites in support of this view. In **section 4.4** I point out why Huemer thinks that his realist explanation for these data is superior to that of antirealists, by highlighting three

³⁷ To what extent moral truths are epistemically accessible, assuming realism, depends on the realist view in question, as I illustrate in **section 4.2**. Realists are not committed to the view that moral truths are accessible to anybody at any moment in time. But all realists maintain that moral truths are *at least somewhat* accessible, such that the skeptical implications pressed by Street (2006, 2016) can be resisted.

³⁸ In this chapter I use the label ‘antirealism’ in the same broad sense as Huemer (2016, p. 1983), as denoting all views which either deny that there are mind-independent moral truths, or that we can have knowledge of such truths. Note that this is an encompassing set of views, which includes versions of non-cognitivism, error theory, as well as the constructivist view that moral truths are underwritten by mind-dependent moral properties or facts (in **chapter 5** I specifically discuss an antirealist view in the latter sense).

challenges Huemer poses to antirealists. In **section 4.5** I address these challenges, and argue that antirealists can draw upon powerful and underappreciated resources to explain the historical trend of moral convergence. In **section 4.6** I move from defence to offense: not only can antirealists overcome Huemer's challenges, but in fact their account has several explanatory virtues, whereas Huemer's account has several explanatory shortcomings. I conclude (**section 4.7**) that the relevant historical data are best explained by assuming an antirealist metaethics.

4.2 Realism and Diffusing Explanations

Realists have fashioned various responses to the challenge of moral disagreement (e.g. Shafer-Landau 2003; Enoch 2009). Similar responses play a role in realist arguments in favour of moral convergence. Typically, such responses involve two components: an empirical challenge to the claim that widespread moral disagreement exists, and an explanation of extant disagreement in terms compatible with realism. Following Doris and Plakias (2008), let's call the latter a *diffusing explanation*.

Diffusing explanations aim to show that moral disagreement is due to contingent historical, cultural or epistemic factors rather than the absence of mind-independent moral facts. They may serve to show, for instance, that

- 1) application of the same moral principles yields different verdicts in different cultural contexts (e.g. in situations of scarcity versus situations of plenty);
- 2) apparent moral disagreement is actually due to disagreement over non-moral issues (e.g. the truth of theism);
- 3) individuals are making moral judgments on an epistemically deficient basis (e.g. by echoing cultural biases).

As these diffusing explanations reveal, realists need not commit themselves to the claim that moral inquiry necessarily leads to the successful discovery of moral facts. Indeed, such a commitment would be extremely difficult to maintain in the light of recent work in moral psychology, which shows that moral reasoning is susceptible to various biases (e.g. Schwitzgebel and Cusman 2012; Fitzpatrick 2014). But in order to resist global scepticism, realists *should* maintain that moral inquiry is *at least somewhat* reliable: we are not hopeless at tracking the mind-independent moral facts.

Diffusing explanation (3) plays a key role in the version of moral realism – *rationalist intuitionism* – that Huemer (2008, 2016) defends. The central tenet of rationalist intuitionism is that human beings have a capacity to acquire *a priori* knowledge of mind-independent moral facts. The means by which we acquire such knowledge are our moral intuitions, which Huemer characterizes as ‘initial intellectual appearances, that is, states of its seeming to one that something is the case upon intellectual consideration (as opposed to sensory observation or introspection), but prior to reasoning’ (Huemer 2008, pp. 370-371). Not all moral intuitions are reliable; acquiring moral knowledge is a fallible process, just like acquiring scientific or

mathematical knowledge.³⁹ There are various types of psychological bias, reasoning errors and cultural dogmatism that might lead us to have defective intuitions. These biases are pervasive and not easily overcome; attempting to overcome them has taken up the bulk of moral history and is an ongoing process.

On Huemer's view, then, moral progress is made by overcoming biases. In the distant past, most people had a very misguided set of moral beliefs – just like most people had a very misguided set of scientific beliefs. But a capacity to recognize – however vaguely – mind-independent moral facts has created a slow yet systematic convergence of people's moral beliefs towards the moral truths. This push was triggered by moral reformers, who diverged from social dogma: their 'intellectual seemings' were less biased than those of others, and their views spread. Thanks to these reformers – whose role is not unlike that of protagonists in the history of science and mathematics – societies have gradually changed their moral stances and have adopted roughly the same set of liberal values. This is not to say that the process of moral development has reached its end and that, at present, we have become acquainted with the precise or complete set of moral truths. But we do know, Huemer claims, that our current liberal values are better than the values of the past. Whatever the precise moral truths turn out to be, it is likely that the historical shift towards liberal values has taken us a long way towards recognizing them.

4.3 The Historical *Explanandum*

By *liberal values*, Huemer does not mean the values of political liberalism. Instead, what he has in mind is a broad ethical orientation which consists of three related commitments: (i) recognition of the moral equality of persons; (ii) respect for the dignity of the individual; (iii) opposition to gratuitous coercion and violence. Huemer cites several empirical trends in support of the view that societies have been converging towards this liberal orientation, many of which are familiar from Steven Pinker's discussion in *The Better Angels of our Nature* (2011). Key data are the lowering rates of violence, war and murder from prehistoric societies up until the present; the widespread abolition of torture and execution over the past four centuries; the universal outlawing of slavery over the past two centuries; the ongoing opposition against racism and sexism, particularly over the last century; the spread of democracy, particularly over the last few decades; and the non-violent means by which former colonies of liberal states have gained sovereignty in modern times.

Of course, there are counterexamples to the trends Huemer cites. Unfortunately, in spite of legal prohibitions, practices of slavery, racism and sexism are still widespread. Over the past decade, some countries have clearly lessened their commitment to i–iii. Over the past century, we have witnessed two world wars and many instances of violence, as well as genocide. Over the past millennia, commitments to human equality have waxed and waned. Rather than following a straight line leading up to the adoption of liberal values, our moral history has been full of twists and turns. Undoubtedly, recognizing a historical trend towards liberalism is partly

³⁹ A challenge for Huemer is to elucidate how people are able to differentiate between truth-apt intuitions and intuitions clouded by bias. Huemer's (2008) main strategy for doing so is to differentiate between intuitions at different levels of generality and to argue that intuitions about some levels (e.g. intuitions about concrete cases) are more susceptible to bias than others (e.g. formal intuitions).

a consequence of selection bias and partly the assumption of a Western perspective.

One strategy that critics of Huemer's (2016) argument could employ, therefore, is to dismiss the supposed trend towards liberal values on empirical grounds. But we should keep in mind that specific historical counterexamples need not suffice to undermine Huemer's general empirical claim: diffusing explanations might help realists to explain local episodes of moral regress. Perhaps, at certain points in history, some cultures have not become *less*, but *more* clouded by bias. As long as the *overall* historical trend suggests that there has been a trend of moral convergence, realists might argue, the empirical data still give support to liberal realism.

In what follows I shall, for the most part, pursue a different strategy. While I think that the cited trend is less robust than Huemer suggests, it would be too quick to dismiss the data that Huemer (2016, section 3) puts forward as a mere fluke. Huemer makes a good case that there has been a general moral development, accelerating over the last two centuries, in a direction that we typically regard as progressive – an *explanandum* that is metaethically relevant. Moreover, what makes this trend particularly striking is the fact that it can be recognized in different societies and with regard to different issues (war, murder, slavery, democracy, women's suffrage, racial segregation, torture, execution, colonization), each moving in a similar direction.

How should we explain this coincidence? The best explanation, Huemer contends, is that over time different societies have come to accept the same mind-independent liberal moral truths. I shall argue, by contrast, that the relevant historical data are better explained in antirealist terms.

4.4 Why Huemer takes the Realist *Explanans* to be Superior

It is important to keep in mind that in speaking of the best explanation in this context, what interests us is the best *historical* explanation of the relevant data. This historical explanation, in turn, may proceed against a background of specific metaethical assumptions – in the present case either by positing the existence of mind-independent moral truths (moral realism), or by denying that such truths exist, or that they are mind-independent (moral antirealism). Hence, we are interested in the following question: from a historian's point of view, are the relevant data better explained against the background of a realist or an antirealist metaethics?⁴⁰

Huemer attempts to establish the greater plausibility of a realist explanation *inter alia* by advancing three challenges for antirealist explanations.⁴¹ I outline these challenges in the present section, and reply to them in **section 4.5**.

4.4.1 Evolutionary Explanations and the 'Scope Challenge'

In recent years it has become fashionable among antirealists to explain the basic contents of our moral judgments from a broadly evolutionary perspective (e.g. Joyce 2006; Street 2006;

⁴⁰ Naturally, to accord with the historian's best explanation is not the only *desideratum* for metaethical theories. But it is certainly a weighty *desideratum*, which may tilt the overall balance of metaethical plausibility.

⁴¹ In fact, Huemer (2016, section 4) discusses – and dismisses – not three but four possible antirealist explanations for the shift towards liberalism. In the present section I highlight only three of these explanations – those that strike me as most plausible.

Kitcher 2011). But are evolutionary explanations also suitable to account for more recent developments, such as the historical trend towards liberalism? Huemer thinks that they are not: this recent trend lies beyond the scope of evolutionary explanations (see also FitzPatrick 2012; Buchanan and Powell 2015). We may call this Huemer's *scope challenge*: the challenge for antirealists to show that recent value changes are not beyond the explanatory scope of evolutionary theory.

Consider the anti-racism movement in the United States, which only led to the abolishment of the Jim Crow laws in the 1960s. It is highly implausible that this moral accomplishment had much to do with the spread of specific genes. Huemer argues that the realist's alternative is much more plausible: what explains the success of the anti-racist campaign is that the campaigners discovered a moral truth, and that this truth – racism is morally wrong – was generally recognized during the second half of the 20th century.

It should be kept in mind that realists are not committed to the view that this moral shift was *entirely* driven by the recognition of this mind-independent moral truth, or that evolution plays *no role* whatsoever in explaining our moral beliefs (cf. FitzPatrick 2012, p. 174). The point of the *scope challenge* is that appeals to evolution do not *suffice* to explain the contents of our moral endorsements. Especially where moral values are concerned that have only recently been adopted, such explanations seem problematic. Therefore, Huemer claims, the antirealist's evolutionary account falls short.

4.4.2 Cultural Explanations and the 'Coincidence Challenge'

Apart from appealing to evolutionary influences, antirealists can also appeal to cultural influences to explain the contents of our moral judgments. Indeed, for many moral judgments – especially recent ones – this seems to be a more obvious explanation. Huemer, too, concedes that the shift towards liberalism can largely be explained in terms of cultural forces. Not all individuals have independently figured out that slavery and racism are wrong. Instead, many people are simply taught that this is true as part of their moral upbringing. But Huemer also insists that cultural forces don't provide us with an exhaustive explanation, since they do not illuminate *why* cultures worldwide have converged towards a liberal orientation. Let's call this Huemer's *coincidence challenge*: the challenge for antirealists to explain why different cultures have moved globally in a similar direction with regard to several moral issues. According to Huemer, this coincidence is too salient to be left unexplained:

Are we to believe it is coincidence that, at the same time that slavery was becoming economically inefficient, some other trend was leading women's suffrage to become more popular (...), another trend was causing democracy to spread across the world, another was causing war to seem less glorious, another made torture seem less beneficial, and so on? This is not just a series of unrelated changes; they are all changes in line with a certain coherent ethical perspective: all the changes fit together, in one way or another, with the value of equal respect for the dignity of persons. (Huemer 2016, p. 1999)

Antirealists do not have an obvious explanation for this coincidence, Huemer claims. Realists, on the other hand, do: different cultures have converged by recognizing the same mind-independent moral truths.

4.4.3 Rational Explanations and the ‘Inference Challenge’

A third possible explanation for the shift towards liberalism refers to our rational capacities. In progressive moral transitions, moral reasoning presumably plays an important role. Consider the view advanced by Peter Singer, according to which moral progress is driven by an expansion of people’s circle of moral concern. That we should expand our circle of concern, Singer holds, is a truth that we can track through moral reasoning. But Huemer points out that it is not a truth of logic. In fact, it seems inescapable to appeal to intuitive moral truths in explaining progressive moral shifts. We may call this Huemer’s *inference challenge*: the challenge for antirealists to explain how moral reasoning, unaided by intuitions about mind-independent moral truths, can lead to progressive moral changes.

By way of example, consider the following argument:

- I) *Prior judgment*: We should respect the interests of individuals within our tribe.
- II) *New moral insight*: There is no morally relevant difference between individuals within and outside our tribe.
- III) *Posterior judgment*: We should respect the interests of individuals outside our tribe.

On Singer’s view, this kind of inference is illustrative of progressive moral change. But where does the new moral insight (II) come from? There appears to be no good explanation for this insight, other than an appeal to a deep-seated moral intuition – precisely as Huemer’s non-rationalist intuitionism suggests. Hence, it appears that realist intuitionists are better able than antirealists to elucidate the role that rational intuitions play in progressive moral change.

4.5 How Antirealists can Counter Huemer’s Challenges

In the previous section I outlined Huemer’s main considerations for thinking that, assuming an antirealist metaethics, we cannot adequately explain the historical shift towards liberalism. In this section I challenge these considerations. I will argue that antirealists can meet Huemer’s *coincidence challenge* and *inference challenge*. The *scope challenge* cannot easily be met, but for many antirealists this won’t be much of a concern, since the challenge does not target their view.

If my arguments in this section will be along the right lines, they undercut Huemer’s historical argument for the presumed advantage of realism over antirealism. Moreover, in **section 4.6** I will proceed to argue that the antirealists’ historical explanation is in fact superior.

4.5.1 Answering the ‘Scope Challenge’

Let’s start with the *scope challenge*. We should agree with Huemer that the contents of some – perhaps many – moral judgments are beyond the explanatory scope of evolutionary theory. But it should also be noted that antirealists are not typically committed to thinking otherwise. Kitcher (2011), for example, holds that a capacity for normative guidance originally served the

evolutionary functions of overcoming a lack of responsiveness to others and fostering human cooperation. But Kitcher also holds that the emergence of moral thought eventually prompted an ongoing ethical discussion, which can go in any number of directions – and is not strictly tied to its evolutionary function. Or take Street (2006), who argues that our ‘basic evaluative tendencies’ have been directly shaped by natural selection. But Street is not committed to the adaptationist view that the same holds for the contents of our ‘full-fledged evaluative judgments’. By contrast, she grants that various social, cultural and historical influences, as well as the *sui generis* influence of rational reflection, have influenced these contents (Street 2006, p. 114). For these antirealists, then, it won’t be much of a concession to admit that values that are of relatively recent origins are beyond the scope of evolutionary explanations. What matters is whether they can adequately be explained in other antirealist-friendly terms – that is, without positing the existence of mind-independent moral truths.

4.5.2 Answering the ‘Coincidence Challenge’

A promising strategy for antirealists to dismantle Huemer’s *coincidence challenge* is to point out that the historical processes Huemer cites are entangled. Here I pursue a strategy along these lines, in three steps. First, I argue that the coincidence poses less of an explanatory burden for antirealists than Huemer suggests, because the coinciding phenomena directly influence each other. Second, I point out that the coinciding phenomena are indirectly influenced by shared common causes, which further reduces the antirealist’s explanatory burden. Additionally, as a third step, I point to the directionality of historical change: for what remains of a coincidence, antirealists are well placed to explain it in functional terms.

4.5.2.1 First Step: Direct Cultural Influence

Coincidences are in need of explanation insofar as the coinciding phenomena are independent (if not, this explains their coincidence). But for many of the historical trends that Huemer highlights – the abolition of torture, execution and slavery; the opposition to racism and sexism; decolonization and the spread of democracy – such independence is questionable. As a result, the coincidence constitutes less of an explanatory burden for antirealists than might initially seem to be the case.

First, these trends are not fully independent because they have influenced each other. For instance, throughout history slavery has tended to be a racist practice; the opposition to slavery and racism are historically entangled. Similarly, the rise of democracy was affected by the process of decolonization and the emerging ideal of popular sovereignty. The arguments, tactics and language of empowerment used by oppressed groups have affected the emancipatory struggles of others. Consider the 19th and 20th century struggles for racial and sexual equality in the United States: even if these movements did not always support each other’s cause, they did influence each other with their mutual appeal to rights.⁴² While such

⁴² Rather than being understood as direct cultural influences, historical connections such as these may also be understood as stemming from a common cause – e.g. from the Enlightenment ideal of egalitarianism, or from a common belief in natural rights. In explaining complex historical dynamics, the distinction between direct causes and common causes is often difficult to make. For the purposes of the present argument, however, this distinction is only of secondary importance; antirealists can call upon either kind of influence to call into question the independence of convergent historical trajectories.

influences do not explain the *direction* of the historical shift itself, they do explain, to a large extent, the *commonality* of this direction with regard to many distinct issues. Rather than being seen as a concurrent series of directional changes, then, the historical trend that Huemer cites can largely be regarded as an interconnected historical process.

Second, societies have influenced each other. Rather than being developments that were driven by internal social dynamics, there is ample evidence that many of the convergences Huemer cites are actually due to the fact that societies have pushed each other in certain directions – by setting examples, by using soft power, by means of force and through negotiation. The ideals of the American Revolution set a precedent for its French counterpart; the Indian independence movement inspired many others. In a world with a global traffic of goods and information, in which societies depend on each other for resources and share mutual goals, it should not come as a surprise that they also come to adopt a roughly shared moral outlook. Indeed, given that social conformity is a very common phenomenon in moral reasoning (e.g. Haidt and Bjorklund 2008), and that people often internalize the norms of their cultural environment, it would have been a surprise if, after centuries of globalization, we had *not* seen values shift in a roughly shared direction.

Perhaps the most dramatic example of a global shift towards liberal values has been the adoption of the Universal Declaration of Human Rights (1948), which has the defence of human dignity and human rights at its very core. Obviously, not all countries independently came to support this declaration. Instead, its near-global endorsement was the outcome of deliberation and negotiation through a global platform: the United Nations. The same goes for many other international treaties and declarations, which have contributed to the joint adoption of a broad moral outlook: they are not the product of individual trajectories, but of mutual cultural influences.

4.5.2.2 Second Step: Non-Moral Convergence

The Universal Declaration of Human Rights (UDHR) brings us to another major explanatory component of the shared direction of moral change: the common historical experience of different societies. The UDHR was adopted after the Second World War, which provided the immediate impetus for its creation. Historical experience can provide a rich source for gaining moral insight – insight about which norms lead to human flourishing and to human suffering, and which institutions are effective in fostering human cooperation. Moreover, historical experiences are often shared by different societies and may constitute a common cause for convergent moral trajectories.

In fact, there are several common causes that may explain, at least in part, the global shift towards liberalism. For instance, over time many societies have developed similarly in terms of material wealth, technology, healthcare and scientific knowledge. These non-moral convergences, in turn, may explain much of the apparent convergence in moral terms. Of course, societies have not universally moved in the same direction – but neither is the trend towards liberalism a universal trend. That said, regarding Western countries, where the shift towards liberal values is most apparent, there have also been notable parallels in terms of non-moral changes.

Consider one of the trends that Huemer cites as part of his *explanandum*: the lowering rates of violence, war and murder over the past millennia. This trend is clearly correlated with

an increase in wealth, health, education and commercial activity, as Pinker (2011, p. xxiii) recounts. Moreover, while the exact causal dynamics will be difficult to pin down, we may hypothesize that the latter developments can also *explain* – at least in part – why violence, war and murder have declined: plausibly, conditions of poverty, poor health and poor education more easily give rise to violent struggle than conditions of affluence do. Or consider the abolition of slavery, which is correlated with the transition, at least in many Western countries, from agricultural to industrial societies. Huemer (2016, p. 1999) is aware of this correlation, but thinks it would be speculative to regard the relation as causal. But in fact it is not: Pleasants (2010) provides clear historical evidence that the rise of the market economy proved to be a crucial trigger for an effective abolitionist campaign (see also **section 4.6.2**).

4.5.2.3 Third Step: Functional Explanation

Similar material and informational conditions often invite similar social problems, and lend themselves to similar solutions – including solutions along moral lines. Many metaethicists – realists and antirealists alike – maintain that moral norms and values are typically functional. For instance, Gibbard (1990) argues that moral norms serve the function of coordinating our actions so that they comply with adaptive rules of conduct. Copp (2008) claims that morality has the function of enabling a society to meet its needs. Kitcher (2011) holds that moral norms serve the function of solving ‘altruism problems’ and fostering social cooperation. Even if their accounts differ regarding whether they provide a predominantly evolutionary or sociocultural explanation for the function of moral norms, these authors – along with many other metaethicists – stand united in arguing that the contents of moral norms can typically be explained in functional terms.

Let’s assume that these accounts are along the right lines and that moral norms are typically functional – in evolutionary, social or other terms. This observation is relevant for predictions about moral convergence, since functions give rise to constraints. If moral codes tend to be fitness-enhancing, then we should expect that in different societies, under similar conditions, these codes will typically evolve along convergent trajectories. Likewise, if moral norms serve to overcome social cooperation problems, then we should expect that in different societies facing similar cooperation problems, such norms will be similar in outline. Convergence need not be perfect; some problems can be solved in multiple ways. Often, though, similar problems call for similar solutions. Therefore, some degree of historical moral convergence should be our default expectation – for realists and antirealist alike.

With regard to the broad patterns of convergence that Huemer highlights, this broad appeal to moral functionalism provides antirealists with another instrument to explain the empirical data. Recall that the broadly liberal orientation in which Huemer is interested – of respecting human dignity, recognizing equality and opposing gratuitous coercion – gained global adherence over the last few centuries, with accelerating speed. Two major historical developments that took place during the same centuries were the transition to an industrial economy and the emergence of a global network of reliable and increasingly rapid information exchange. These developments, antirealists might argue, are very much entangled with our shifting moral perspectives. First, they initiated processes and made salient problems that were quite specific to this globalizing and industrializing world, such as the sustained interaction between people of different ethnicities and the emergence of global inequality. Second, the

economic and industrial changes afforded new forms of human interaction and cooperation, such as the transition from a system of slave labour to a system of wage labour.

Against this background, antirealists are well positioned to explain why over the course of history a coherent shift in our moral outlook has taken place, in a broadly liberal direction. In part, this shift can be explained as the result of cultural pressures and non-moral convergences; in part, the adoption of liberal values may be regarded as solving some of the cooperation challenges specific to a globalizing world. Additionally, to a large extent this shift has been a response to lessons learned from history. The outcome of the two world wars played a pivotal role in the accelerated adoption of liberal values during the second half of the 20th century: Europe's traumatic experience of rule under the decidedly violent and racist Nazi regime fuelled the liberal wave in the decades that followed. Western rivalry with – and the later collapse of – the Soviet Union subsequently aided the spread of liberal values.

4.5.3 Answering the 'Inference Challenge'

To provide a comprehensive explanation for the shift towards liberalism, antirealists also have to account for the role that reasoning played in it, without appealing to intuitions about mind-independent truths. Hence, they have to overcome what I have called Huemer's *inference challenge*. I will argue that, contrary to what Huemer suggests, antirealists have a powerful – and underappreciated – resource for doing so: the phenomenon of consistency reasoning (Campbell and Kumar 2012).

Applied moral reasoning is often driven by a norm of consistency – treating like cases alike. If cases are not treated alike, then a morally relevant difference between them should be found. Consider the following example:

Suppose that although I am an avid dog owner, I also eat meat. A vegetarian friend, however, presses me: what's the difference between factory farming and practices that I already consider abhorrent, like dog fighting? Faced with this challenge, and provided that I am disposed to trust my interlocutor, I should either decline to condemn dog fighting or, more likely, change my opinion about factory farmed meat. Unless I can find some morally relevant difference between these two practices, I should treat like cases alike. (Kumar forthcoming)

As Kumar points out, the upshot of exposing moral inconsistencies can be twofold. Either the moral agent revises (or is pressed to revise) one of her extant beliefs, such that consistency is restored, or she imposes consistency upon her extant beliefs by arguing that there is a morally relevant difference between them. In theory, of course, it is also possible to reject the metanorm that we should hold moral beliefs that are consistent, but in practice this option is rarely favoured: striving for consistency is an aim that moral reasoners are typically unwilling to give up.

Campbell and Kumar (2012) hypothesize that over the course of history dialogical appeals to consistency have served as an important engine of progressive moral change. This hypothesis seems plausible, too, with regard to the shift towards a liberal moral orientation. For instance, the demand to apply the same norms and privileges consistently to different people, irrespectively of gender, race or sexual orientation, has been a key argument in modern

emancipatory struggles. Similarly, many of Singer's influential arguments in support of animal welfare – which constitute a recent extension of Huemer's trend towards liberalism – are consciously driven by a demand for consistency.

Against this background, let's revisit the *inference challenge*: the challenge for antirealists to explain how moral reasoning, unaided by intuition, can lead to progressive moral change. Taken by itself, the phenomenon of consistency reasoning does not suffice to explain such change in non-moral terms. That we should expand our circle of moral concern is not a truth of logic, as Huemer correctly observes, but a moral premise. Instead, how we specify the circle of partners to whom we should extend our concerns, and which issues or agents we single out as demanding consistent moral treatment, is itself driven by moral considerations.

However, an appeal to consistency reasoning does help to explain the directionality of moral change *as such*. Moreover, when paired with other factors that trigger such change – such as the lessons we learn from history – we can also explain why consistency reasoning has been channelled in support of specific judgments. Recall our earlier example of an inference that characteristically accompanies progressive moral change:

- I) *Prior judgment*: We should respect the interests of individuals within our tribe.
- II) *New moral insight*: There is no morally relevant difference between individuals within and outside our tribe.
- III) *Posterior judgment*: We should respect the interests of individuals outside our tribe.

Huemer criticizes antirealists for their inability to explain how moral reasoners arrive at this new moral insight (II). But antirealists can actually come up with a clear account of reasoning-induced moral change without appealing to intuitions about mind-independent moral truths: such change comes about through consistency reasoning.

As said, a demand for consistency does not *necessitate* that moral reasoning leads to greater inclusivism. In fact, it might just as well lead to *less* inclusivism – and historically, there have been plenty of examples of societies moving in this alternative direction too. Additionally, it is not *necessary* that moral reasoners adhere to the metanorm of being consistent. But given our psychological inclination to favour a consistent moral stance, it is perfectly explicable, in antirealist terms, that moral reasoning has often been a trigger for progressive moral change.⁴³ The search for consistency has imposed upon the history of moral debate its own internal dynamic, which provides antirealists with yet a further ingredient to explain the broad trend towards liberalism.

4.6 Why the Antirealist's *Explanans* is Superior

In the previous section I have argued that Huemer's three main challenges for antirealists to explain the historical convergence of moral values can all be met. If this argument has been

⁴³ Of course, in order to provide a full account for this, antirealists also need to specify what makes moral change progressive. For an antirealist defence of moral progress, see, for example, Kitcher 2011.

successful, then Huemer can no longer claim that realists are better placed than antirealists to explain this convergence.

Moreover, in this section I will argue that a historical *explanans* which assumes an antirealist metaethics is in fact superior to Huemer's realist *explanans*. I do so by highlighting specific virtues of the former explanation, as well as shortcomings of the latter. Some of these virtues and shortcomings are closely related, and perhaps reducible to each other. I distinguish between them for purposes of exposition, but their dialectical weight is most evident when considering them in tandem: they show that in accounting for the relevant historical phenomena, antirealists have the better explanation.

4.6.1 Antirealist Explanations Are More Robust

One reason to favour an antirealist explanation is that it can illuminate a broad range of historical *explananda*, including historical episodes of moral divergence and regress, which are more difficult to account for assuming a realist metaethics. As mentioned in **section 4.3**, there are various counterexamples which do not easily fit the historical trend of convergence towards liberal values, such as the extreme racism of Nazi-Germany. Additionally, depending on the historical period under consideration and the examples chosen, we can identify several instances of increasing moral divergence, both within and between populations. Historical explanations which assume antirealism – i.e. which assume that either there are no moral truths, or moral truths are mind-dependent – can accommodate the many twists and turns of our moral history fairly straightforwardly, whereas realist explanations for divergence and regress (e.g. the Nazi's were extremely clouded by bias) typically provide little historical insight (see also **section 4.6.4**), and are sometimes *ad hoc*. If antirealist accounts can explain a broader range of historical cases, and are less dependent on cherry-picking the right examples, then we should favour them on historical grounds.

4.6.2 Antirealist Explanations Have a Better Fit with the Data

Not only are antirealists better able to explain episodes of moral regress and divergence – historical *explananda* that go beyond the trend which Huemer has picked out. Also with regard to some of the episodes of progress and convergence which Huemer highlights, antirealist explanations have a better fit with the data. Consider the abolition of slavery. What made abolitionist campaigns succeed? Assuming rationalist intuitionism, we would expect that the intellectual recognition of the moral fact that slavery is bad was the trigger for change. But in fact the disputed moral status of slavery was recognized well before the abolitionist movement gained momentum. Pleasants (2010) argues, on the basis of detailed engagement with historical sources, that the success of abolitionism required that a competitive alternative to slavery and other forms of coerced labour could be offered. With the emergence of the market economy and the system of wage labour such an alternative arose, and the conditions were put in place to advance strong arguments against slavery. Hence, it was not merely a novel intellectual recognition that triggered moral change; practical barriers to envisioning slavery's dispensability had to be overcome.

These data fit well with an antirealist account, in which moral reasoning – driven by a demand for consistency – is socially embedded and takes place against various economic and technological constraints. What typically drives moral progress, antirealists maintain, is not

just the acquisition of a new moral insight but also, at least in part, changing social and material conditions. Of course, realists can also appeal to changing social and material conditions to explain moral change. But on realist accounts such factors can only be of secondary importance; the most important explanatory factor is people's recognition of mind-independent moral truths.

4.6.3 Antirealist Explanations Are More Predictive

If a historical explanation is predictive of actual historical changes, this counts as an important virtue of the explanation. There is at least one respect in which antirealist accounts clearly have this explanatory virtue, and there is at least one respect in which realist accounts do not.

Assuming antirealism, an important predictor of shifts in moral value is the variety of attitudes taken into account in moral discussion. Antirealists typically think that novel moral insights come about through moral dialogue; moral progress can be made by examining the merits and weaknesses of the arguments put forward. In turn, the variety of people participating in this dialogue will influence which voices are heard and which arguments are put forward. When the variety of attitudes taken into account in moral discussion increases or diminishes greatly, we should expect that this will typically result in the endorsement of a very different set of moral values. We should keep in mind that greater inclusion does not invariably lead to greater consensus, but antirealists are not committed to thinking otherwise: they can give a unified explanation of both convergence and divergence. What matters is that, assuming antirealism, changing levels of inclusivity in moral discussion will typically coincide with changes in society's prevailing moral stance. This is indeed what we find when looking at historical data. Especially with regard to the recent value shifts Huemer alludes to – e.g. changing stances towards slavery, racism and sexism – the inclusion of previous out-groups in moral discussion provides a clear predictor of moral change (Anderson 2016).

Assuming realism, and assuming – as Huemer seems to do – that discovering moral truths is akin to discovering non-obvious scientific or mathematical truths, we would expect to find, throughout the course of history, small groups of moral experts who defend and debate non-obvious moral claims. But this is not what we actually find. The historical protagonists of moral change were not typically individuals solving moral conundrums, but interest groups pushing to be given their due. Admittedly, we occasionally do find moral protagonists who, on the basis of a reasoning process, arrive at moral views that a majority of people finds too radical to endorse. But antirealists are well placed to account for this in terms of consistency reasoning. As Campbell and Kumar (2012, p. 288) observe, 'the motive to be consistent can keep one from conforming to the group, as when one bravely calls attention to and seeks to resolve an inconsistency that others in one's community refuse to face'. Rather than having a superior capacity to discern mind-independent moral truths, antirealists may argue that the distinctive quality of moral reformers is their capacity to apply consistency norms in novel contexts, thereby taking moral discussions in new directions.

4.6.4 Realist Explanations Are Insufficiently Detailed

As the case of slavery suggests, antirealist explanations are often richer in detail than realist explanations. With concern to historical explanations, such richness of detail typically counts as an explanatory asset. Historical explanations that build on a realist metaethics, by contrast,

have a tendency to remain very general, and contribute less to our historical understanding. Consider the question of why slavery was abolished. Following Huemer's account, what explains its abolition is first and foremost the fact that slavery is unjust. But in and by itself this is not very illuminating, as becomes apparent when we further inquire about the details of slavery's abolition. Why was the unjustness of slavery only generally recognized during the 19th century? Why then, and not before? This is all the more surprising given that arguments against slavery date from much earlier periods. If these earlier arguments were fuelled by the recognition of a mind-independent moral truth, why did they not persuade people? The antirealist offers a more articulate hypothesis here, by underscoring the importance of changing socio-economic conditions as a precondition for moral change.

4.6.5 Realist Explanations Are Incomplete

Another weakness of realist explanations stems from the vagueness of a realist epistemology. How do we acquire moral knowledge, assuming realism? As noted, rationalist intuitionists argue that acquisition of such knowledge is typically culturally induced but that the ultimate sources of moral knowledge are the intuitions of moral protagonists, who have recognized mind-independent moral truths. This explanation leaves unclear, however, how these mind-independent truths are intuited in the first place and how other people know that the intuitions of moral reformers are reliable. To elucidate this seems necessary to shed sufficient light on the *explanandum*; in the absence of such elucidation, the realist's explanation is incomplete.

Perhaps such elucidation is forthcoming, but the prospect of realists to specify how the intuitions of moral reformers allow them to gain reliable access to mind-independent moral truths is fraught with difficulties. Huemer, for instance, maintains that apart from being mind-independent, moral truths are causally inert: they do not causally influence the intuitions of moral reformers. But if moral facts exist independently of the intuitions of moral reformers and do not causally influence these intuitions, then it is unclear how these reformers could have reliably tracked moral truths – a problem reminiscent of the Benacerraf-Field challenge for mathematical Platonism.

Note that this is familiar problem for realism, and not specific to the argument I have presented. But it is related to the present argument, since absent a clear account of how we are able to acquire moral knowledge, it remains equally unclear how we have been able to acquire such knowledge over the course of history.

4.6.6 Realist Explanations Are Unclear

To explain the historical shift towards liberalism, antirealists can appeal to cultural changes (e.g. the changing variety of people involved in moral discussion), material changes (e.g. the consequences of the Industrial Revolution) and changes in moral attitudes (e.g. recognizing that one group has been morally privileged, and arguing, by way of consistency reasoning, that other groups are entitled to the same privileges), as well as various other factors. Realists may appeal to the same factors, but must additionally appeal to a recognition of mind-independent moral truths. But if the foregoing criticisms have been along the right lines, then such an appeal does not help the explanation: it makes it vaguer than explanations in which mind-independent truths do not figure. For instance, if we lack a comprehensive account of how mind-independent moral truths are grasped (4.6.5), then the resultant explanation is bound to be more

obscure than an explanation that does not refer to mind-independent truths. Hence, unless the realist can overcome these previous criticisms, the antirealist's explanation should also be preferred for reasons of clarity.

4.6.7 Realist Explanations Contain Redundancies

Realist accounts of historical moral change invoke an explanatory factor which antirealist accounts do not invoke: over the course of history, we have tracked mind-independent moral truths. Moreover, realists hold that this is a necessary component of a comprehensive historical explanation. By implication, realists maintain that historical explanations that proceed against the background of moral antirealism are incomplete.

In **section 4.5**, however, I have argued that this contention is mistaken. I have shown that antirealists have ample resources to account for the apparent coincidence of our global value shift, and for the role that reasoning plays in progressive moral change. If my argument has been along the right lines, then the antirealist's explanation is by no means incomplete; it can shed light on the *explanandum* in a comprehensive way. Moreover, if the antirealist has indeed provided a comprehensive explanation, then arguably the realist's explanation contains a redundancy. Parsimony considerations come into play: if all other things are equal, apart from the fact that the realist's explanation contains a redundancy whereas the antirealist's explanation does not, then the latter should be preferred.⁴⁴

Realists may want to object to the relevance of parsimony and argue that in deciding which philosophical view to favour, parsimony considerations are irrelevant and question-begging (cf. FitzPatrick 2015). However, we should keep in mind that the present debate does not turn on the question of which metaethical view we should favour *all things considered*, but on the question of what is the better *historical explanation*: one that proceeds against the background of realism, or antirealism. With regard to competing historical explanations, parsimony considerations are certainly relevant: if mind-independent truths need not be appealed to in order to provide an adequate historical explanation, then historians should not appeal to them.

Alternatively, realists may want to reason as follows: one can only submit that the best historical explanation does not appeal to the existence of mind-independent moral truths if one has *already* rejected moral realism. If one has not, then mind-independent moral truths may still be regarded as explanatorily indispensable. Hence, we arrive at a stalemate in the debate: whether one thinks that an appeal to mind-independent moral truths is required for our best explanation depends on one's underlying metaethical commitments (cf. FitzPatrick 2014, p. 248).

But in fact the present dialectic is quite different. I have discussed the realist's best arguments, as presented by Huemer (2016), for thinking that mind-independent truths are

⁴⁴ Some realists may want to deny that other things are equal, and argue instead that they substitute the antirealist's historical explanation with a completely different explanation. If the explanations of realists and antirealists turn out to be different in many respects, then parsimony considerations are indeed beside the point. They only come into play if, apart from its appeal to mind-independent truths, the realist's explanation is (almost) identical to that of the antirealist.

indispensable in explaining historical moral change, and found them wanting. At this point in the dialectic, the burden of argument is with realists; they need to show that my criticisms are mistaken and that mind-independent moral truths are explanatorily indispensable after all. Unless and until such an argument is given, we may presume that the arguments I have provided are along the right lines and that mind-independent moral truths are indeed redundant in explaining historical moral change.

4.6.8 Realist Explanations Have a Poor Fit with Other Disciplines

An additional drawback of Huemer's realist explanation is that it does not easily fit with our background knowledge from other disciplines, such as psychology and cognitive science. For instance, there is no indication from these fields of research that moral reformers engage in a process of intuiting mind-independent facts. By contrast, cognitive scientists do find that inconsistency is typically a trigger for belief change, which lends support to the antirealist's view of reasoning-induced moral change. Realists may appeal to the same phenomenon of moral consistency reasoning, but they *additionally* need to appeal to the existence of mind-independent moral facts – an appeal that finds no support in findings from other disciplines. If antirealist explanations can more easily be unified with our scientific background knowledge than realist explanations, this provides a further reason to favour the former over the latter.

4.7 Conclusion

Huemer (2016) recognizes the relevance of historical data for assessing the adequacy of metaethical theories, and appropriately so. But contrary to what Huemer suggests, when taking into account all of the evidence, these data provide better support for an antirealist than a realist metaethics. Huemer's criticisms of antirealist explanations are unconvincing; moreover, antirealist explanations have several virtues compared to realist accounts.

That moral antirealism has a better fit with our historical background knowledge than moral realism should not be taken to conclusively *vindicate* the former, or to *debunk* the latter. After all, more considerations are relevant for evaluating the adequacy of metaethical theories than their compatibility with our best historical explanations. But the arguments given, if successful, do raise the overall plausibility of antirealism compared with realism, as they suggest that antirealists are better positioned to shed light on an important historical *explanandum*. Other things being equal, we should prefer some version of moral antirealism over Huemer's realist intuitionism on historical grounds.

References

- Anderson, E. 2016. The Social Epistemology of Morality: Learning from the Forgotten History of the Abolition of Slavery. In: M.S. Brady and M. Fricke, eds. *Contribution to The Epistemic Life of Groups*. Oxford: Oxford University Press, 75-94.
- Buchanan, A. and Powell, R. 2015. The Limits of Evolutionary Explanations of Morality and Their Implications for Moral Progress. *Ethics* 126: 37-67.
- Campbell, R. and Kumar, V. 2012. Moral Reasoning on the Ground. *Ethics* 122: 273-312.

- Copp, D. 2008. Darwinian Skepticism about Moral Realism. *Philosophical Issues* 18: 186-206.
- Dewey, J. and Tufts, J.H. 1908. *Ethics*. New York: Henry Holt.
- Doris, J. and Plakias, A. 2008. How to Argue about Disagreement: Evaluative Diversity and Moral Realism. In: W. Sinnott-Armstrong, ed. *Moral psychology, vol. 2*. Cambridge: MIT Press, 303-331.
- Enoch, D. 2009. How is Moral Disagreement a Problem for Realism? *Journal of Ethics* 13: 15-50.
- Fitzpatrick, S. 2014. Moral Realism, Moral Disagreement, and Moral Psychology. *Philosophical Papers* 43: 161-190.
- FitzPatrick, W. 2012. Review of Philip Kitcher. The Ethical Project. *Ethics* 123: 167-174.
- FitzPatrick, W. 2014. Why There Is no Darwinian Dilemma for Ethical Naturalism. In: M. Bergmann and P. Kain, eds. *Challenges to Moral and Religious Belief: Disagreement and Evolution*. Oxford: Oxford University Press, 237-255.
- FitzPatrick, W. 2015. Debunking Evolutionary Debunking of Ethical Realism. *Philosophical Studies* 172: 883-904.
- Gibbard, A. 1990. *Wise Choices, Apt Feelings*. Oxford: Oxford University Press.
- Haidt, J. and Bjorklund, F. 2008. Social Intuitionists Answer Six Questions about Moral Psychology. In: W. Sinnott-Armstrong, ed. *Moral psychology, vol. 2*. Cambridge: MIT Press.
- Huemer, M. 2008. Revisionary Intuitionism. *Social Philosophy and Policy* 25: 368-392.
- Huemer, M. 2016. A Liberal Realist Answer to Debunking Skeptics: The Empirical Case for Realism. *Philosophical Studies* 173: 1983-2010.
- Joyce, R. 2006. *The Evolution of Morality*. Cambridge: MIT Press.
- Kitcher, P. 2011. *The Ethical Project*. Cambridge: Harvard University Press.
- Kumar, V. Forthcoming. The Ethical Significance of Cognitive Science. In: S.-J. Leslie and S. Cullen, eds. *Current Controversies in Philosophy of Cognitive Science*. New York: Routledge.
- Mackie, J.L. 1977. *Ethics: Inventing Right and Wrong*. New York: Penguin.
- Pinker, S. 2011. *The Better Angels of Our Nature: Why Violence Has Declined*. New York: Viking.
- Pleasants, N. 2010. Moral Argument Is Not Enough: The Persistence of Slavery and the Emergence of Abolition. *Philosophical Topics* 38: 159-180.
- Schwitzgebel, E. and Cusman, F. 2012. Expertise in Moral Reasoning? Order Effects on Moral Judgment in Professional Philosophers and Non-Philosophers. *Mind & Language* 27: 135-153.
- Shafer-Landau, R. 2003. *Moral Realism: A Defense*. Oxford: Clarendon Press.
- Street, S. 2006. A Darwinian Dilemma for Realist Theories of Value. *Philosophical Studies* 127: 109-166.
- Street, S. 2016. Objectivity and Truth: You'd Better Rethink It. In: R. Shafer-Landau, ed. *Oxford Studies in Metaethics, vol. 11*. Oxford: Clarendon Press, 293-334.

5

Chapter 5

Two Accounts of Moral Objectivity: From Attitude-Independence to Standpoint-Invariance

Published in *Ethical Theory & Moral Practice* 20, 2017, pp. 763-780.
Special issue on 'Objectivity in Ethics'.

Keywords

Moral objectivity; Humean constructivism; ideally coherent eccentrics; moral relativism; moral metaphysics

Abstract

How should we understand the notion of moral objectivity? Metaethical positions that vindicate morality's objective appearance are often associated with moral realism. On a realist construal, moral objectivity is understood in terms of mind-, stance- or attitude-independence. But realism is not the only game in town for moral objectivists. On an antirealist construal, morality's objective features are understood *in virtue of* our attitudes. In this chapter I develop a particular version of this antirealist construal of moral objectivity and make its metaphysical commitments explicit. I do so by building on Sharon Street's version of Humean constructivism. Instead of the realist notion of *stance-independence*, the antirealist account of moral objectivity that I articulate centres on the notion of *standpoint-invariance*. While constructivists have been criticized for compromising on the issue of moral objectivity, I make a preliminary case for the thesis that, armed with the notion of standpoint-invariance, constructivists have resources to vindicate an account of objectivity at just the right level of strength, given the commitments of ordinary moral thought and practice. In support of this thesis I highlight recent experimental findings about folk moral objectivism. Empirical observations about the nature of moral discourse have traditionally been taken to give *prima facie* support to moral realism. I argue, by contrast, that from what we can tell from our current experimental understanding, antirealists can capture the commitments of ordinary discourse at least as well as realists can.

'The task of making more exact a vague or not quite exact concept used in everyday life or in an earlier stage of scientific or logical development, or rather of replacing it by a newly constructed, more exact concept, belongs among the most important tasks of logical analysis and logical construction. We call this the task of explicating, or giving an explication for, the earlier concept.'

– Rudolf Carnap (1956)

5.1 Introduction

There is a broad consensus among philosophers that morality has certain objective features and that it is a metaethical *desideratum* to account for this. For example, many judgments in ordinary moral discourse purport to be objective; many metaethicists, in turn, regard this as a datum that their theories should accommodate. *How* to accommodate morality's objective aspirations, however, is an issue of ongoing debate, fuelled at least in part by controversy over how the notion of 'moral objectivity' should be understood.⁴⁵ Objectivity has various connotations; its precise meaning may differ across contexts and different domains. Nichols (2014, p. 734) contends that '[g]iving a precise characterization of objectivity is itself a major philosophical endeavour.' Timmons (2010, p. 544) even characterizes the issue of moral objectivity as 'the central topic of metaethics.'

On one prominent account, moral objectivity is understood in terms of mind-independence, stance-independence, or – my term of choice in this chapter – *attitude-independence*. Call this a realist understanding of moral objectivity (e.g. Shafer-Landau 2003).⁴⁶ On a realist understanding, to say that a moral judgment purports to be objective is to say that this judgment purports to have an attitude-independent truth-value (Finlay 2007, p. 822). There are objective facts about what is right and wrong, good and bad, and these facts do not depend on anyone's attitudes. Consider the claim that the Earth revolves around the sun: this claim is true, it states a fact and this truth or fact is fully independent of what any agent thinks or feels about it. Similarly, moral realists maintain that moral truths or facts are fully independent of the attitudes of any agent.

A recurring criticism against a realist understanding of objectivity, which has been pushed in the domains of both ethics and mathematics (Clarke-Doane 2014), is that it raises epistemic quandaries. For example, Sharon Street (2006, 2016) has forcefully argued that if moral truths or facts are fully attitude-independent, then we have no guarantee that our moral judgments reliably track these truths or facts. Other critics, such as John Mackie (1977), argue

⁴⁵ Another potential source of controversy is the vagueness of 'morality', thus characterized. I give a more precise characterization of morality's objective features in **sections 5.2 and 5.3**.

⁴⁶ As noted in the introduction of this dissertation, although moral realism is a metaethical term of art, philosophers tend to characterize it in slightly different ways. As a result, the variety of metaethical claims labelled 'realist' is rather diverse (see Finlay 2007 for an overview). For a helpful recent characterization of moral realism, see Leibowitz and Sinclair eds. (2016), especially pp. 4-5, 60-61, 126-127. The realist view of moral objectivity I discuss in this chapter, according to which moral objectivity is understood in terms of attitude-independence, sometimes goes by the name *robust realism*.

that a realist view is not only epistemically but also metaphysically problematic.⁴⁷ However, whatever difficulties moral realism might face, there seems to be a fair amount of consensus among philosophers that at least concerning the aim of capturing moral appearances, it does a good job. For instance, the realist has resources to argue that in cases of moral disagreement, at least one of the disagreeing parties must be mistaken, since his or her judgment does not correspond with the attitude-independent moral facts or truths. *Prima facie*, this seems to accord with people's intuitions. Consider an imaginary agent who values an action that many people would regard as decidedly immoral – say an agent who values torturing innocent human beings for fun. Presumably, many people would want to claim that this agent is morally wicked, and objectively so. Torturing innocent people for fun is morally wrong, no matter what this agent thinks or feels about it, or so our intuition presumably tells us. Since realists take objective moral truths to be attitude-independent, they can procure this intuition in a straightforward manner.

Many metaethicists agree that realist theories are good at accommodating morality's 'objective pretensions', and that in this respect they have an advantage over rival theories (e.g. Finlay 2007, p. 844). Some philosophers implicitly treat a realist understanding as the default understanding of moral objectivity (e.g. Mackie 1977; Enoch 2014), or even equate moral realism and moral objectivism (Kahane 2011), thus implying that all moral objectivists are moral realists.⁴⁸ This is an uncharitable move, however, since there has been an extensive, predominantly Kantian tradition of providing alternative accounts of moral objectivity which do not centre on the notion of attitude-independence. Call these *antirealist* accounts of moral objectivity. What unites antirealist accounts of moral objectivity, as I shall understand them, is that they understand morality's objective features *in virtue* of our attitudes, or at least partly in virtue thereof. For example, an antirealist might argue that moral judgments are objective if they result from a procedure that elicits a dialogical understanding (Habermas 1995). Or she might maintain that moral judgments are objective if they are independent of the whims of anyone's *particular* attitudes: 'Objectivity (...) is not so much a "view from nowhere," but a "view of no one in particular"' (Sen 1993, p. 129).

My primary aim in this chapter is to develop a specific version of this antirealist view of moral objectivity and to make its metaphysical commitments explicit. Extant antirealist accounts are underdeveloped in this respect. For example, defining objectivity as a view of no one in particular is insufficiently informative: the definition fails to provide a positive answer to the question *in virtue of what* moral judgments are objective. Building on the metaethical position that Sharon Street (2008b, 2009, 2010, 2012) has developed under the header Humean constructivism (abbreviated in this chapter as HC), I seek to give antirealist theories of objectivity a more solid theoretical foundation. I expand on Street's constructivist account by framing moral objectivity in terms of standpoint-invariance: moral judgments that purport to be objective purport to have the property of being standpoint-invariant. Distinguishing between

⁴⁷ Error theorists maintain that a substantial subset of moral judgments (e.g. all affirmative moral judgments) are systematically false or systematically unjustified, since their truth-makers – objective properties that are at the same time prescriptive – are too queer to be instantiated (cf. Mackie 1977). In making this claim, error theorists assume that ordinary discourse is committed to a realist notion of objectivity – i.e. objectivity understood in terms of attitude-independence.

⁴⁸ Kahane 2011; see esp. p. 121, fn. 1 and p. 123, fn. 39.

realist and antirealist accounts of moral objectivity, then, comes down to distinguishing between an account of objectivity in terms of attitude-independence and an account of objectivity in terms of standpoint-invariance.

Street herself has expressed some doubts whether Humean constructivists can vindicate a sufficiently strong form of objectivity.⁴⁹ Indeed, there seems to be a consensus among many metaethicists that realists have a leg up on antirealists in accommodating morality's objective pretensions. The argument in support of this consensus goes roughly as follows. Realists advance a strong notion of objectivity, and this strong notion appears to capture the commitments of our ordinary moral discourse and thinking. If we assume that a metaethical account of moral objectivity scores 'plausibility points' (Enoch 2011) if it accords with ordinary discourse and thinking about moral objectivity, this finding counts in favour of realism (I discuss this assumption in **section 5.2**). Antirealist accounts, by contrast, lack the strength of their realist counterpart. *Prima facie*, antirealist objectivity merely seems to boil down to intersubjective agreement. The realist's strong notion of attitude-independence captures the prescriptive force and inescapability that seem characteristic of objective judgments in the moral domain; the antirealist view seems to capture neither. Hence, a realist account of objectivity accords better with the intuitions of participants in moral discourse than the weaker account of the antirealist.

Contrary to this claim, I shall advance several reasons for thinking that the antirealist is at least as well positioned as the realist in accounting for folk intuitions and in capturing ordinary discourse.⁵⁰ First, building on recent experimental findings, I suggest that ordinary discourse and thinking are not committed to the realist's strong sense of objectivity but rather to a graded form objectivism. Second, elaborating on Street's Humean constructivism, I argue that the antirealist's account of objectivity is much stronger than appears at first glance, and that the antirealist may well be able to procure the intuition that moral judgments have an inescapable authority – at least, to the extent that experimental research indicates that it *should* procure this intuition. Hence, instead of compromising on the issue of moral objectivity, I argue that the antirealist is well-equipped to *vindicate* an account of objectivity at just the right level of strength, apt for the moral domain.

The chapter proceeds as follows. In **section 5.2** I begin by clarifying how findings about ordinary moral discourse might be relevant for metaethical theorizing. In **section 5.3** I give a brief overview of recent experimental research on folk moral objectivism, which allows us to refine our *explanandum* – i.e. morality's objective features. In **section 5.4** I outline the metaethical position that Street has developed under the header Humean constructivism. In

⁴⁹ Street has touched upon the topic of objectivity in various places (e.g. 2010, p. 380; 2012, p. 41; 2016a, pp. 42-43), expressing slightly different views. Her general position, as I understand it, is that whereas so-called Kantian versions of constructivism vindicate a strong form of moral objectivity, the Humean version of constructivism that she defends *compromises* on moral objectivity (Street 2010, 2012; presentation at the workshop *Objectivity in Ethics* 2016).

⁵⁰ Naturally, there may be other respects in which realists and antirealists could be better or worse positioned; which position is the most plausible all things considered will have to be settled on the basis of comprehensive grounds. Traditionally, it has been held that whereas moral realists are better at capturing moral appearances, antirealists are better at reconciling morality with our understanding of the mind and world (Finlay 2007, p. 844). In this chapter, I seek to cast doubt on the former claim; I have discussed the latter claim in the previous chapters.

section 5.5 I use Street's position as the basis from which to articulate an antirealist account of moral objectivity: objectivity as standpoint-invariance. I then go on to address three objections against HC that have been raised in the literature – the relativity objection (**section 5.6**), the alignment objection (**section 5.7**) and the revision objection (**section 5.8**) – and outline the resources that the constructivist has to counter these objections. In the concluding section, **section 5.9**, I contrast the antirealist account of moral objectivity with its realist counterpart and make a preliminary case for the thesis that, even if a realist understanding of objectivity may be appropriate in other domains, in the moral domain an antirealist account of objectivity has a better fit.

5.2 The Metaethical Relevance of Ordinary Discourse

Many metaethicists regard it as a *desideratum* to account for morality's objective appearance. In this section and the next one, I outline various findings about ordinary moral thought and talk that have been cited in support of morality's apparent objectivism. But first we should be clear about the metaethical relevance of these findings. Why should the commitments of ordinary moral discourse carry any metaethical weight?

It can be tricky to take people's ordinary thought and talk about moral objectivity at face value. After all, folk beliefs might be mistaken and folk discourse might be deficient. Instead of mirroring people's ordinary beliefs about moral objectivity, our best metaethical theory might force us to revise these beliefs; instead of vindicating the commitments of ordinary discourse, our best metaethical theories might tell us to debunk them. Philosophers may hold views about the metaphysics of moral objectivity that are incompatible with laypeople's intuitions, and it may well be that philosophers are correct in doing so and that the laypeople are mistaken.

How moral agents themselves conceive their own discourse and what is the best philosophical understanding of that discourse, then, are two distinct issues. But even though these issues are logically distinct, they are not unrelated. Ordinary moral thought and talk are not beyond the purview of metaethical analysis. On the contrary: these are among the very phenomena that metaethicists seek to shed theoretical light upon. Whether or not ordinary discourse and laypeople's beliefs about moral objectivism are ultimately vindicated, it is the metaethicist's task to accommodate them at the level of theory. As Blackburn (1984, p. 180) notes, metaethicists aspire to explain 'why our discourse has the shape it does'. Facts and findings about ordinary discourse constitute metaethical data; in turn, metaethicists seek to provide an adequate theoretical framework to account for them. In the final analysis, the plausibility of a metaethicist's framework will depend on how well this framework fits with our best theoretical understanding of the moral domain, all things considered – and findings about actual moral discourse can influence how many plausibility points a metaethical theory scores in this respect. Hence, while ordinary thought and discourse are not infallible, this does not preclude their metaethical relevance.

Indeed, metaethicists have long regarded it as a theoretical *desideratum* to accommodate the features of ordinary discourse. Consider some of the findings that have traditionally been cited in support of morality's apparent objectivism. These include the observation that moral judgments are often stated in the declarative mode, just like statements of fact (Bagnoli 2015); the observation that just like in theoretical deliberation, in practical

deliberation we tend to assume that there are correct and mistaken answers about what we have reason to do (Blackburn 1984; Shafer-Landau 2003); and the observation that in terms of phenomenology, moral judgments seem to be more objective than judgments of taste, and almost or equally objective as factual judgments (Enoch 2014). These observations all take input from colloquial moral intuitions, moral practices and moral language. They are the findings that give rise to the metaethical *explanandum* of accounting for morality's objective appearance.

Lastly, note that even though the commitments of ordinary discourse need not be taken at face value, it takes effort to discard them. All other things being equal, a metaethicist who claims that large parts of ordinary discourse are in need of revision faces a larger burden of proof than a metaethicist who vindicates moral discourse. If there is a good fit between some metaethical theory and some moral practice, then, unless and until arguments to the contrary have been given, this fit serves to corroborate the theory and to justify the practice. For example, if many people regard moral judgments as objective and a metaethicist can vindicate their objectivity, then there is a good fit between practice and theory. By contrast, if many people regard moral judgments as objective but a metaethical theory does not justify morality's purported objectivity, then this theory faces the more daunting task of accounting for the fact that many people's intuitions about moral objectivity are mistaken.

5.3 Refining the *Explanandum*

I started the previous section by highlighting some of the findings that have traditionally been cited in support of the claim that metaethicists have to account for morality's objective appearance. Only over the last decade or so, however, have moral psychologists and experimental philosophers begun to scrutinize the objectivist commitments of ordinary moral speakers in a rigorous, experimental fashion. This has resulted in a fairly extensive and ever-growing body of studies, with subjects that have been tested across different cultures and age groups (e.g. Nichols 2004; Goodwin and Darley 2008; Sarkissian et al. 2011; Goodwin and Darley 2012; Wright et al. 2013; Beebe 2014; Khoo and Knobe 2018). While these studies give some support to the traditional consensus that moral discourse purports to be objective, they provide a more fine-grained understanding of the sense in which it does. In this section I highlight some of the findings that the new wave of experiments has brought to light and show how they refine the traditional consensus. Rather than being straightforwardly objectivist, I argue, the new experiments suggest that moral discourse is committed to a *graded* sense of objectivism.

Experimenters typically test the objectivist commitments of ordinary respondents by first presenting them with a range of personal, conventional, moral and scientific judgments. In some studies, respondents are asked to classify these judgments in one of these categories themselves. For example, a judgment predominantly classified as moral is that 'consciously discriminating against someone on the basis of race is unacceptable' (Wright et al. 2013, p. 340). After being presented with a list of such statements, respondents are typically asked whether in cases of disagreement, they think that one of the disagreeing parties is mistaken. If the respondents think that it is possible for both of two disagreeing parties to be correct about an issue, this is regarded as a denial of objectivity. If the respondents think that at least one of the disagreeing parties must be mistaken, this is regarded as an ascription of objectivity.

There is much more to say about the experimental set-up of these studies and their potential pitfalls; I will discuss these more thoroughly in **chapter 6**. For now, let it suffice to say that while I do not think that these experiments are unproblematic in all respects, their results constitute the best data about folk moral objectivism currently available. While these results corroborate some prior ideas about the objectivist commitments of ordinary discourse, such as the idea that moral judgments purport to be more objective than judgments of taste, they also indicate that there are complexities about folk objectivism that metaethicists have heretofore overlooked. I highlight three of them.

First, the *degree of objectivity* attributed to moral judgments varies substantially depending on the moral issue in question. For example, in Wright et al.'s (2013) study a majority of subjects self-classified the (un)acceptability of first trimester abortion as a *moral* issue, but it was classified as an *objective* issue only in 15% of the cases. The (un)acceptability of robbing a bank to pay for a holiday, on the other hand, which was also self-classified as a moral issue, was regarded as objective in 60% to 85% of the cases. This diversity in attributions of moral objectivity has also been found by other studies, suggesting that we might be mistaken to think that all moral judgments purport to be equally objective. Instead, experimental research suggests that in ordinary discourse, judgments on some moral issues are regarded as more objective than judgments on other moral issues (Goodwin and Darley 2012).

A second complication concerns the *scope* of appeals to objectivity. Philosophers have traditionally held that moral claims purport to be inescapable, in the sense that they are binding on all agents. Again, while empirical research supports this intuition at a general level, more fine-grained analyses suggest a more diverse picture. Asked about the scope of application of moral judgments, experimentally tested subjects typically exhibit a *graded* commitment to objectivism: while they have absolutist intuitions when moral judgments concern individuals from their own culture, they express increasingly relativist intuitions when these judgments concern individuals from increasingly different cultures or ways of life, such as aliens from outer space (Sarkissian et al. 2011). People tend to be cultural absolutists but galactic relativists.

Expanding on the study by Sarkissian et al., Khoo and Knobe (2018) have recently highlighted a third complication. They show that there are cases of moral conflict in which people are inclined to say both that the two speakers disagree and that it is not the case that at least one of them must be saying something incorrect. To reflect the moral semantics of ordinary discourse, they propose, we should distinguish between two types of disagreement: a type of disagreement in which the disagreeing parties are thought to occupy positions that exclude each other, and a type of disagreement where the positions of the disagreeing parties are non-exclusionary. In at least some cases of moral disagreement, the study suggests, ordinary respondents regard the positions of disagreeing parties as non-exclusionary.

To sum up, as the current state of research on folk objectivism indicates, moral discourse is committed to a *graded* sense of objectivity, in at least three ways. First, some moral judgments are held to be very objective, others less so. Second, even for those types of judgments that are held to be very objective, people do not think that their moral authority applies to all intelligent creatures, such as non-human aliens. Third, some moral disagreements may be non-exclusionary, in the sense that none of the disagreeing parties are mistaken. The

task for metaethicists who want to shed light on ordinary moral discourse, then, is to provide an *explanans* for this commitment to *graded objectivism*.

5.4 Humean Constructivism

In the following sections I develop an antirealist account of objectivity that is particularly well suited to making sense of this commitment to *graded objectivism*. I do so by departing from Sharon Street's version of Humean constructivism. Following Street, in this section I present metaethical constructivism as a theory about normative practical reasons. In the next section I return to reasons that are specific to the moral domain and go beyond Street's account in articulating an antirealist account of moral objectivity.

Normative practical reasons are reasons to pursue, value or favour something. For example, to make the normative judgment '*X* is a reason for *A* to *Y*' can be understood as 'for agent *A*, *X* counts in favour of *Y*'. Normative reasons should be distinguished from psychological reasons (also known as motivating reasons): *X* can be a normative reason for *A* to *Y*, whether or not *A* is psychologically inclined to *Y*.⁵¹ But while normative reasons and psychological reasons should not be equated, the constructivist does emphasize that normative reasons are always reasons *for someone*. The judgment '*X* is a reason to *Y* full stop' is underdetermined: to know whether this judgment holds true we need to know *whose* reason it is. For Johnny, hearing a song by Prince is a reason to turn up the volume; for Jackie, it is not. For Louis, hearing that a dish contains bluefin tuna is a reason not to order it; for Laura, it is not.

What is it for an agent to judge *correctly* that *X* is a normative reason to *Y*? In other words, what are the truth-makers of normative judgments? On a straightforward subjectivist account, if *A* thinks that *X* is a reason for her to *Y*, her thinking makes it so. Metaethical constructivism is *not* straightforwardly subjectivist. According to the constructivist, whether *X* provides agent *A* with a reason to *Y* depends on whether this judgment would withstand scrutiny from agent *A*'s full set of normative judgments, in combination with the non-normative facts. Call the full set of normative judgments that an agent makes the agent's *evaluative standpoint*.⁵² What follows from this standpoint reflects the idealized commitments of a moral agent – i.e. what her commitments will be, once they are coherent and informed. The constructivist maintains that the truth of any normative judgment is relative to the evaluative standpoint of the agent who makes the judgment, combined with the non-normative facts. Hence, it's not Johnny's *actual mental state* but rather Johnny's *evaluative standpoint* that sets the standards of correctness for his normative judgment.

What is it for an agent to *mistakenly* judge that *X* provides her with a reason to *Y*? To commit a normative mistake is to make a judgment that *does not* withstand scrutiny from an agent's evaluative standpoint, even if the agent, in making this judgment, *thinks* it does. There are several ways in which normative mistakes can come about. For example, an agent might be ignorant or misinformed about the relevant non-normative facts, she might not have given

⁵¹ Following the current fashion in metaethics, the concept of a 'normative reason' may be regarded as a primitive concept of normative analysis (cf. Scanlon 1998, p. 19).

⁵² Adopting Street's terminology, we may alternatively call this an agent's practical point of view, an agent's practical standpoint, or her full set of evaluative attitudes. Here and in what follows, I stick with the term evaluative standpoint.

her judgment sufficient reflection and fails to see the full consequences of her existing commitments or be incoherent in any other sort of way (cf. Street 2017). Let's consider these mistakes in some more detail, taking Laura – who thinks that the fact that a dish contains bluefin tuna (X) is a reason to order it (Y) – as our example.

5.4.1 Factual Misinformation

First, let's assume, plausibly, that it follows from Laura's standpoint that she has reason to be informed about the relevant non-normative facts. That is to say, as a self-governed epistemic norm, Laura cares about being correctly informed by the facts that are relevant to her normative judgment: if her normative judgment at t_1 rests on poor or false information, then at t_2 she might retroactively realize that her judgment was mistaken in terms of her own commitments. For example, at t_1 Laura doesn't know that bluefin tuna is an endangered species (Z). But once she becomes aware of this, she readily submits that she has reason not to order it. In other words, once at t_2 Laura becomes aware that Z and realizes that for her, X in combination with Z entails *not* Y . Since she judged herself to have reason to be informed about the relevant non-normative facts when she made her original judgment at t_1 , she now realizes that, by her own lights, her judgment at t_1 was mistaken.

5.4.2 Incoherence

Second, even if Laura is aware of the relevant non-normative facts, she might still err in her normative judgment. For example, although Laura is aware that bluefin tuna is an endangered species, and although she endorses the judgment that one shouldn't order dishes that contain endangered species, she fails to take this into explicit consideration when placing her order. Hence, while Laura judges that X and Z and thinks that she has normative reason to Y , she isn't alert to the fact that Z gives her instrumental reason to *not* Y . But once she thinks harder about her judgment, she realizes that *not* Y , instead of Y , is what actually follows from her evaluative standpoint when it is stripped of inconsistencies. In making her initial judgment, then, Laura was mistaken.

Given the complexity involved in making normative judgments and the myriad non-normative facts that might be relevant to a judgment there are numerous ways in which an agent might err by her own lights (cf. Flanagan et al. 2014). Indeed, the Humean constructivist maintains that human agents make such mistakes all the time. Human beings are not ideally coherent; we rarely make normative judgments by weighing all the relevant options, reflecting on all the relevant norms and gathering all the relevant non-normative facts. Instead we apply shortcuts and rules of thumb and make decisions under conditions of less than full information, and as a result we regularly make mistakes in terms of our own evaluative standpoint.

To see just how broad this range of potential normative mistakes is, consider mistakes that an agent could make due to failures of imagination. Suppose that it follows from an agent's standpoint that she has reason to be imaginative about the normative judgments that she could make. That is to say, the agent values being the sort of person who does not necessarily act upon one of the reasons that she is already familiar with. Instead, she values using her

imagination to probe whether she might act upon reasons which she has not heretofore considered. Take the case of Laura. Perhaps Laura has not spent much time reflecting on moral issues; until t2 she has simply taken it for granted that *X* is a reason for her to *Y* (Laura's parents regularly ordered bluefin tuna). But with hindsight, she blames herself for not considering the possibility that *X* might be taken as a reason to *not Y* – she openly admits that she *could* have reasonably considered this, and since it followed from her evaluative standpoint that she had reason to be imaginative, she *should* have done so. Thus, even if up until t2 Laura has not even contemplated the possibility that *X* might be taken as a reason to *not Y*, it might still be the case that she is mistaken to judge at t1 that *X* is a reason for her to *Y*, since it already followed from her evaluative standpoint at t1 that she had reason to be imaginative about the normative judgments that she could make, including the judgment that *X* is a reason to *not Y*.

What follows from an agent's evaluative standpoint may be unveiled by answering the following question: what judgment *would* this agent make if she were ideally coherent, imaginative and informed about the relevant non-normative facts? If the judgment she would make under these conditions coincides with the judgment that she actually makes, then her actual judgment is correct. If being more coherent or informed would lead her to make a different judgment, then her actual judgment is mistaken. Given the many ways in which such mistakes can come about, the Humean constructivist submits that, at least among the human agents in our world, ideal coherence is likely to be a rare trait.

5.5 Objectivity as Standpoint-Invariance

With our brief exposition of HC in place, let's scrutinize which facets of normative objectivity the constructivist can procure. Start with the weakest sense of objectivity: intersubjective agreement. On the constructivist's view, even though normative agreement is not guaranteed *in principle*, it is quite likely that, at least in our world, there will be an extensive overlap between the normative judgments that withstand scrutiny from the standpoints of different agents. After all, all human beings share the vast majority of their evolutionary history; as a result, we have many biologically grounded needs and preferences in common (cf. Street 2006). Cross-culturally endorsed norms typically have an adaptive rationale: their universality can often be explained as a consequence of the fact that they are the kinds of norms that tended to promote the survival and reproductive success of human ancestors. Moreover, human agents also share several social practices and institutions, as well as a reservoir of factual knowledge – and increasingly so, in our globalizing world. As a result, there is a substantial overlap between the normative judgments that different agents endorse.⁵³

⁵³ To appreciate this, it is helpful to think of the vast number of possible normative judgments any agent *could* make (cf. Street 2006, 2008a, 2016). *A priori* there are innumerable judgments that might follow from an evaluative standpoint, from the judgement that infanticide is laudable, to the judgment that plants are more valuable than human beings, to the judgement that the fact that something is purple is a reason to scream at it (Street 2006, p. 133). Relative to this vast set of possible judgments, the variation in the normative judgments that human beings *actually* make is quite limited. For example, there may well be universal agreement on the judgment that seeing a fire is a reason not place your hand in it: this is a normative judgment that we may find so obvious that ordinarily we do not bother to reflect on it explicitly.

Naturally, just as there is pervasive agreement on some normative issues, we should keep in mind that there is also pervasive disagreement on other normative issues. The constructivist does not

However, normative judgments in the moral domain typically make a stronger claim to objectivity than mere intersubjective agreement. Apart from being intersubjectively valid, moral judgments typically purport to have an inescapable authority: they express norms that are ‘thought to give us reasons to act in conformity with them, even if we have no desires or interests that are served by those norms’ (Handfield 2016, p. 62). Or, using Joyce’s (2006) term, moral judgments seem to have ‘practical clout’. Can the constructivist procure the idea that moral judgments are strongly objective, in the sense of having this inescapable authority?

Prima facie this seems to be a difficult task. According to HC, the standards of correctness of moral judgments are standpoint-relative: their correctness depends on whether they withstand scrutiny from the standpoints of individual agents. Therefore, it may appear that HC cannot procure the intuition that moral judgments can have normative force irrespective of what an agent wants. Agent *A* only has reason to *X* if *X* follows from her own standpoint: the ultimate source of a normative judgment’s prescriptive authority, the constructivist maintains, comes from within, not from some external source.

However, recall that what follows from an agent’s evaluative standpoint need not coincide with what the agent *thinks* follows from her standpoint. The constructivist maintains that there are certain standards that govern whether a normative judgment is correct or mistaken, irrespective of whether an agent *thinks* of her judgment as being correct or mistaken. It is not the agent’s actual mental states but the agent’s *evaluative standpoint* that legislates whether or not her normative judgment is correct. As a result, an agent’s actual normative judgments may legitimately be criticized by other agents: agent *B* may be in a position to argue that it follows from agent’s *A* standpoint that she has reason to *X*, even if agent *A* has not yet realized this.

As discussed thus far, the constructivist’s view of what makes normative judgments correct seems rather individualistic. The crux of extending it with an account of objectivity is to broaden the diversity of standpoints from which a moral judgment should withstand scrutiny. In order to hold objectively, a moral judgment should not only be able to withstand argumentative scrutiny from the agent’s own idealized standpoint but also from the evaluative standpoints of other agents – and the more diverse these other standpoints are, the more objective the agent’s judgment may be taken to be. This is what it means for moral claims to hold objectively: such claims should be able to withstand a process of *intersubjective scrutiny* in a diverse community of moral agents. Hence, if an ideally coherent agent makes the moral judgment that *p*, and her judgment withstands scrutiny from diverse angles, then *p* may be regarded as an objective moral truth.

Crucially, on this account of objectivity, the truth or falsity of the moral judgment that *X* is not *independent* of the standpoint of moral agents. Rather, if the truth of *X* is highly objective, then this truth (almost) *invariably* follows from different evaluative standpoints. This invariance is the distinctive feature of an antirealist account of objectivity and sets it apart from its realist counterpart. On the antirealist account, if a moral judgment purports to be objective, it purports to withstand scrutiny from a diverse set of evaluative standpoints. A ‘fully objective’ moral judgment purports to be fully standpoint-invariant: no matter how the standpoint of a

deny this; indeed, she acknowledges that some disagreements may rest on fundamental differences between evaluative standpoints (see also **section 5.6** and the end of **section 5.9**).

moral agent varies, the truth of X should follow from it. But the constructivist can accommodate the finding, outlined in **section 5.3**, that not all moral judgments aspire to make such a strong claim to objectivity. Instead, some moral judgments are held to be more objective than others, and their inescapability is a matter of degree. The antirealist accounts for this *datum* by maintaining that the *degree* of objectivity of a moral judgment is a function of the diversity of evaluative standpoints under which its truth or falsity remains fixed. The more diverse these standpoints are – diverse, that is, in terms of the set of input judgments that constitute the initial standpoint – the more objective the resultant judgment. In this way, the antirealist makes room for a *graded* understanding of moral objectivity, which, we have seen, is precisely the sense of objectivity that a metaethical account *should* procure, as suggested by experimental findings.

Lastly, note that on an antirealist account, although some moral judgments may aspire to be objective in a very strong sense, this aspiration is not unbounded. Conceivably, there will be agents for whom the truth of X *does not* follow from their standpoint, even though many other agents regard X as an objective moral truth. The constructivist concedes that this is a genuine theoretical possibility – and in the most extreme of such cases, she will maintain, it is most sensible to conclude that fully coherent agents who endorse very bizarre normative judgments should not be regarded as *moral* agents. This last point brings us to thought experiments that involve ‘ideally coherent eccentrics’ (ICEs), which we discuss in the next section.

5.6 The Relativity Objection

One objection that has been raised against HC is that the view is too relativist. Borrowing Gibbard’s (1999) example, consider Caligula, who judges that he has reason to torture other people for fun. Let’s stipulate that Caligula is ideally coherent, imaginative and informed about the non-normative facts and that his judgment that he has reason to torture people for fun withstands reflective scrutiny from his evaluative standpoint.⁵⁴ The Humean constructivist is committed to the claim that Caligula is *correct* in making this judgment: after all, to make a normative judgment that withstands scrutiny from one’s own evaluative standpoint is to make a correct normative judgment. But this seems to run counter to common sense. As Bratman (2012, p. 84) notes, it is ‘plausible that our ordinary thinking about reasons for action supports Gibbard’s denial that Caligula has a normative reason to torture even if Caligula’s standpoint wholeheartedly supports torture and does not involve relevant and false non-normative belief’. The worry can be formulated in a different way: our ordinary thinking seems to support the idea that some normative judgments – typically moral judgments – are objectively binding, in the sense that that *all* agents should be responsive to them. HC, on the other hand, makes the truth of moral judgments subject-related. In doing so, the constructivist fails to capture morality’s binding authority.

Note two things about the relativity objection. First, the objection is fostered by observations about ordinary moral thinking: since our colloquial intuitions support the idea that

⁵⁴ Alternatively, we might say that Caligula has *most* reason to torture other people for fun, or that he has reason to torture others for fun *all things considered*. From a constructivist viewpoint, however, these additions are not strictly necessary; they are implied by the stipulation that this judgment withstands scrutiny from Caligula’s evaluative standpoint.

eccentric creatures like Caligula are making a moral mistake, a plausible metaethical theory should procure this idea too. Second, note that the supposed commitment of our ordinary thinking is quite strong. Naturally, the constructivist could maintain that other agents have reason to prevent Caligula from torturing other agents; after all, it might well follow from *their* evaluative standpoint that Caligula is making a moral mistake. But according to the relativity objection, this would not be enough. Allegedly, our ordinary thinking is not only committed to the idea that it would be *right* if people tried to stop Caligula but also that Caligula is *mistaken* in his judgment. In other words, allegedly, our ordinary thinking about morality supports the realist's strong notion of objectivity rather than the antirealist's weaker notion.

What resources does the constructivist have to dispel this objection? First, she might question whether our colloquial intuitions have any relevance for the plausibility of thought experiments that involve ideally coherent creatures with eccentric moral preferences. This is not because ordinary discourse is irrelevant for metaethics – in **section 5.2** I have provided several considerations for thinking otherwise – but because the thought experiments involve creatures that are radically different from the agents that people are familiar with (cf. Street 2009). As a result, it is doubtful that our intuitions about such thought experiments have epistemic value. Arguably, we should suspend our pre-theoretical intuitions about what Caligula has most reason to do, and simply inquire what follows from our best metaethical theory.

Second, the constructivist may grant that our ordinary thinking about ICEs is metaethically relevant but question whether our ordinary thinking does in fact support the realist's strong notion of objectivity. If the constructivist takes this second route, the experimental research on folk objectivism should be taken into account. After all, these experiments provide the most rigorous findings about what people's ordinary thinking about morality is committed to. Specific scenarios involving ICEs have not yet been experimentally tested, but the extant research cited in **section 5.3** provides some preliminary indication that the folk notion of moral objectivity is weaker than that of the realist. For example, as Khoo and Knobe (2018) show, ordinary discourse supports the idea that there can be cases of moral disagreement in which none of the disagreeing parties is mistaken. The Caligula scenario may well be such a case: perhaps we genuinely disagree with Caligula about what he has moral reason to do, but do not think that Caligula is making a mistake in upholding his eccentric judgment.

Moreover, we should be aware that comparisons with the historical Caligula are likely to be distorting where the thought experiment is concerned. After all, it is not at all obvious that the historical Caligula was ideally coherent. In fact, if the constructivist position is along the right lines, there is reason to think that the historical Caligula was probably not ideally coherent. While the historical Caligula did have a reputation for killing on a whim, there is no indication that he was highly imaginative, factually informed and very reflective. Given the many coherence failures that human agents are liable to make, we may, therefore, hypothesize that when the historical Caligula judged that he had reason to torture others for fun, he was *mistaken* by his own lights.

Ideal coherence is a very strenuous condition. In **section 5.4** we discussed some of the many failures that less-than-ideal human agents can make in the light of their own evaluative standpoint. By stipulation, ICEs commit none of these mistakes. Hence, by stipulation our

ideally coherent Caligula is fully aware of the fact that torture causes pain; he is fully aware that his judgment will lead others to condemn and penalize him; he is imaginative and able to think of many other things that he could do for fun. In spite of all of this, he still judges that he has most reason to torture others for fun. We should be alert to the fact that this makes Caligula *a very exotic kind of creature*, who inhabits some possible world which may be quite distant from the world that humans inhabit. He may not share our evolutionary history, he may not be familiar with our sociocultural practices and he may not share our factual knowledge – for if he did, it seems implausible to suppose that his queer judgment would withstand scrutiny from his own standpoint. As Street (2009, p. 281) notes, ‘an accurately imagined ICE will look more like an interesting visitor from another planet than a human being’. If this is correct, then the experiments by Sarkissian et al. (2011) provide further reason to think that our ordinary thinking *does not* support the intuition that Caligula is objectively mistaken. After all, their studies suggest that where aliens are concerned, ordinary respondents do not think that objective moral judgments retain their binding authority. Or, in terms of the constructivist position outlined in **section 5.5**, on our ordinary thinking about morality, creatures like Caligula may not be regarded as *moral* agents.

To sum up, there are grounds for doubting that our colloquial notion of moral objectivity accords particularly well with the strong notion of the realist. Therefore, unless and until further arguments are given, the objection that the antirealist’s notion of objectivity is too relativist is under-supported. The antirealist might even make a preliminary case that the graded relativism typical of ordinary moral discourse is *in tension* with the realist’s strong objectivism. However, these are still the early days of experimental research on folk moral objectivism, and we should be careful not to overstate the implications of this research. What I hope to have achieved, at this stage of the dialectic, is at least to have cast reasonable doubt on the claim, traditionally upheld by many metaethicists, that realists have a leg up on antirealist theories in accounting for the objectivist features of ordinary discourse.

5.7 The Alignment Objection

Another objection that has been pressed against HC concerns its purported inability to make sense of the idea that agents can criticize the normative judgments of other agents (Berker 2014). Consider once more the case of Laura and Louis. Laura judges that she has reason to order tuna, and the Humean constructivist contends that whether this judgment holds true depends on whether it withstands scrutiny from Laura’s evaluative standpoint. But now consider Louis, who judges that Laura has reason *not* to order bluefin tuna. We should, of course, be able to make sense of the idea that Louis can be correct about this judgment. But it may seem that HC cannot make sense of this idea: whether or not Laura has reason to order bluefin tuna depends on whether this judgment withstands scrutiny from Laura’s – not Louis’s – evaluative standpoint. Presumably, their evaluative standpoints do not fully align. Therefore, it seems that Louis is in no position to judge what Laura has reason to do.

Call this the alignment objection.⁵⁵ According to the alignment objection, the fact that

⁵⁵ This objection is distinct from what Bratman (2012, pp. 81, 85, 95) calls ‘the problem of alignment’. By this, Bratman means the challenge for HC to provide an account of the agent’s evaluative standpoint that aligns with insights from the philosophy of action.

we have imperfect knowledge about the evaluative standpoints of other agents precludes us from making correct judgments about the reasons that those agents have. However, in **sections 5.4 and 5.5** we have already seen that the constructivist has resources to dispel this objection. Recall that the constructivist readily admits that agents can make mistakes *by their own lights*. After all, there is a potential mismatch between the reasons that an agent *takes* herself to have and the reasons that she *actually* has, in terms of what follows from her evaluative standpoint. As a result, the reasons that an agent *takes* herself to have can legitimately be criticized if they do not withstand scrutiny from the agent's own evaluative standpoint.

Moreover, other agents may be well positioned to press such criticism. They may do so, for example by updating another agent's non-normative beliefs or by inviting her to take a different perspective and fuelling her imagination. Note that the incentive to change other people's attitudes is typical of *moral* judgments but much less so of other evaluative judgments, such as judgments of taste, convention or aesthetics. While there may be intersubjective agreement on the purported fact that ice cream is tasty, agents will typically show little motivation to push for social convergence on this issue. For moral judgments, on the other hand, there tends to be a much greater incentive to influence people's attitudes. As Gibbard (1990, p. 171) notes, in making a claim to moral objectivity, 'I do more than simply evince the norms that I accept; what I do is to exert a conversational pressure. In effect I demand acceptance of what I am saying.' HC can accommodate this observation; it can procure the idea that at least in terms of appearance, objective moral judgments typically have practical clout. If an agent defends a moral truth that she regards as objective, she typically thinks that this truth *should* withstand scrutiny from the evaluative standpoints of other moral agents, irrespective of whether they think it does. Hence, the agent is committed to achieving social convergence on this point (cf. Mitchell-Yellin 2015, p. 5).

Consider Louis, who is a preservationist. Louis judges that preserving the bluefin tuna as a biological species is valuable. Moreover, he thinks that others have reason to value this as well: not only Laura, but *no* human agent has reason to order a dish that contains bluefin tuna, Louis maintains. In other words, he regards the judgment that human agents have reason not to order bluefin tuna as inescapable: it is binding on all human agents. Other agents, such as Laura, may of course disagree: she does not think that this judgment withstands scrutiny from her own evaluative standpoint. But for Louis, whether Laura *thinks* that this judgment withstands scrutiny from her own standpoint is quite beside the point: what concerns him is whether it *does*. And since he reckons that it *should* follow from Laura's standpoint that she has reason not to order tuna, the challenge for Louis is to update Laura's attitudes and non-normative beliefs in such a way that she comes to realize that, by her own lights, she has reason not to order tuna.

What the alignment objection correctly brings out is that Louis's criticism is fallible. Whether it *does* follow from Laura's evaluative standpoint that she has reason not to order tuna may not be fully transparent to Louis. The Humean constructivist does not deny this. There is likely to be some subjective uncertainty about what does in fact follow from the standpoint of other agents – and unless an agent is ideally coherent, there will also be some subjective uncertainty about what follows from her *own* evaluative standpoint. Indeed, it is this very uncertainty that allows the constructivist to procure the intuition that we can be *mistaken* about our normative judgments, regarding our own reasons as well as the reasons of others.

But the consequences of this uncertainty are not as dire as the alignment objection suggests. Even if evaluative standpoints are not fully transparent, they are not fully opaque either. Agent *A* is familiar with her own attitudes, and she will generally have a clear – if fallible – intuition about what does in fact follow from her own evaluative standpoint. For other agents, agent *A*'s evaluative standpoint may be slightly more opaque, but again, not fully so. In particular, agents who are acquainted with agent *A* are likely to be familiar with many of the values that constitute her standpoint, and even agents who are not acquainted with agent *A* but who do share their evolutionary history with her (as well as, perhaps, much of their cultural upbringing and factual knowledge) may be sufficiently well positioned to assess what follows from her evaluative standpoint.

Hence, even if different standpoints do not fully align, they can still be sufficiently similar to warrant intersubjective criticism. But the constructivist may also grant that in some cases they are not, and that some moral disagreements are unresolvable.

5.8 The Revision Objection

A third objection that has been raised against HC is that the view cannot account for types of hindsight that involve radical changes of value (Jaeger 2015). Consider a former member of the Ku Klux Klan, whose attitudes have radically transformed over time: at *t*₁ he judged that he had reason to discriminate against black people, but at *t*₂ he despises his former racist attitudes. An adequate theoretical account should be able to preserve the idea that the Klan member, looking back at his former self, can judge that he was mistaken about his reasons at *t*₁. But HC cannot preserve this idea, or so it seems. The values that the former Klan member endorses at *t*₂ differ radically from the values he endorsed at *t*₁ – but the evaluative standpoint of the Klan member at *t*₁, rather than *t*₂, sets the standard for what he has reason to do at *t*₁. If, according to his own evaluative standpoint at *t*₁, the Klan member had no reason to discard his racist attitudes, then his later self at *t*₂ cannot criticize him for not having done so. But this is unintuitive, or so it seems: supposedly, we would want to procure the idea that one's past self can be mistaken about his reasons, even if the past self could not have recognized these reasons.

The objector is right to point out that according to HC, the standards of correctness for an evaluative judgment are relative to the further set of values that the agent endorses at the time of making this judgment. But the constructivist has resources to argue that the implications of this view are less implausible than the objector suggests. Instead, she might counter, the objector presents a scenario that is implausible in its own right. To see this, note that on a constructivist reading, if the Klan member at *t*₁ had no reason *at all* to detest his own racist attitudes, this implies that even if the Klan member had been ideally coherent, imaginative and informed, he would still have stuck with his racist values. In other words, the example stipulates that there really is *no ground whatsoever* for criticizing the values of the Klan member based on his own evaluative standpoint at *t*₁. But if this is the case, it seems problematic to suppose that at *t*₂, the former Klan member thinks that he has reason to criticize his former self. After all, if the agent's later self is aware that his former self could not have come up with *any* reason not to *X*, then it seems peculiar to suppose that his later self would nonetheless criticize his former self for not having come up with a reason not to *X*. The constructivist may argue that we had best conclude that the agent's future self is confused, or that the stipulated scenario is simply incoherent.

Alternatively, the scenario may be such that the Klan member at t2 can *legitimately* criticize his former self: there *really was* a reason for the Klan member to detest his own racism, even though he wasn't aware of this reason. Recall, once more, that HC distinguishes between the reasons that an agent *takes* herself to have and the reasons she *actually* has, as determined by what follows from her evaluative standpoint. Because of the potential mismatch between them, it may well be the case that it *did* in fact follow from the Klan member's standpoint at t1 that he had reason to detest his own racist attitudes but that this reason did not become apparent to the agent up until t2. Thus understood, HC has no problem accommodating the scenario: the constructivist can procure the idea that one's past self can be mistaken about his reasons, provided that this past self plausibly *could have* recognized his mistake.

5.9 Conclusion

In this chapter I have highlighted two accounts of moral objectivity: a realist and an antirealist account. On a realist account, morality's objective features are understood in terms of their *attitude-independence*. Just like the judgment that the Earth revolves around the sun, but unlike the judgment that ice cream is tasty, the moral realist maintains that the truth or falsity of a moral judgment is independent of what people think or feel about it.

By contrast, on an antirealist account, morality's objective features are understood in terms of their *standpoint-invariance*. Invariance differs from independence. According to the antirealist, the objectivity of moral judgments does *depend* on people's attitudes, but is constituted by the fact that these judgments withstand scrutiny from a diverse set of evaluative standpoints. As a result, moral judgments are not objective in the same sense as facts about planetary orbits: their truth or falsity is not similarly attitude-independent. But moral judgments also differ from judgments of taste, convention and aesthetics in the sense that moral judgments typically purport to have an inescapable authority, which these other classes of evaluative judgment do not. If the constructivist account is correct, we are mistaken in thinking that moral judgments *either* have to be judgments of fact *or* judgments of taste. Instead, at least where their objectivist aspiration is concerned, moral judgments belong to a different category altogether. This idea fits well with research about folk objectivism: while moral judgments are generally regarded as much more objective than judgments of taste, most people regard them as slightly less objective than factual judgments (Goodwin and Darley 2008; Wright et al. 2013).

It might, perhaps, be objected that the antirealist fails to vindicate the 'real' sense of objectivity. By definition, the objector might maintain, objectivity means attitude-independence; anyone who settles for less has changed the subject. We should keep in mind, however, that objectivity is a rich notion which is appealed to in different senses and across a wide variety of discourses (Nozick 2001; Burge 2010). It is not obvious that one account should have to be vindicated at the cost of all others. Instead, it seems to be more fruitful to assess which account of objectivity is most suitable for which specific domain. While there may certainly be domains in which an understanding of objectivity in terms of attitude-independence is most apt, I have presented some considerations for thinking that in the moral domain, an understanding of objectivity in terms of standpoint-invariance is more suitable.

It has been beyond the scope of this chapter to provide a full comparison of both realist and antirealist accounts. It may, of course, turn out that the realist can accommodate the

experimental findings I have presented in this chapter equally as well as the antirealist. However, if my argument has been successful, I have at least been able to show that, contrary to the prevailing opinion in metaethics, a realist understanding of moral objectivity need not be superior in accounting for morality's internal features. Instead, I have articulated an alternative, antirealist account and have presented some grounds for thinking that this account is in fact better suited to capturing the commitments of ordinary moral thought and practice.

On the account that I have presented, the objectivity of moral judgments may be quite strong, but it is not unbounded. ICEs who are correct in their eccentric moral judgments constitute a genuine possibility, at least in theory. Moral conflicts can have a tragic dimension, the antirealist submits, in the sense that there is no ultimate ground for deciding which of the conflicting parties is right. In practice such cases of deep conflict may be rare – ideal coherence, we have noted, is an achievement that might hardly ever be achieved by human agents in our actual world – but they constitute a theoretical possibility nonetheless. Those with realist inclinations may consider this a weakness of the account. Those who think that a vindication of moral objectivity need not preclude some degree of moral relativism, on the other hand, might consider it one of its deeper appeals. The tragic dimension of moral disagreement is a dimension that the antirealist admits to exist – and in a different sense of realism, I submit, she is all the more realistic in doing so.

References

- Bagnoli, C. 2015. Moral Objectivity: A Kantian Illusion? *Journal of Value Inquiry* 49: 31-45.
- Beebe, J.R. 2014. How Different Kinds of Disagreement Impact Folk Metaethical Judgments. In: J.C. Wright and H. Sarkissian, eds. *Advances in Experimental Moral Psychology: Affect, Character, and Commitments*. London: Continuum, 167-187.
- Berker, S. 2014. Does Evolutionary Psychology Show that Normativity Is Mind-Dependent? In: J. D'Arms and D. Jacobson, eds. *Moral Psychology and Human Agency: Essays on the Science of Ethics*. New York: Oxford University Press, 215-252.
- Blackburn, S. 1984. *Spreading the Word: Groundings in the Philosophy of Language*. New York: Oxford University Press.
- Bratman, M. 2012. Constructivism, Agency, and The Problem of Alignment. In: J. Lenman and Y. Shemmer, eds. *Constructivism in Practical Philosophy*. Oxford: Oxford University, 81-98.
- Burge, T. 2010. *Origins of Objectivity*. Oxford: Clarendon Press.
- Carnap, R. 1956. *Meaning and Necessity*. Chicago: University of Chicago Press.
- Clarke-Doane, J. 2014. Moral Epistemology: The Mathematics Analogy. *Noûs* 48: 238-255.
- Enoch, D. 2011. *Taking Morality Seriously: A Defense of Robust Realism*. Oxford: Oxford University Press.
- Enoch, D. 2014. Why I Am an Objectivist about Ethics (And Why You Are, Too). In: R. Shafer-Landau, ed. *The Ethical Life*. Oxford: Oxford University Press, 192-205.
- Finlay, S. 2007. Four Faces of Moral Realism. *Philosophy Compass* 6: 820-849.

- Flanagan, O., Ancell, A., Martin, S. and Steenbergen, G. 2014. Empiricism and Normative Ethics: What Do the Biology and the Psychology of Morality Have To Do with Ethics? *Behaviour* 151: 209-228.
- Gibbard, A. 1990. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Oxford: Clarendon Press.
- Gibbard, A. 1999. Morality as Consistency in Living: Korsgaard's Kantian Lectures. *Ethics* 110: 140-164.
- Goodwin, G.P. and Darley, J.M. 2008. The Psychology of Meta-Ethics: Exploring Objectivism. *Cognition* 106: 1339-1366.
- Goodwin, G.P. and Darley, J.M. 2012. Why Are some Moral Beliefs Perceived to be More Objective than Others? *Journal of Experimental Social Psychology* 48: 250-256.
- Habermas, J. 1995. Reconciliation through the Public Use of Reason: Remarks on John Rawls's Political Liberalism. *The Journal of Philosophy* 92: 109-131.
- Handfield, T. 2016. Genealogical Explanations of Chance and Morals. In: U.D. Leibowitz and N. Sinclair, eds. *Explanation in Ethics and Mathematics: Debunking and Dispensability*. Oxford: Oxford University Press, 58-82.
- Jaeger, G. 2015. The Wisdom of Hindsight and the Limits of Humean Constructivism. *Philosophical Explorations* 18: 45-57.
- Joyce, R. 2006. *The Evolution of Morality*. Cambridge: MIT Press.
- Kahane, G. 2011. Evolutionary Debunking Arguments. *Noûs* 45: 103-125.
- Khoo, J. and Knobe, J. 2018. Moral Disagreement and Moral Semantics. *Noûs* 52: 109-143.
- Leibowitz, U.D. and Sinclair, N. eds. 2016. *Explanation in Ethics and Mathematics: Debunking and Dispensability*. Oxford: Oxford University Press.
- Mackie, J.L. 1977. *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin.
- Mitchell-Yellin, B. 2015. Aligning with the Good. *Journal of Ethics and Social Philosophy*, 1-8.
- Nichols, S. 2004. After Objectivity: An Empirical Study of Moral Judgment. *Philosophical Psychology* 17: 5-28.
- Nichols, S. 2014. Process Debunking and Ethics. *Ethics* 124: 727-749.
- Nozick, R. 2001. *Invariances: The Structure of the Objective World*. Cambridge: Belknap Press of Harvard University Press.
- Sarkissian, H., Parks, J., Tien, D., Wright, J.C. and Knobe, J. 2011. Folk Moral Relativism. *Mind & Language* 26: 428-505.
- Scanlon, T.M. 1998. *What We Owe to Each Other*. Cambridge: Harvard University Press.
- Sen, A.K. 1993. Positional Objectivity. *Philosophy & Public Affairs* 22: 126-145.
- Shafer-Landau, R. 2003. *Moral Realism: A Defence*. New York: Oxford UP.
- Street, S. 2006. A Darwinian Dilemma for Realist Theories of Value. *Philosophical Studies* 127: 109-166.
- Street, S. 2008a. Reply to Copp: Naturalism, Normativity, and the Varieties of Realism Worth Worrying About. *Philosophical Issues* 18: 207-228.

- Street, S. 2008b. Constructivism about Reasons. *Oxford Studies in Metaethics*, vol. 3: 207-245.
- Street, S. 2009. In Defense of Future Tuesday Indifference. Ideally Coherent Eccentrics and the Contingency of What Matters. *Philosophical Issues* 19: 273-298.
- Street, S. 2010. What Is Constructivism in Ethics and Metaethics? *Philosophy Compass* 5: 363-384.
- Street, S. 2012. Coming to terms with Contingency: Humean Constructivism about Practical Reason. In: J. Lenman and Y. Shemmer, eds. *Constructivism in Practical Philosophy*. Oxford: Oxford University Press, 40-59.
- Street, S. 2016. Objectivity and Truth: You'd Better Rethink It. In: R. Shafer-Landau, ed. *Oxford Studies in Metaethics*, vol. 11. Oxford: Clarendon Press, 293-334.
- Street, S. 2017. Nothing 'Really' Matters, but That's Not What Matters. In: P. Singer, ed. *Does Anything Really Matter: Parfit on Objectivity*. Oxford: Oxford University Press, 121-148.
- Timmons, M. 2010. Objectivity in Moral Discourse. In: A. Barber and R. Stainton, eds. *Concise Encyclopedia of Philosophy of Language and Linguistics*. Oxford: Elsevier Ltd., 539-545.
- Wright, J.C., Grandjean P.T. and McWhite, C.B. 2013. The Meta-Ethical Grounding of Our Moral Beliefs: Evidence for Metaethical Pluralism. *Philosophical Psychology* 26: 336-361.

6

Chapter 6

The Metaethical Significance of Experiments about Folk Moral Objectivism

Forthcoming in *Philosophical Psychology*.

Keywords

*Experimental metaethics; moral psychology; folk pluralism; moral semantics;
presumptive arguments*

Abstract

The metaethical commitments of laypeople – specifically their commitment to the objectivity of moral claims – have recently become subject to empirical scrutiny. Experimental findings suggest that people are metaethical pluralists: there is both inter- and intrapersonal variation with regard to people’s objectivist commitments. What metaethical implications, if any, do these findings have? I point out that current research does not directly address traditional metaethical questions: the methods used and distinctions drawn by experimenters do not perfectly match those of metaethicists. But I go on to argue that, in spite of this mismatch, the research findings should be of interest to moral philosophers, including metaethicists. Not only do these findings extend the field of moral psychology with new data and hypotheses, but they also provide tentative evidence that touches on the adequacy of theses in moral semantics and moral metaphysics. Specifically, I argue that they put pressure on presumptive arguments in support of moral realism.

‘There are many ways to approach the question, and I don’t want to deny that sitting in one’s armchair and speculating may be among them (such activity often casts up promising ideas) –but the ultimate arbiter must be the body of a posteriori data issuing from such disciplines as psychology, anthropology and experimental economics. Even an evolutionary genealogy may shed light on the subject, since an understanding of how having a moral sense advanced the reproductive success of our ancestors may have a lot to do with what role(s) it continues to play.’

– Richard Joyce (2006b)

6.1 Introduction

Over the last decade several philosophers and psychologists have begun to test non-philosophers’ commitments concerning the objectivity of moral judgments in an experimentally rigorous way (e.g. Nichols 2004; Goodwin and Darley 2008, 2012; Sarkissian et al. 2011; Wright et al. 2013, 2014; Beebe 2014; Fisher et al. 2017; Theriault et al. 2017; Khoo and Knobe 2018). Empirical research on people’s metaethical commitments is not entirely new; the moral–conventional distinction has long been a topic of experimental research (e.g. Turiel 1983), and specific kinds of moral violations, such as the violation of harm norms, have also been studied experimentally (e.g. Haidt et al. 1993). However, only in recent years have experimenters begun to inquire systematically how lay individuals think about issues that have traditionally been focal points of metaethical theorizing – e.g. about the nature of moral motivation, moral disagreement, moral truth and moral objectivity.

In this chapter I focus on one specific topic of this new wave of research – folk moral objectivism – and assess its metaethical implications. These implications are not straightforward. Critics of experimental philosophy (e.g. Kauppinen 2007; Ludwig 2010; Bengson 2013) even question whether experimental research can touch upon traditional philosophical questions *at all*. As these critics point out, the survey methodology typically employed by experimenters is ill-suited to scrutinizing people’s reflective commitments. By contrast, when metaethicists make a claim about laypeople’s objectivist commitments, their claim typically *does* concern people’s reflective commitments. Hence, there seems to be a mismatch between the *explanandum* of experimenters and that of metaethicists. Moreover, this mismatch is aggrandized by the fact that the distinctions drawn by experimenters on folk moral objectivism do not always perfectly match those drawn by metaethicists.

I submit that these worries should be taken seriously, and call for great caution in extrapolating the results of experimental findings to a metaethical context. But I will argue, nonetheless, that there may be fertile interactions between philosophical metaethics and psychological studies of people’s metaethical commitments. The primary upshot of the new wave of experiments is extending the field of moral psychology with a new set of questions, for instance regarding the psychosocial strategies underlying people’s metaethical commitments. The theories thereby advanced – such as Wright et al.’s (2013, 2014) theory of ‘folk pluralism’ – should not be regarded as *alternatives* to traditional metaethical theories: they shed light on the *psychology* of moral judgment, whereas metaethicists have traditionally been interested in the *concept* of moral judgment. In this respect, the new findings and

hypotheses are complementary to traditional metaethical work. But additionally, some experimental findings provide indirect evidence that touches upon issues in moral semantics and moral metaphysics.

The chapter proceeds as follows. In **section 6.2** I summarize the new wave of experimental research on folk moral objectivism and outline some of its most well-established findings. In **section 6.3** I make some critical remarks about the methodology used by several experimenters and explain why their findings do not directly address traditional metaethical concerns. But I proceed to argue that, nonetheless, experimental research *does* have potential to shed light on some traditional metaethical theses. In **section 6.4** I discuss the *concept* of moral judgment and argue that although the conceptual implications of experimental findings can only be indirect, these findings create argumentative pressures in extant debates, for example about the tenability of error theory. In **section 6.5** I turn to the relevance of experimental findings for moral metaphysics and argue that they give reason to question presumptive arguments in support of moral realism. In **section 6.6** I highlight recent hypotheses about the psychosocial function of metaethical judgments, building on the experiments discussed. These hypotheses are more naturally understood as belonging to the field of moral psychology than metaethics; nonetheless, I argue that they may be used as a further resource to argue against the presumptive support for moral realism. I conclude that although research on folk moral objectivism is still in its early stages, and various methodological improvements are called for, this research is beginning to fill an empirical hiatus and has the potential to become an important resource in metaethical theorizing.

6.2 Experimental Findings: An Overview

In this section I highlight some of the main findings of the recent wave of experiments on folk objectivism, assuming the validity of the underlying research methodology (I discuss the limitations of this methodology in **section 6.3**). Such experiments have been conducted across different cultures and with different age groups, using different research methods and asking respondents a variety of questions. The most common research method has been to present respondents with surveys that are designed to test people's objectivist commitments. For instance, respondents might be asked whether or not they regard the statement 'Consciously discriminating against someone on the basis of race is unacceptable' as a moral statement, whether they think that the statement is true, false or a matter of opinion and whether they think that in cases of disagreement one of the disagreeing parties must surely be mistaken or whether both parties might be equally justified in holding their view (e.g. Goodwin and Darley 2008; Wright et al. 2013). Experimenters commonly operationalize objectivity as the combined commitment to moral statements being either true or false, and as moral disagreement being indicative of a mistake. For example, if respondents judge that the claim 'Consciously discriminating against someone on the basis of race is unacceptable' is true and that in cases of disagreement one of the disagreeing parties must surely be mistaken, they will be classified as objectivists. If they think that this claim is a matter of opinion and that in cases of disagreement both parties might be equally correct, they will be classified as non-objectivists.

Concerning the moral domain in general, it has been found that people typically regard moral judgments as much more objective than judgments of taste, somewhat more objective than judgments of convention and slightly less objective than judgments of fact (Goodwin and

Darley 2008; though see Theriault et al. 2017 for a somewhat different result). However, there is substantial variation with regard to the objectivity attributed to specific moral judgments, which depends *inter alia* on the topic under consideration (Wright et al. 2013). Hence, rather than being regarded as uniformly objective, research suggests that people perceive moral objectivity as being on a sliding scale: some moral issues are regarded as being as objective as matters of fact; others are regarded as being more akin to questions of taste.

The variability of test subjects' objectivist endorsements has been one of the most notable and consistent findings of the experimental research. The fact that people are pluralists with regard to their commitment to moral objectivity is now well established. Indeed, a good case can be made that the results of earlier studies which suggested that people are objectivists should be reinterpreted as supporting the view that people are metaethical pluralists (Pözlner 2017). This pluralism manifests itself in different dimensions.

6.2.1 Interpersonal Variation

First, variation exists in the objectivist leanings of different people. This variation has been found to correlate with differences in personality traits (Feltz and Cokely 2008) as well as with the different stage in the lifespan that people are at (Beebe and Sackris 2016): people tend to become gradually more objectivist over time, but increasingly relativist during their early twenties. Goodwin and Darley (2010) found that a religious grounding of ethics is another predictor of an objectivist metaethical stance, whereas a better capacity for 'disjunctive thinking' correlates with diminished moral objectivism.⁵⁶

6.2.2 Intrapersonal Variation

Second, variation in the objectivity ascribed to moral judgments also has intrapersonal dimensions. One of these is the aforementioned variability in terms of topic. Depending on the moral issue under consideration, moral judgments may be regarded as very objective, or much less so (Wright et al. 2013). Typically, moral issues that are regarded as subjective tend to be issues that are politically controversial – e.g. abortion and assisted suicide among American respondents. Some experimenters have found that perceived social consensus is an important predictor of this variability (Goodwin and Darley 2012; Beebe 2014): the less social consensus there is on a given moral issue, the less objective people regard the issue. Moreover, at an early age the idea is already emerging that disagreements about widely shared moral beliefs have only one right answer, whereas disagreements about controversial moral beliefs do not (Heiphetz and Young 2017).

⁵⁶ A disjunctive reasoning ability denotes the capacity to actively unpack alternative possibilities when reasoning, which is tested, for instance, by presenting respondents with the 'five blocks' task:

There are five blocks in a stack; the second one from the top is green and the fourth one from the top is not green. Is a green block definitely on top of a non-green block?

Participants have to choose between three response options: yes, no and cannot tell. The intuitive response 'cannot tell' is incorrect, as becomes clear after explicitly considering what follows if the third block is green and what follows if the third block is not green.

The level of objectivity attributed to moral judgments also varies with contextual and situational factors. Sarkissian et al. (2011) found that if two disagreeing parties are depicted as belonging to the same cultural group, then in situations of disagreement respondents typically judge that one of the disagreeing parties must be mistaken. But when the disagreement involves individuals from different cultural or different evolutionary backgrounds, then the answers to moral disputes are regarded as more relativist, and respondents are inclined to judge that both parties to the disagreement can be correct. Hence, standards of moral correctness are taken to be at least somewhat relative to a specific culture: objectivist intuitions vary because of cultural distance, with increased distance leading to decreased attributions of objectivity. Using a similar research set-up, Khoo and Knobe (2018) found that there are moral conflict cases in which people are inclined to say both that two speakers disagree and that it is not the case that at least one of them must be saying something wrong. Hence, it seems that people allow for the possibility that there are moral disagreements in which the claims of the disagreeing parties do not exclude each other.

Goodwin and Darley (2012) found that greater moral objectivity is associated with ‘closed’ rather than ‘open’ responses in the face of moral disagreement. Following up on this finding, Fisher et al. (2017) found that, at least with regard to controversial issues, people’s mode of social interaction influences their construal of moral truth. When people adopt a cooperative ‘argue-to-learn’ mindset, they typically think that controversial moral issues admit more than one correct answer. When people adopt a competitive ‘argue-to-win’ strategy, they typically take on a more objectivist stance.

Lastly, there have been some findings suggesting that people’s metaethical commitments affect their moral behaviour and everyday moral decision-making. For instance, Wright et al. (2014) found that subjects stated to be less willing to interact with and to help people who disagree about moral issues, especially if these issues are taken to be objective. The authors suggest that regarding a moral issue as objective is related to less tolerance for disagreement, whereas non-objectivism leaves an opening for diversity and dialogue.

These findings shed light on a new area of moral psychology – the psychology of metaethics. They may provide insight into the psychological and social functions that attributing a certain metaethical status to one’s judgments might serve (see **section 6.6**, as well as **chapter 7**). But none of the experiments reviewed here engages directly with traditional metaethical objects of study, such as the *concept* of moral judgment or the *metaphysics* of moral objectivity. As a result, and as I will emphasize in the next section, attempts to draw *direct* metaethical conclusions based on current experimental findings are often problematic.

6.3 Metaethical Implications? Pitfalls and Promises

The aforementioned findings and hypotheses are steadily beginning to fill an empirical hiatus in discussions about moral objectivity. But their metaethical promise should not be overstated, as recent commentators have pointed out (Beebe 2015; Pölzler 2018, chapter 3). In fact, some philosophers maintain that the ambitions of experimentalists to shed light on traditional philosophical questions is misguided in general (e.g. Kauppinen 2007; Ludwig 2010; Bengson 2013). Building on these criticisms, in this section I highlight some of the pitfalls of experiments on folk moral objectivism and explain why current studies do not lend themselves

to draw *direct* metaethical conclusions.⁵⁷ These limitations need to be taken seriously. Nonetheless, I will conclude the section by arguing that they do not fully undermine the promise of current experimental findings to serve as a resource in the metaethical debate.

One potential pitfall of many of the studies mentioned in **section 6.2** is their reliance on a survey methodology. Several critics of experimental philosophy have underlined the deficiencies of surveys as a means to investigate people's philosophical stances. When philosophers make an assertion about the contents of people's intuitions, or about people's conceptual commitments, they typically mean to make a claim about these intuitions or commitments under idealized conditions. Performance errors, semantic incompetence and other philosophically irrelevant factors may detract from these idealized conditions in survey-based research. Indeed, some philosophers maintain that the kinds of intuitions that philosophers are interested in can only become apparent in Socratic dialogue (Kauppinen 2007). The survey methodology used by experimental philosophers, by contrast, typically serves to probe respondents' immediate responses, often prompted by means of a forced choice paradigm. Such responses are typically quick guesses, hypotheses and emotional reactions (Bengson 2013), which clearly differ from the 'intellectual intuitions' that philosophers are interested in (Ludwig 2010, p. 430).

This criticism also applies to research on folk moral objectivism, which tends to rely heavily on a survey methodology. Surveys seem to capture people's *first-off* metaethical responses. As such they differ from metaethicists' traditional object of analysis, which is to capture people's *endorsed* commitments. Therefore, the experimental findings cannot directly shed light on traditional metaethical questions.

Adding to this mismatch is the fact that the distinctions experimenters draw do not always capture existing metaethical distinctions. To give just one example, consider experimenters' operationalization of moral objectivity as the dual commitment to a moral statement being either true or false, and of moral disagreement being indicative of mistake (e.g. Goodwin and Darley 2008). On a standard metaethical classification, the former part of the experimentalist's operationalization is compatible with various metaethical views, including moral realism, subjectivism and constructivism. Yet many metaethicists regard only the first of these views as a form of objectivism, in virtue of the realist's commitment to moral truths being stance-independent (Shafer-Landau 2003). Hence, what many experimentalists regard as a condition that adds to a moral statement's objectivity, many metaethicists do not. This illustrates, once more, that experimentalists tend to study different phenomena from those that metaethicists have traditionally been interested in.

Insofar as they aspire to shed light on traditional metaethical theses about moral objectivity, then, we should conclude that current experiments suffer from serious flaws. But if we look at the aspirations of these experiments from another angle, as shedding light on a different set of phenomena altogether – people's immediate responses rather than their reflective intuitions; moral objectivity in the experimenter's psychological, rather than the

⁵⁷ As Hannon (forthcoming) rightly stresses, the philosophical significance of findings from experimental philosophy should be evaluated on a case-by-case basis. While the general observations made in this section apply to *some* studies on folk moral objectivism, it should be kept in mind that they do not apply to *all* of them; some existing studies are more careful in drawing metaethical distinctions than others.

metaethicist's philosophical sense – then this mismatch does little to discredit the experiments. This, I submit, is the most charitable interpretation of extant experiments involving folk objectivism: they shed light on a topic that lies in the vicinity of, but differs from, traditional metaethical concerns.

Apart from being criticized for its mismatch with traditional metaethics, the methodology of experiments on folk moral objectivism has also been criticized on other grounds. As Beebe (2015) points out, in several existing studies the questions posed by experimenters can be read in multiple ways, and there is no guarantee that the answers respondents give track the metaethically relevant distinctions. More generally, it may be questioned whether current studies really succeed in pulling apart task demands, experimenter effects and people's metaethical commitments. Additionally, many current studies suffer from the 'stimuli as fixed effects fallacy' (Judd et al. 2012): they generalize from findings about specific examples to conclusions about a domain in general, even though there is no statistical basis for this generalization.

All of these methodological flaws may be overcome, at least in principle (see, for example, Theriault et al. 2017), but they do detract from the general reliability of existing results. In what follows, I will proceed on the assumption that upon overcoming these methodological flaws, the main research findings, as outlined in **section 6.2**, will remain intact. It should be kept in mind, however, that future research may prove this assumption wrong.

In closing this section, let's start to consider whether, and how, current findings may still be metaethically relevant, in spite of the criticisms outlined above. While these findings do not have *direct* metaethical implications, they do lie in the vicinity of traditional metaethical concerns, and I will argue that the relations are worth exploring – both with regard to traditional metaethical questions (see also **section 6.4** and **section 6.5**) and with regard to the prospect of new philosophical questions that psychological research might give rise to (see also **section 6.6**).

To start with the latter, one of the upshots of the experiments is that they highlight previously unrecognized patterns of stability in folk metaethical judgments. These patterns might be relevant for metaethical practice – for instance, by highlighting how the choice of examples can influence metaethical analysis (see **section 6.4**). Additionally, these patterns may themselves be proper targets of metaethical analysis and may be connected with theoretical insights about the psychological, social and evolutionary functions of folk metaethical judgments. While some of these insights go beyond the strictly metaethical domain, they can be connected with theses that have traditionally been advanced in a metaethical context. Consider, for instance, Michael Ruse's (1986) argument in favour of a version of error theory which rests on an empirical hypothesis about the biological function of objectivist moral beliefs. While the philosophical tenability of Ruse's argument is contested (see, for example, Joyce 2016), Ruse's empirical hypothesis might be of metaethical interest nonetheless. Research on folk moral objectivism has recently contributed to the formation of an elaborate novel hypothesis in the same evolutionary ballpark (see Stanford 2018) which raises questions that are of interest to moral psychologists, and possibly to metaethicists as well. For instance, as I will argue in **section 6.6**, Stanford's suggestion that the objectivist features of our moral phenomenology can be evolutionarily explained can be taken up in an argument against moral realism.

Secondly, with regard to the relevance of psychological experiments for traditional metaethical questions, and in defence of the survey method used by experimenters, it should be noted that although people's prompted responses about moral objectivity *may* come apart from their considered intuitions, it is by no means guaranteed that they *will*. In fact, there is some evidence that the patterns in philosophical responses unveiled by experimental research on first-off intuitions do not disappear after greater reflection (Weinberg et al. 2012). Further experimental research is needed here; awaiting such research, we can only guess how people's immediate responses relate to their reflective intuitions. But given the pervasiveness of rationalization in moral reasoning, it seems reasonable to guess that findings about people's first-off moral responses provide at least *tentative* evidence about the contents of their reflective intuitions.

Thirdly, in response to the criticism that experimenters often do not frame moral objectivism in the same way that metaethicists – particularly realists – tend to do, it should be pointed out that there is in fact substantial variation in the way that metaethicists think of moral objectivity. To give just one example of this diversity, the aforementioned realist account of moral objectivity in terms of stance-independence has been challenged – partly on experimental grounds – by an antirealist account, which frames moral objectivity in terms of stance-*invariance* (see my arguments in this dissertation, especially **chapter 5**, which has been published as Hopster 2017). The fact that different metaethicists adhere to different accounts of moral objectivity should give some pause to critics who argue that experimentalists fail to capture a metaethically relevant phenomenon, since they do not capture moral objectivity in the realist's sense. As Beebe (2015, p. 27) rightly points out, 'the investigation of folk views about the objectivity of morality should be broadened to encompass more than just the issue of moral realism'. Perhaps, depending on context, moral judgments may purport to be objective in a variety of ways (e.g. in being stance-independent; stance-invariant; authority-independent, *etc.*). At the very least, the metaethical primacy of a realist conception of moral objectivity is not sufficiently well established to regard research about alternative conceptions as metaethically irrelevant.

Lastly, the fact that there exists substantial diversity among the views of professional metaethicists, all of whom have dedicated much time to reflecting on the topic of moral objectivity, may be taken as suggestive – if anecdotal – evidence in support of the hypothesis that the pluralism of folk respondents does not merely reflect their incompetence, but constitutes a real feature of people's metaethical intuitions.

6.4 The Concept of Moral Judgment

Suppose that current experimental findings (**section 6.2**) are indeed along the right lines and that the metaethical status that people attach to moral judgments varies along multiple dimensions. How might this be relevant for traditional metaethical analysis?

First, the findings are in tension with traditional philosophical approaches to metaethics, many of which do not allow for much flexibility or hybridization. As Gill (2009) argues, much of 20th-century metaethics proceeded on the assumption that ordinary moral discourse is uniform and determinate. Classical metaethical questions – for example, whether there is a necessary connection between moral judgment and moral motivation, or whether moral language is cognitivist or non-cognitivist – were assumed to have invariant and

determinate answers. Experimental findings provide preliminary evidence that this assumption is mistaken: there is inter- and intrapersonal variation concerning the metaethical status of moral claims, which problematizes any view which assumes that this status is invariant (cf. Loeb 2008; Beebe 2014; Khoo and Knobe 2018).

Secondly, the findings give pause to metaethicists who defend a thesis on the basis of a narrow set of examples. One important lesson for metaethicists to learn from the experimental findings is that their choice of examples might heavily influence the resultant analysis (cf. Sarkissian 2017, p. 574). Experiments show that situational and contextual factors substantially impact respondents' objectivist proclivities, which should provide caution to metaethicists about regarding a few core examples as being representative of moral discourse as a whole. For instance, if a metaethicist's example of choice is a particularly uncontroversial case ('torturing innocent people is morally bad'), then people's metaethical intuitions will tend to be more objectivist than with regard to disputed moral claims. Similarly, if an example is presented in a situation of intra-cultural conflict, then people's metaethical intuitions will be more objectivist than in a situation of inter-cultural dialogue. Here metaethicists can clearly profit from experimental findings: by providing a better understanding of the general correlates of people's moral intuitions, these experiments can serve as a corrective to a biased choice of examples and can provide a better understanding of the respects in which alleged paradigm cases of moral judgment are in fact representative of moral judgment across the board.

Thirdly, the findings may help to shed light on the contents of the *concept* of moral judgment. It seems plausible that, with regard to our concepts, there has to be a fairly tight connection between their actual contents and the contents that most people *take* them to have. If this connection were very loose, then it would seem that in analysing the contents of our concepts, philosophers are simply changing the topic: they are no longer analysing the concepts that people ordinarily use (cf. Loeb 2008). Metaethicists during the latter half of the 20th century have often assumed that in making moral judgments, people are conceptually committed to a claim to objectivity. The finding of folk pluralism, however, suggests that this may not always be the case; at the very least, this shifts the burden of argument to metaethicist who maintain that people have this conceptual commitment.

One possible response for metaethicists is to withdraw to the armchair and deny that experimental findings have any conceptual relevance at all. This position may, for example, be advanced by externalists regarding the contents of moral concepts. According to content-externalists, features of the world function as reference-fixers for moral concepts. What laypeople make of the concept of MORAL JUDGMENT need not carry much evidential weight: its true meaning may simply be opaque to them.

Most metaethicists, however, find it more plausible to pursue the route of content-internalism: to view moral concepts as mental constructs that are transparent to reflection (Laskowski and Finlay 2017). For content-internalists who claim that the concept of MORAL JUDGMENT commits us to a claim to objectivity, experimental findings pose an argumentative challenge. Perhaps this challenge can be met, for reasons similar to those mentioned in **section 6.3**. For instance, content-internalists commonly assume that upon reflection, competent participants in moral discourse will tend to agree whether or not a proposition (such as the proposition that moral judgments make a claim to objectivity) should be regarded as a platitude. By contrast, existing experiments scrutinize respondents' immediate commitments to moral

objectivity rather than their stance upon reflection (also, they do not check for the conceptual competence of test subjects). As a result, present experimental evidence may be too shaky to have a great impact on this debate. With improved experimental methodologies, however, it could potentially have such an impact.

The outcome of this debate, in turn, is also relevant for the tenability of broader metaethical theories, such as John Mackie's (1977) error theory. According to Mackie, our concept of MORAL JUDGMENT entails that it refers to objectively prescriptive properties. Current experimental findings put some pressure on Mackie's empirical claim by suggesting that people's commitment to objectivism is both more moderate and more flexible than Mackie assumes. We should keep in mind that no research has yet been conducted that specifically questions whether people intend the concept of MORAL JUDGMENT to refer to objectively prescriptive properties (cf. Fraser 2014). Instead, people's objectivist commitments are tested using a somewhat different measure, which can only provide indirect evidence for or against Mackie's error theory. But even if current evidence is tentative, it provides preliminary reason to think that Mackie's conceptual claim does not hold up – at least not for *all* moral judgments (cf. Sarkissian 2016, p. 222). In this indirect way, current experimental findings can be relevant for extant metaethical debates by shifting the burden of argument. However, their capacity to do so often depends on additional metaethical assumptions and may have to partly await further empirical exploration (see Pölzler forthcoming for further discussion).

6.5 The Metaphysics of Moral Judgment

In the previous section I pointed out that with regard to our conceptual commitments, it seems plausible that there has to be a fairly tight connection between what these commitments are and what people take them to be. With regard to our metaphysical commitments, the same reasoning may not apply. The fact that many people once thought that witches exist, or think that God exists, does little to enhance the probability that witches or God actually exist (cf. Joyce 2015). Similarly, even if it were experimentally shown that many people believe in the existence of objectively prescriptive properties, this does not establish that such properties indeed exist. They might simply be mistaken.

We should, therefore, expect experimental findings about folk objectivism to be of little relevance to the metaphysics of moral judgment. Nonetheless, as I will argue in the present section, given the actual dialectic of the metaethical debate, such findings are not completely mute. An influential type of argument that several metaethicists have advanced *does* purport to derive metaphysical conclusions on the basis of a claim about folk judgments: the so-called experiential or presumptive argument in support of moral realism (e.g. Brink 1989; Shafer-Landau 2003; see Kirchin 2003; Loeb 2007; Sinclair 2012; Pölzler 2017 for criticisms). Since experiments involving folk moral objectivism touch upon the adequacy of this claim about folk judgments, they also touch upon the adequacy of the presumptive argument in support of realism.

There are, in fact, various kinds of presumptive arguments that play a role in metaethics; they engage with different aspects of moral discourse and practice, and differ in how they claim these aspects to be metaethically relevant (e.g. Loeb 2007). Some presumptive arguments claim to best accommodate moral discourse in the sense of best *explaining* it; others claim to best accommodate moral discourse in the sense of best *justifying* it. In what follows, I shall be

interested in the latter version of the argument, as used in support of moral realism. Following Sinclair (2012, pp. 158-159), this argument relies on three premises:

- *Premise 1*: A claim about the face-value commitments of moral discourse and practice;
- *Premise 2*: The claim that moral realism best justifies these face-value commitments;
- *Premise 3*: The claim that it is a *desideratum* of metaethical theories to justify the face-value commitments of moral discourse and practice.

The first claim is empirical and concerns findings of the sort that experiments involving folk objectivism shed light upon. The second is a philosophical claim about the justificatory power of moral realism versus rival theories. The third claim concerns the metaethical relevance of people's metaphysical intuitions (it contradicts the suggestion with which this section began – that such intuitions are of limited metaethical relevance).

The presumptive argument in support of moral realism is of particular importance to the metaethical debate, as it has an alleged potential to shape the philosophical dialectic. If the argument holds up, then the starting point of metaethical inquiry should be to vindicate some form of moral realism. Indeed, this is precisely what David Enoch (2018), another defender of this type of argument, has suggested in recent work. In the remainder of this section I shall critically discuss Enoch's argument on empirical grounds.

According to Enoch, moral discourse is implicitly committed to a realist – and specifically non-naturalist – ontology. By way of illustration, he considers the following six considerations about the face-value commitments of moral discourse (*idem*, pp. 29-42):

- 1) Moral language functions much like representational language in other domains;
- 2) Moral discourse seems to exhibit objective purport. For example, we are comfortable applying moral standards to other people, without inquiring about their own moral inclinations, and we don't withdraw once we find out that their moral commitments differ from ours;
- 3) We seem to endorse – pre-theoretically – counterfactuals that do not sit well with response-dependence (if eating meat is morally wrong, then presumably it would have been wrong even had no one ever acknowledged its wrongness);
- 4) We think that in cases of moral disagreement only one of the disagreeing parties can be right;
- 5) Paradigmatic moral facts and properties appear to be very different from paradigmatic natural ones;
- 6) Moral evaluation is not just the evaluation of something-qua-something, but of goodness *simpliciter*.

As Enoch is aware, these considerations do not *vindicate* non-naturalist realism. Other metaethical theories might also be able to accommodate the face-value commitments. Nonetheless, they do seem to give non-naturalist realism an advantage over rival theories, as they have a particularly good fit with the data. Hence, Enoch contends that all other things

being equal, face-value observations make non-naturalist realism ‘the view to beat’ in metaethics.

6.5.1 Criticism of *Premise 1*

Is Enoch’s presumptive argument in support of moral realism tenable when taking into consideration the current evidence about folk moral objectivism? Let’s assess how Enoch’s six considerations hold up in the light of this research:

- 1) Whether people think that moral language in general is representational is a question that has not yet been a specific topic of experimental research. Indeed, some might want to argue (e.g. Sinclair 2012, p. 168) that this is not the kind of question that is apt for experimental research. After all, that the function of moral language is representational is a sophisticated philosophical claim and unlikely to be part of folk metaethical discourse. But if this is correct, then Enoch’s presumption begs the question: he simply presumes that people are moral realists in a technical sense.

Alternatively, it may be argued that the question of whether moral language is representational is implicit in moral discourse and can be experimentally tested. Several experiments have tested whether participants think that moral statements are true or false, or matters of opinion (e.g. Goodwin and Darley 2008, 2012; Wright et al. 2013, 2014). The findings of these studies do not corroborate Enoch’s contention. By contrast, they suggest that the face-value commitments of moral discourse are a thoroughly mixed bag and that whether respondents think that moral claims are truth-apt or not is *inter alia* dependent on the topic under consideration. These findings suggest that the claim that moral language functions much like representational language is likely to hold up only under a specific choice of examples.

- 2) Does moral discourse exhibit objective purport? Again, there is experimental evidence that the answer depends on contextual factors. Whether people are comfortable applying moral standards to other people depends on the cultural and evolutionary proximity of these people (Sarkissian et al. 2011). This suggests that in some situations moral discourse *does* exhibit objective purport, but in other situations it does not.
- 3) What moral counterfactuals people are inclined to endorse has not been experimentally studied. It won’t come as a surprise that realists like Enoch endorse counterfactuals that do not sit well with response-dependent views, but it would not be a surprise either if countervailing counterfactuals were endorsed by antirealists (e.g. Street 2009). Before further research is done, it seems premature (and question-begging) to attach much argumentative weight to these suspicions.
- 4) Enoch’s claim about the semantics of disagreement has not been borne out by Khoo and Knobe’s (2018) experimental research. In fact, their findings suggest that contrary to what was previously thought, many people think that there can be moral disagreements in which none of the disagreeing parties are mistaken. As a result, if we want to accommodate people’s face-value commitments, there is pressure to justify the intuition that non-exclusionary moral disagreement is possible.
- 5) Whether paradigmatic moral facts appear to be different from paradigmatic natural facts may depend on what one regards as paradigm cases. Therefore, in the absence of further argument, not much argumentative weight can be attached to single examples.

- 6) That moral evaluation is generally regarded as the evaluation of goodness *simpliciter* is not brought out by experimental findings. In fact, current research suggests that various contextual features are relevant to moral evaluation (e.g. Sarkissian et al. 2011; Fisher et al. 2017), which puts some pressure on Enoch's claim.

6.5.2 Criticism of *Premise 2*

The face-value commitments of moral discourse can only provide support for moral realism if realists are best able to account for these findings. If rival theories are able to account for them equally well, then the presumptive support for moral realism is undermined. Traditionally, the purported objectivism of moral discourse has been cited as the most obvious face-value commitment in support of realism. However, antirealist proposals to justify this purported objectivism have also been advanced (Hopster 2017). Therefore, further argument is needed to get the presumptive argument in support of realism off the ground.

Importantly, realism *would* have a clear advantage over rival theories if the distinctive metaphysical commitment of realism – the existence of mind-independent properties – were presupposed by ordinary moral discourse and practice. But thus far experimental findings have provided no indication that the face-value commitments of moral discourse and practice support a distinctively realist metaphysics. Indeed, as Kirchin (2003, pp. 249-51) argues, it is implausible that the existence of mind-independent moral facts is part of the contents of our moral experience. Similarly, as Sinclair (2012, p. 168) observes:

Everyday moralisers seldom show appreciation of metaethical issues and categories, let alone appreciation of a particular metaethical theory. (...) Everyday moralisers hold that some things are right, others wrong, that some moral judgments are true, that there are correct answers to moral questions, that moral disagreement is possible and so on. But these are not categories that define realism. What defines realism is the view that moral judgments have a characteristic linguistic function, express states of mind with a characteristic representational function and (therefore) that their truth consists in correspondence between the representational content of such states and the moral way of the world.

On this more technical characterization, we could properly distinguish realist commitments from the commitments of rival views. But it is implausible that realism in this more technical sense is presupposed by moral discourse.

6.5.3 Criticism of *Premise 3*

Even if our moral discourse were clearly realist seeming, it is an open question whether our discourse is justified.⁵⁸ As previously stated, people's metaphysical intuitions may be

⁵⁸ As Sinclair (2012, p. 163) observes, it is plausible that large parts of moral discourse and practice are *pragmatically* justified. That is, we have good reason to go on engaging in this practice, as it allows us to relate to the world and to others in worthwhile ways. Moreover, the fact that practically all known human societies have developed some sort of moral practice suggests that this practice has clear functional benefits. But none of this gives us reason to think that our moral practice is also *epistemically* justified – that it allows us to relate to the world and each other in a veridical way.

erroneous. Indeed, it is not obvious why we should think that common intuitions about the (non-)existence of mind-independent moral properties should be on-track. Argument is required as to why people's commitments constitute any argumentative weight – why metaethicists should accommodate face-value intuitions in the sense of *justifying* them. After all, these are merely pre-theoretical intuitions concerning a subject matter for which the importance of pre-theoretical intuitions is not evident. Plausibly it's not much of an intellectual cost to rid ourselves of such intuitions and replace them with a reflective understanding of moral metaphysics.

Consider a weaker version of premise 3: conformity to our metaphysical intuitions provides *some* inductive support for a metaethical view. This weaker premise may be defensible: just like the plausibility of a scientific claim might be somewhat enhanced if it conforms to our scientific intuitions, a metaphysical claim might derive some support from conformity to intuition. The extent to which it does, however, crucially depends on the issue under consideration. For instance, ordinary people's intuitions about the behaviour of particles at the micro level should carry extremely little weight for scientists studying micro-level phenomena: after all, there is little reason to think that such intuitions are truth-apt. How about metaphysical intuitions concerning the reality of mind-independent moral properties? Perhaps such intuitions may carry more weight, for instance if these properties are directly perceptible. But assuming that they are is a metaethical assumption, which lies at the heart of the debate between realists and antirealist. It would be question-begging to rely on this assumption in defence of a presumptive argument for moral realism. If people's metaphysical intuitions are products of philosophical argument, then the metaethical *desideratum* is not to capture these intuitions; rather, it is to evaluate the cogency of the argument that generates them.

6.6 The Psychology of Metaethics

The findings highlighted in **section 6.2** have generated new questions about the cognitive processes associated with metaethical judgments as well as with the evolutionary and psychosocial functions that metaethical commitments serve. These questions fall in the domain of what we might call the psychology of metaethics. In this section I highlight some of the hypotheses that this new area of research has given rise to. They give an impression of the various avenues of theorizing triggered by research on folk moral objectivism – research avenues that go beyond the traditional subject matter of metaethicists but are nonetheless likely to attract philosophical interest.

One hypothesis, formulated by Wright and colleagues (2013, 2014, forthcoming), addresses the widespread intrapersonal variation with regard to the objectivity assessment of moral judgments. Wright et al. hypothesize that judgments about moral objectivity regulate how tolerant individuals and communities are of divergent moral judgments. If people think that a judgment is not to be tolerated, they adopt an objectivist stance; if they are willing to enter into discussion about the judgment, they ground the judgment subjectively. This tendency is particularly clear with regard to moral issues – and much less so, for instance, for issues that are regarded as matters of convention or taste. As the authors explain, the capacity to view moral issues pluralistically modulates

the level of permissible choice and dialogue about moral issues, both within and between sociocultural groups. Viewing a moral issue as objectively grounded removes it from the realm of legitimate personal/social negotiation (i.e., individual and/or social attempts to condone it will be deemed unacceptable, and censorship/prohibition will be supported). Viewing a moral issue as non-objectively grounded, on the other hand, allows people to acknowledge its moral significance (i.e., that it is not simply a personal matter), while at the same time maintaining room for choice, dialogue, and debate. (Wright et al. 2014, p. 31)

This hypothesis gains further support from Fisher et al.'s (2017) finding that a competitive mindset invites an objectivist construal of moral truth, whereas a cooperative mindset invites a subjectivist construal of moral truth. In a competitive setting, moral issues are taken out of the realm of open discussion, people's attitudes become less flexible and disagreeing parties are judged to be increasingly immoral. In a cooperative setting people become more willing to consider different moral viewpoints.

Naturally, metaethical judgments may serve more than one psychosocial function. Consider the earlier-mentioned relation between increased social consensus concerning a moral truth and increased objectivism concerning that truth – a relation that is, in fact, causal (Goodwin and Darley 2012; Beebe 2014). Perhaps, apart from signalling that a judgment is not to be tolerated, judging a moral issue to be objective can also serve as a means to signal allegiance to a certain in-group.

But the former signalling hypothesis merits special interest, as it ties in with recent hypotheses (DeScioli and Kurzban 2013; Stanford 2018) from evolutionary psychology. If an agent signals that a certain moral judgment is not to be tolerated, this might additionally be taken as a public signal that she is a partner who can be relied on to cooperate with regard to that issue. After all, the agent thereby signals a willingness to further a given moral goal and not to cooperate with agents who do not. This signal is at the basis of Stanford's (2018) recent hypothesis about the evolutionary origins of the phenomenological perception of morality's objectivity. Stanford argues, in part on the basis of the research outlined in **section 6.2**, that 'externalizing' moral demands has played a key role in the evolution of human cooperation. Following Joyce (2006a), he observes that moral demands have a distinct phenomenology: they often come across as if they are imposed on us externally and as if we have to comply with these demands regardless of our subjective preferences and desires. How should we account for this feature of our moral experience? Stanford argues that externalizing our moral motivation was favoured in ancestral environments because it allowed prosocial, altruistic and cooperative agents to quickly and efficiently correlate their interactions, paving the way for a more cooperative form of social life:

[E]xperiencing moral demands and obligations as externally imposed simultaneously on both ourselves and others ensures that if I myself come to be motivated to conform to a particular norm or standard of behavior that I experience as distinctively moral in character, I *automatically* demand that others conform to it as well, judging them to be less attractive potential partners in social interaction generally if they do not. (...) From an evolutionary point of view, our characteristic externalization of moral motivation

thus represents a mechanism for establishing and maintaining *correlated interaction under plasticity*. (Stanford 2018, pp. 8-9, his italics)

This hypothesis belongs to the realm of evolutionary moral psychology; like the hypotheses about the psychosocial function of metaethical judgments mentioned above, it has no *direct* implications for the conceptual and metaphysical issues that tend to be at the heart of metaethical debate. Stanford (*idem*, p. 45) himself considers the suggestion that his account might constitute an error theory of morality, but concludes it does not; he contends that people's ordinary views about moral objectivity are insufficiently articulate to be proven false.

Nonetheless, Stanford's hypothesis, if correct, *does* add to the psychological challenge to moral realism that I have outlined in the previous section. As noted, several metaethicists, such as Enoch (2018), support their case for moral realism with appeals to intuition, phenomenology and the seemingly objectivist features of moral discourse. Allegedly, these features are best explained on the assumption that moral realism is true. Indeed, this suggestion is what fuels the presumptive support for realism. Suppose that Enoch's phenomenological observation is apt, and that it generalizes: people generally perceive moral demands as highly objective, as if they were imposed on them externally. Stanford's evolutionary account suggests that explaining the characteristic objective-seeming phenomenology of moral discourse can be explained in purely evolutionary terms. While in and by itself this does not discredit moral realism, it *does* undercut any presumed metaethical plausibility points that realist theories might have been thought to have, in virtue of the fact that they vindicate morality's seeming objectivity. Indeed, this seeming objectivity may be little more than the culmination of various cognitive mechanisms, biases and illusions placed on us by evolution.

As this example suggests, there may be interesting ways in which psychological findings can touch upon traditional metaethical theses. From a philosophical viewpoint, we might have expected these realms to be clearly distinct, at least where the tenability of largely metaphysical doctrines – such as moral realism – is concerned. But since realists commonly defend their views by appealing to empirical or phenomenological claims, there turns out to be some overlap between them. Further exploring this area of overlap – where it lies, and whether experimental work might touch upon philosophical arguments – is among the future tasks of the psychology of metaethics.

6.7 Conclusion

I have shown that the recent wave of experiments involving folk moral objectivism has opened up various avenues of research in moral psychology, leading to novel hypotheses about the psychological function of our metaethical commitments. Moreover, I have argued that, notwithstanding the mismatch between the experimenter's and the metaethicist's typical objects of study, this research has promise as an important resource of empirical insight that touches on traditional metaethical concerns. Whether this promise can be fully delivered is, as of yet, an open question which depends *inter alia* on the success of experimenters in correcting for existing methodological flaws. But its current limitations notwithstanding, this research is beginning to fill an empirical hiatus and has already delivered some suggestive findings which may be indirectly relevant for analyses in moral semantics and moral metaphysics. Specifically, the findings provide resources that can be used to challenge metaethical arguments that heavily

rely on the alleged objectivist commitments of ordinary moral discourse, such as Mackie's argument for error theory or Enoch's presumptive argument in support of moral realism. Also, they serve as a reminder to philosophers that concerns about moral objectivity are more encompassing than concerns about moral realism. There may be other notions of objectivity at play in moral discourse which merit the attention of both psychologists and metaethicists.

References

- Beebe, J.R. 2014. How Different Kinds of Disagreement Impact Folk Metaethical Judgments. In: J.C. Wright and H. Sarkissian, eds. *Advances in Experimental Moral Psychology: Affect, Character, and Commitments*. London: Continuum, 167-187.
- Beebe, J.R. 2015. The Empirical Study of Folk Metaethics. *Etyka* 15: 11-28.
- Beebe, J.R. and Sackris, D. 2016. Moral Objectivism across the Lifespan. *Philosophical Psychology* 29: 912-929.
- Bengson, J. 2013. Experimental Attacks on Intuitions and Answers. *Philosophy and Phenomenological Research* 86: 495-532.
- Brink, D. 1989. *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- DeScioli, P. and Kurzban, R. 2013. A Solution to the Mysteries of Morality. *Psychological Bulletin*, vol. 139: 477-496.
- Enoch, D. 2018. Non-Naturalistic Realism in Meta-Ethics. In: T. McPherson and D. Plunkett, eds. *The Routledge Handbook of Metaethics*. New York: Routledge, 29-42.
- Feltz, A. and Cokely, E.T. 2008. The Fragmented Folk: More Evidence of Stable Individual Differences in Moral Judgments and Folk Intuitions. In: B.C. Love, K. McRae and V.M. Sloutsky, eds. *Proceedings of the 30th Annual Conference of the Cognitive Science Society*. Austin: Cognitive Science Society, 1771-1776.
- Fisher, M., Knobe, J., Strickland, B. and Keil, F.C. 2017. The Influence of Social Interaction on Intuitions of Objectivity and Subjectivity. *Cognitive Science* 41: 1119-1134.
- Fraser, B. 2014. Moral Error Theories and Folk Metaethics. *Philosophical Psychology* 27: 789-806.
- Gill, M.B. 2009. Indeterminacy and Variability in Meta-Ethics. *Philosophical Studies* 145: 215-234.
- Goodwin, G.P. and Darley, J.M. 2008. The Psychology of Meta-Ethics: Exploring Objectivism. *Cognition* 106: 1339-1366.
- Goodwin, G.P. and Darley, J.M. 2010. The Perceived Objectivity of Ethical Beliefs: Psychological Findings and Implications for Public Policy. *Review of Philosophy and Psychology*. 1: 161-188.
- Goodwin, G.P. and Darley, J.M. 2012. Why Are some Moral Beliefs Perceived to be More Objective than Others? *Journal of Experimental Social Psychology* 48: 250-256.
- Haidt, J., Koller, S.H. and Dias, M.G. 1993. Affect, Culture and Morality, or Is It Wrong to Eat Your Dog? *Attitudes and Social Cognition* 65: 613-628.
- Hannon, M. Forthcoming. Intuitions, Reflective Judgments, and Experimental Philosophy.

Synthese.

Heiphetz, L. and Young, L.L. 2017. Can only one Person be Right? The Development of Objectivism and Social Preferences regarding Widely Shared and Controversial Moral Beliefs. *Cognition* 167: 78-90.

Hopster, J.K.G. 2017. Two Accounts of Moral Objectivity: From Attitude-Independence to Standpoint-Invariance. *Ethical Theory and Moral Practice* 20: 763-780.

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

Joyce, R. 2006b. Metaethics and the Empirical Sciences. *Philosophical Explorations* 9: 133-148.

Joyce, R. 2015. Moral Anti-realism versus Realism: Intuitions. *Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/moral-anti-realism/moral-realism-intuitive.html>

Joyce, R. 2016. *Essays in Moral Skepticism*. Oxford: Oxford University Press.

Judd, C. M., Westfall J. and Kenny, D.A. 2012. Treating Stimuli as a Random Factor in Social Psychology: A New and Comprehensive Solution to a Pervasive but Largely Ignored Problem. *Journal of Personality and Social Psychology* 103: 54-69.

Kauppinen, A. 2007. The Rise and Fall of Experimental Philosophy. *Philosophical Explorations* 10: 95-118.

Khoo, J. and Knobe, J. 2018. Moral Disagreement and Moral Semantics. *Noûs* 52: 109-143.

Kirchin, S. 2003. Ethical Phenomenology and Metaethics. *Ethical Theory and Moral Practice* 6: 241-264.

Knobe, J. 2016. Experimental Philosophy Is Cognitive Science. In: J. Sytsma and W. Buckwalter, *A Companion to Experimental Philosophy*.

Laskowski N. and Finlay, S. 2017. Conceptual Analysis in Metaethics. In: T. McPherson and D. Plunkett, eds. *Routledge Handbook of Metaethics*. New York: Routledge, 536-551.

Loeb, D. 2007. The Argument from Moral Experience. *Ethical Theory and Moral Practice* 10: 469-484.

Loeb, D. 2008. Moral Incoherentism: How to Pull a Metaphysical Rabbit Out of a Semantic Hat. In: W. Sinnott-Armstrong, ed. *Moral Psychology: The Cognitive Science of Morality*, vol. 2. Cambridge: MIT Press, 355-385.

Ludwig, K. 2010. Intuitions and Relativity. *Philosophical Psychology* 23: 427-445.

Mackie, J.L. 1977. *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin Books.

Nichols, S. 2004. After Objectivity: An Empirical Study of Moral Judgment. *Philosophical Psychology* 17: 5-28.

Pözlner, T. 2017. Revisiting Folk Moral Realism. *Review of Philosophy and Psychology* 8: 455-476.

Pözlner, T. 2018. *Moral Reality and the Empirical Sciences*. New York: Routledge.

Pözlner, T. Forthcoming. Can the Empirical Sciences Contribute to the Realism / Antirealism Debate? *Synthese*.

Ruse, M. 1995 [1986]. Evolutionary Ethics: A Phoenix Arisen. In P. Thompson, ed. *Issues in Evolutionary Ethics*. Buffalo, NY: SUNY Press, 225-247.

Sarkissian, H., Parks, J., Tien, D., Wright, J.C. and Knobe, J. 2011. Folk Moral Relativism.

Mind & Language 26: 428-505.

Sarkissian, H. 2016. Aspects of Folk Morality: Objectivism and Relativism. In: W. Buckwalter and J. Sytsma, eds. *A Companion to Experimental Philosophy*. London: Blackwell, 212-224.

Sarkissian, H. 2017. Folk Platitudes as the Explananda of Philosophical Metaethics: Are They Accurate? And Do They Help or Hinder Inquiry? *Journal of Indian Council of Philosophical Research* 34: 565-575.

Shafer-Landau, R. 2003. *Moral Realism*. Oxford: Oxford University Press.

Sinclair, N. 2012. Moral Realism, Face-Values and Presumptions. *Analytic Philosophy* 53: 158-179.

Stanford, P.K. 2018. The Difference Between Ice Cream and Nazis: Moral Externalization and the Evolution of Human Cooperation. *Behavioral and Brain Sciences* 41: 1-57.

Street, S. 2009. In Defense of Future Tuesday Indifference: Ideally Coherent Eccentrics and the Contingency of What Matters. *Philosophical Issues* 19: 273-298.

Theriault, J., Waytz, A., Heiphetz, L., and Young, L. 2017. Examining Overlap in Behavioral and Neural Representations of Morals, Facts, and Preferences. *Journal of Experimental Psychology* 146: 1586-1605.

Turiel, E. 1983. *The Development of Social Knowledge: Morality and Convention*. Cambridge: MIT Press.

Weinberg, J.M., Alexander, J., Gonnerman, C. and Reuter, S. 2012. Restrictionism and Reflection: Challenge Deflected, or Simply Redirected? *The Monist* 95: 200-222.

Wright, J.C., Grandjean, P.T. and McWhite, C.B.. 2013. The Meta-Ethical Grounding of our Moral Beliefs: Evidence for Metaethical Pluralism. *Philosophical Psychology* 26: 336-361.

Wright, J.C., McWhite, C.B. and Grandjean, P.T. 2014. The Cognitive Mechanisms of Intolerance: Do our Meta-Ethical Commitments Matter? In: T. Lombrozo, J. Knobe and S. Nichols, eds. *Oxford Studies in Experimental Philosophy, vol. 1*. Oxford: Oxford University Press, 28-61.



Chapter 7

Evolution and the Perceived Objectivity of Moral Judgments

To be submitted for publication in a special volume based on the Oxford Brookes conference ‘Evolutionary Debunking Arguments: The Nuts and Bolts Approach’ (20–21 July 2018).⁵⁹

Keywords

Experimental metaethics; folk moral objectivity; determinants; personal interest; signalling

Abstract

Michael Ruse has argued that moral objectivity is an illusion that our genes have instilled in us to make us good social cooperators. We argue that Ruse's argument has philosophical shortcomings and is only partially corroborated by experimental evidence. That said, new scientific hypotheses about the psychosocial and evolutionary functions of objectifying morality do give impetus to a reassessment of whether moral realism can be challenged on the basis of evolutionary moral psychology. In this chapter we provide such a reassessment. We do so in the context of presenting the results of an experiment that we have conducted about the determinants of 'perceived moral objectivity'. Our findings show a positive relation between respondents' ratings of the 'personal importance' of a moral issue, and its 'perceived objectivity'. We emphasize that in itself, this result has no metaethical implication. However, we proceed to discuss this result in the context of other recent findings about the psychosocial and evolutionary functions of moral judgment and argue that these findings *do* lend themselves to a metaethical challenge against realism. The challenge differs from Ruse's argument, however: rather than *debunking* realism, we argue that these findings serve to overturn certain *desiderata* of metaethical analysis in a realist-unfriendly way.

⁵⁹ This chapter is based on collaborative work with Michael Klenk. Our division of labour was as follows: Michael and I were both involved in formulating the research hypothesis and setting up the experiment discussed in this chapter (**section 7.3**). Subsequently, Michael collected and analysed the data; I connected our results with recent work about the function of moral judgment and made our presentation suitable for submission to the Oxford Brookes conference *Evolutionary Debunking Arguments: The Nuts and Bolts Approach* (20–21 July 2018). **Subsection 7.2.1** and the experimental part (**section 7.3**) of this chapter were primarily drafted by Michael; the other sections were primarily drafted by me. The chapter will be submitted for a special volume, edited by Helen de Cruz and Johan de Smedt, based on the proceedings of the aforementioned conference *Evolutionary Debunking Arguments: The Nuts and Bolts Approach*.

'As we learn more about our moral psychology, we put ourselves in a position to develop more empirically informed debunking arguments.'

– Shaun Nichols (2014)

7.1 Introduction

7.1.1 Perceived Moral Objectivity: Its Relevance for EDAs

Many evolutionary debunking arguments (EDAs) in metaethics target moral realism, or some version thereof (e.g. Ruse 1986; Street 2006).⁶⁰ Some of them do so by appealing to the perceived objectivity of moral judgments and by arguing that this perceived objectivity is best explained in non-realist terms. For instance, Michael Ruse maintains that our moral experience is underwritten by a strong claim to objectivity:

An important part of the phenomenological experience of substantive ethics is not just that we feel that we ought to do the right and proper thing, but that we feel that we ought to do the right and proper thing because it truly is the right and proper thing. (Ruse, 2010, p. 309)

According to Ruse, we typically perceive moral demands as having a strong authority which does not depend on our own inclinations or even on those of any other human agent: moral demands seem to have an external authority. This apparent external authority, and the strong feeling of obligation to which moral demands give rise, are components of what we will call morality's *perceived objectivity*.⁶¹

Thinking of moral demands as being externally imposed may serve a motivational function. As has been documented by research in social psychology (cf. Nichols 2004, p. 183), when we perceive a demand as objective, we are typically more inclined to comply with it. Complying with moral demands, in turn, is typically evolutionarily advantageous because of the cooperation benefits it serves. Hence, the perceived objectivity of moral demands may be a proximate mechanism to foster fitness-enhancing social interactions.⁶²

Ruse, however, argues that our objectivist moral commitments cannot be substantiated. He contends that it has been evolutionarily beneficial for our ancestors to think that morality

⁶⁰ As elsewhere in this dissertation, here moral realism is understood as the metaethical position that moral statements can be objectively true or false and that this claim to objectivity should be understood in terms of the existence of a stance-independent moral reality to which true moral statements correspond. Note that the aim of Street's (2006) article is to raise the plausibility of antirealism (the view that moral truths are *stance-dependent*) with respect to realism. Ruse's (1995) self-proclaimed target is the view that morality is 'objective' in the sense of having foundations that are external to human beings. The differences between Ruse's and Street's positions notwithstanding, it is harmless to identify the (sub)target of both of their arguments as *moral realism*.

⁶¹ It should be borne in mind that the notion of moral objectivity can be understood in different ways. In this chapter, we are specifically interested in a *psychological* conception of moral objectivity rather than a *metaethical* conception. The perception that moral demands are externally imposed is often taken to be a defining characteristic of this psychological conception (e.g. Stanford 2018).

⁶² As we point out in **section 7.4.1**, however, the motivational function of objectifying moral demands may not be as clear-cut as this picture suggests.

has objective foundations external to us, but that in fact no such objective foundations exist. Our perception that moral demands are objective is merely an illusion that our genes have instilled into us in order to make us good social cooperators. Ruse himself ultimately endorses an error theory of moral judgment: all moral judgments are false.

7.1.2 The Tenability of Ruse's EDA: Philosophical and Empirical Issues

Ruse's EDA can be challenged on philosophical grounds. Even if he is correct to claim that we are biologically prone to perceive moral demands as having an external, mind-independent foundation, this does not indicate that this belief cannot be justified on metaethical grounds. To argue, as Ruse does, that postulating the existence of mind-independent foundations is explanatorily redundant, since we can also explain our objectivist commitments in evolutionary terms, need not be regarded as problematic by moral realists. Realists may counter, for instance, that 'figuring in our best scientific explanations' is not a necessary condition for vindicating existence claims: even if mind-independent moral standards do not figure in explanations of morality's origins, postulating their existence may still be indispensable for other, non-explanatory projects, such as practical deliberation (Enoch 2011): the project of deliberating how to act, which goals to pursue and what values to endorse.

Moreover, Ruse's argument can also be challenged on empirical grounds. There are two empirical claims underlying his EDA:

- People perceive moral judgments as objective in the realist sense of having an external foundation that does not depend on the attitudes of any human agent;
- This trait – morality's perceived objectivity – is an evolutionary adaptation.

Recent experimental work in moral psychology undermines, to some extent, the first claim. Experimental findings from a newly emerging subfield of moral psychology – the psychology of metaethics – suggest that there is considerable inter- and intra-subject variance regarding the perceived objectivity of moral judgments (see **chapter 6** of this dissertation). That is, not all people judge a moral proposition like 'killing is wrong' to be objectively true or false, and some people who judge that 'killing is wrong' is objectively true or false do not think the same about other moral propositions, like 'consensual sex with siblings is morally wrong' (cf. Goodwin and Darley 2008, 2012; Sarkissian et al. 2011; Wright et al. 2013; Beebe 2014; Fisher et al. 2017). These experimental findings indicate that people's commitment to moral objectivity is less clear-cut than Ruse assumes, and suggest that the ambition of an antirealist EDA on the basis of the perceived objectivity of moral judgments should be moderated. A more nuanced version of such an EDA may proceed from the observation that *some* moral judgments are perceived as having an external authority, and may purport to argue that *to the extent* that people have this impression, they are mistaken.

On the other hand, recent work in evolutionary moral psychology *does* provide support for Ruse's second empirical claim: his conjecture that perceiving moral judgments as objective is an evolutionary adaptation. Indeed, evolutionary psychologists have recently come up with fairly fine-grained hypotheses about the evolutionary function of moral objectivism, suggesting that perceiving moral judgments as objective may be functional in more intricate ways than Ruse himself proposed (DeScioli and Kurzban 2013; Stanford 2018). While this does not help

to salvage the objections to Ruse's EDA, this development does give impetus to a reassessment of whether findings about the psychosocial and evolutionary functions of perceiving moral demands as objective might have any implication for the tenability of moral realism. We will return to this question at the end of this chapter (section 7.5).

7.1.3 Outline of the Chapter

Before turning to metaethics, in the ensuing sections of this chapter we will first discuss in greater detail recent findings and theories about the psychology, sociology and evolution of moral objectivism. We will do so, in part, by presenting the results of our own experimental work. This work starts from the hypothesis that the observed variance in the perceived objectivity of moral judgments is at least partly explained by the personal interests of an individual or group: if people find the truth or falsity of a moral issue important, then they are more prone to consider judgments concerning that issue as objective. Thus, we hypothesize that *personal interest* is one of the determinants of perceived moral objectivity.⁶³

Additionally, we will discuss this hypothesis in the context of recent theories about the psychosocial and evolutionary functions of moral judgment. We expand on these existing theories by proposing that presenting moral demands as objective may serve a signalling function: objectifying a moral issue serves as a public signal that one is willing to stand one's ground with regard to this issue. We relate this signalling hypothesis to recent proposals about the evolutionary origins of the perceived authority of moral demands and specifically to the hypothesis that a key evolutionary function of moral judgment is to strategically choose sides in cases of conflict (DeScioli and Kurzban 2013).

We proceed, in section 7.2, by spelling out in further detail our experimental hypothesis. In section 7.3 we present the results of the experiment that we set up to test this hypothesis and discuss some methodological limitations of our experiment. In section 7.4 we discuss our findings in the context of recent hypotheses about the psychosocial and evolutionary functions of moral judgment. In section 7.5 we return to the EDA debate and argue that advances in experimental moral psychology and evolutionary psychology may indeed pose a challenge to moral realism, albeit in a different way than Ruse proposed: these advances may serve to overturn the 'objectivist *desiderata*' of metaethical analysis in a realist-unfriendly way.

7.2 Determinants of Perceived Moral Objectivity: Our Experimental Hypothesis

⁶³ Discussion of the 'perceived objectivity' of moral judgments is common in the experimental metaethics literature (e.g. Goodwin and Darley 2012) and we will follow this usage here. We grant, however, that such discussion may be misleading in the context of our experimental hypothesis: whether test subjects indicate that a moral issue is objective need not be a matter of *perception* but may only be a matter of *presentation*. Our experiment is mute on whether subjects merely present issues as objective or also perceive them as such; we scrutinize the correlates of respondents' ascriptions of objectivity, irrespective of their underlying mechanisms. Readers who feel uneasy with the label 'perceived objectivity' may substitute it with 'ascribed objectivity', which remains neutral on the question of whether moral demands are merely presented as objective or are also perceived as being so.

What makes people regard or present a moral issue as objective, in the sense of having a strong, subject-independent authority? This question has recently gained the attention of experimental moral psychologists. Some have found evidence that social consensus is an important factor in determining people's objectivist commitments. For instance, Goodwin and Darley (2012) presented undergraduates with false information about the percentage of students from the same institution who agreed with them about the correctness of certain moral statements. Respondents presented with low consensus estimates were significantly less likely to regard statements as objective than respondents presented with high consensus estimates. Similarly, Beebe (2014) found that having respondents reflect upon the extent of societal disagreement about moral statements decreased their attributions of objectivity regarding those statements. These findings may lead to the impression that conformity bias plays an important role in ascriptions of moral objectivity: in judging a moral issue to be objective, people simply follow the majority vote.

Neither of these studies, however, suggests that social consensus is the *only* determinant of perceived objectivity. Indeed, there may be various such determinants – and while we grant that perceived social consensus may be one of them, we doubt that it is the *principal* determinant. In our research we will scrutinize the importance of another possible determinant: the *personal interest* of moral agents. We hypothesize that if subjects regard a moral issue as being of personal importance, then this will positively influence their rating of the statement's objectivity.

7.2.1 Social versus Moral Norms

One impetus for looking for further determinants of the perceived objectivity of moral judgments comes from social scientific research on social and moral norms (Bicchieri 2006). Social scientists have found that moral norms typically differ from social norms in that the former, but not the latter, are independent of a subject's expectations about the behaviour and social expectations of others (Bicchieri 2006, p. 29). That is, a social norm like 'dress in black at a funeral' exists because a sufficient number of subjects within a group think that a sufficient number of other group members a) behave accordingly and b) expect others to behave accordingly (and reprimand them if they don't). A moral norm, by contrast, exists independently of these expectations. This contrasts with the proposal that regarding a moral judgment as objective is a function of social consensus, and gives some impetus to search for other determinants.

7.2.2 Self-Serving Bias

Another impetus for considering our hypothesis comes from extant work on the importance of personal interest and self-serving bias in making moral judgments. Various studies indicate that people often strategically advocate moral rules that serve their personal interests and benefit themselves more than others (see DeScioli and Kurzban 2013, p. 48 and DeScioli 2016, p. 25 for references).⁶⁴ For example, it has been found that people generally approve of having

⁶⁴ Personal interest is, of course, a broad category. A person's interests are subject-dependent and can be manifold. In what follows we will use personal interest in this broad and dynamic sense. Note that we focus on a *psychological* conception of personal interest, rather than, for example, an *evolutionary*

driverless cars designed in such a way that they minimize overall harm (e.g. by preventing the car from hitting pedestrians), but that when it comes to buying their own car, the same people will opt for cars designed to save the passengers at all costs (Bonnefon et al. 2016). Our hypothesis is in line with the hypothesis that personal interest influences people's moral judgments. Its novelty is that it extends this hypothesis to the metaethical realm: we hypothesize that personal interests can also influence the 'metaethical status' that people ascribe to moral judgments by attributing greater objectivity to them.

7.3 Experiment

7.3.1 Study

To perform a first test of our hypothesis, we conducted a survey to compare objectivity ratings and importance ratings on a range of moral issues in a 5 (importance rating, 5-point Likert scale anchored at 'Extremely Important' and 'Not Important') \times 3 (objectivity score, operationalized according to the objectivity paradigm of Goodwin and Darley 2008) within-subject design. Subjects were presented with 35 statements, mostly concerning moral issues, such as 'discriminating against someone on the basis of race is unacceptable' (see **appendix**). Question categories (factual, moral, taste, religious) were assigned by the experimenters, but partly informed by the study of Wright et al. (2013), in which subjects self-identified moral issues.

Subjects rated each statement according to how important they found the truth or falsity of the statement and indicated whether the statement was true, false or just an opinion or attitude, and whether another person who disagreed with them about the truth or falsity of the judgment had to be wrong or whether both could be right. The order of individual questions as well as the order of the three question segments (truth, disagreement and importance) was randomized for each subject. The order of presentation had no significant effect on objectivity scores and/or the relation to importance. The survey ended with a demographics question section, in which we asked subjects about their age, biological sex, political orientation (left, right, other) and religiosity (5-point Likert scale anchored at 'Definitely Yes' and 'Definitely No').

A pilot conducted in January 2018 ($n=22$) confirmed the understandability of the questionnaire and served as a coherence check of the planned experimental design. The study was then conducted via Amazon Turk ($n=150$). mTurkers were remunerated for the HIT (Human Intelligence Task) according to Amazon's guidelines for comparable tasks and independently of whether they completed the survey or not, thus ensuring that they could retreat from the survey without any repercussion. Excluding incomplete surveys or surveys that did not confirm the informed consent disclaimer reduced the valid surveys to $n=50$.

Stringent attention checks on three factual questions (Question IDs=1, 29, 32) and two moral questions (IDs=11 & 28, 26 & 27) excluded 29 participants.⁶⁵ Following Goodwin and

conception: our main indicator of personal interest is which statements subjects judge to be important rather than what contributes to their genetic fitness.

⁶⁵ The factual attention check excluded subjects who answered that people who disagreed about IDs 1 or 29 could be 'both right' or 'other' and who answered 29 with 'false' or 'neither true nor false'. Such answers indicate that the subject did not pay attention or failed to understand the idea that two discussants cannot both be right in response to a factual disagreement. The moral attention check

Darley (2008), ‘objectivity scores’ were computed as a categorical variable with 1 point for considering a statement to be true or false and 1 point for judging that one party must be wrong in a disagreement about the statement (hence, 2=max. objectivity score, 1=medium objectivity score, 0=subjectivist).

7.3.2 Results

Pairwise t-tests found a significant positive relation between the degree to which a subject judges a moral issue to be objective and how important the subject judges the truth or falsity of the proposition to be, with a considerable effect for medium scores ($t = 2.6822$, $df = 21$, $p\text{-value} = 0.01395$) and a strong effect for high objectivity scores ($t = 7.4187$, $df = 21$, $p\text{-value} = 2.7e-07$). Figure 1 compares mean importance scores with categorical objectivity scores.

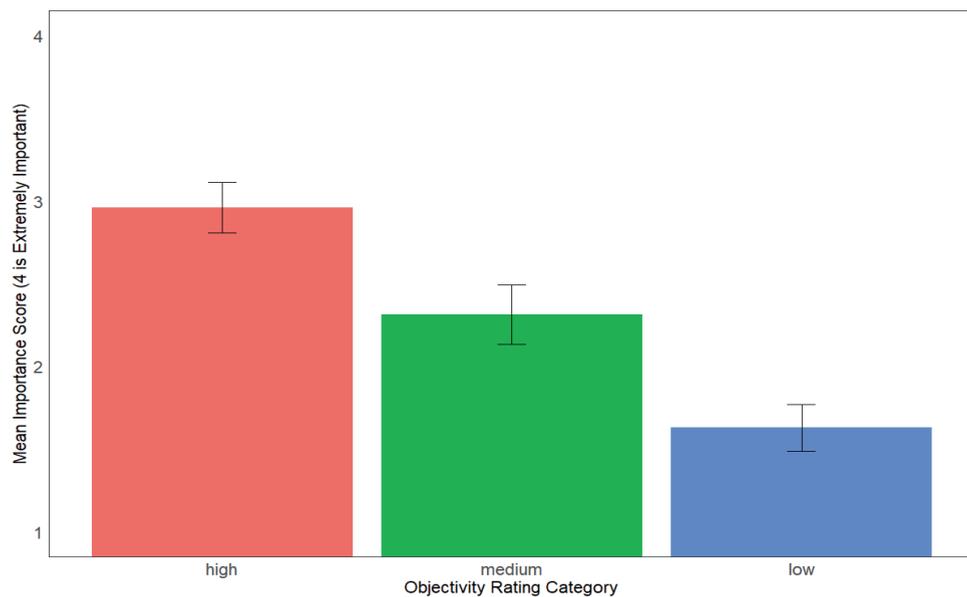


Figure 7.1: Mean Importance Score versus Objectivity Score

7.3.3 Methodological Implications

Our findings should allow for some reflection on Goodwin and Darley’s (2008) objectivity paradigm. Preliminary analyses suggest that measures of disagreement ratings explain importance nearly as well as computed objectivity scores. This raises the question of how to better disentangle extensions of the concept beyond its relation to disagreement for use in experimental studies. We recommend that future studies do not follow Goodwin and Darley in using an objectivity paradigm that combines questions about the truth of moral statements and about the possibility of disagreement because distinguishing between these two features fits better with extant metaethical distinctions.

excluded subjects who gave inconsistent answers to the question pairs 11 and 28 and 26 and 27, which asked about the permissibility of gay marriage and killing of innocents, respectively.

7.3.4 Limitations of the Study

Our findings provide support for the hypothesis that subjects who regard the truth of a statement as being of personal importance are prone to consider the truth or falsity of judgments concerning that statement as more objective. However, we are aware that our study has certain limitations and does not provide outright support for the hypothesis that greater personal interest is the cause of increased perceived objectivity. The following three caveats should be kept in mind.

First, our experiment leaves the relationship between personal interest and perceived objectivity unclear. It could be that perceived objectivity positively influenced the importance that a subject accorded to the truth of an issue, or the other way around, as we suspect. There are two ways to establish causal inference. The first is based on inductive inference. If many studies corroborate that there is a correlation between personal importance ratings and perceived objectivity ratings, this provides an inductive base for judging that there is *some* causal link between the two factors. It does not show, however, how the relation is directed. The second, more straightforward, way is to manipulate experimentally a subject's personal interest in the truth of a given statement and then see whether this increases the subject's perceived objectivity of the statement. Importantly, the manipulation must not in itself increase the perceived objectivity of a statement. One way forward on this question is to look to the literature on economic games to find a way of expediently manipulating a subject's personal interest in the truth of a statement.

Second, we assume that respondents do not regard personal importance – or at least not typically – as an indicator of social consensus. Instead, we assume that personal importance is understood as an indicator of 'personal interest' – broadly, whatever needs and desires individuals have, and whatever goals they seek to pursue. While we grant that occasionally conforming to social consensus may be among individuals' personal interest, it seems plausible that this is not typically the case. To disentangle 'personal interest' and 'social consensus' more thoroughly is one of the tasks for future work.

Third, we are aware of the general limitations of using survey studies as a means to investigate the metaethical commitments of folk respondents (see **chapter 6** of this dissertation). For instance, arguably the prompted answers to survey questions do not provide us with insight about people's endorsed metaethical commitments (cf. Bengson 2013). This criticism equally applies to many other studies about folk moral objectivism, such as those of Goodwin and Darley (2008, 2012). In what follows, we shall assume that this caveat does not fully undermine the relevance of existing survey-based research. But we do acknowledge that future studies should attempt to overcome this potential methodological pitfall.

7.3.5 Do Our Findings Provide Support for Specific Metaethical Views?

It should be noted that our characterization of objectivity leaves open a vast number of metaethical positions. We operationalized objectivity on an ordinal scale from no objectivity (0 points) via medium objectivity (1 point) to strong objectivity (2 points). For the medium level of objectivity, we do not distinguish between the two mixed positions: the 'truth-focused' medium objectivists, regarding a statement S, judged that the statement was either true or false but that disagreement about the truth of the statement need not involve mistaken discussants; the 'disagreement-focused' medium objectivists, regarding a statement S, judged that the

statement was neither true nor false but that at least one discussant in a disagreement about the truth of the statement had to be mistaken.

The coherence of either view depends on the theoretical commitments we ascribe to medium objectivists – neither view is obviously logically incoherent. The first type of medium objectivist says that moral statements are truth-apt but that moral disagreements can be faultless. Such a view is compatible with various types of moral relativism: the truth-makers of moral judgments are determined by a reference frame (such as the attitudes of an individual, the culture of a society, etc.). As a result, moral statements are true but disagreements are faultless, because disputants from different reference frames are talking past each other. The second type of medium objectivist says that moral statements are not truth-apt but that, nonetheless, moral disagreements cannot be faultless. On this view, differences in attitudes are sufficient for disagreement to exist, and there is a way to ascertain which attitudes are correct in any given disagreement. The view might be squared with, for example, views like Blackburn's (1993) quasi-realism, according to which moral statements are not truth-apt but state (quasi-)moral facts nonetheless. Of course, our experimental design allows us to see whether subjects are medium objectivists of type 1 or type 2, but it does not allow us to infer whether or not they took this view on the basis of metaethical reflection.

Apart from this concern, it should be clear that being an objectivist in our sense does not allow us to infer much about the metaethical construal of objectivism among our subjects. An objectivist in our sense could be a metaethical constructivist of both Kantian (e.g. Korsgaard 1996) and the procedural stripe (e.g. Habermas 1991). Both types of constructivism imply that within a group (variously defined), moral statements have truth-values and disagreements will involve mistaken discussants, thus satisfying both of our criteria. Furthermore, an objectivist in our sense could be a naturalist realist (e.g. Jackson and Pettit 1995) or a non-naturalist realist (e.g. Enoch 2011). Both of these positions entail the view that moral statements are true or false (satisfying our first criterion) and that at least one party in a moral disagreement will be mistaken (satisfying our second criterion).

Given the indeterminacy with regard to the metaethical background theory, the present study does not provide reason for or against any of these theories. What our study does suggest, however, is that there is a minimal psychological sense in which one can be an objectivist (understood as thinking about the truth and disagreement criteria outlined above) and that objectivism in this sense is not uncommon. When interpreting our results, it is crucial to keep this in mind: our evidence concerns a broad psychological sense of objectivism, and not objectivism in any specific metaethical sense.

7.4 General Discussion: The Psychosocial and Evolutionary Functions of Perceived Moral Objectivity

In the previous section we presented experimental evidence to show that personal interest is indeed one of the determinants of regarding moral statements as objective. In this section we turn to the question of what psychosocial function objectifying moral issues might serve and what this function might tell us about morality's evolutionary origins. While we think that objectifying moral demands may serve multiple psychosocial functions, we will highlight some of the functions that we think are particularly prevalent. We grant that some of these proposals

are somewhat speculative, and are not directly implied by our experimental work, but we will discuss them nonetheless, as we think that they have a good fit with our experimental findings and merit further theoretical reflection and experimental scrutiny.

7.4.1 What Is the Function of Externalizing Moral Demands?

Perhaps the first metaethicist to suggest that objectifying moral judgments serves a psychosocial function was John Mackie (1977). According to Mackie, it often serves interpersonal relations well to behave as morality requires: ‘We need morality to regulate interpersonal relations, to control some of the ways in which people behave towards one another, often in opposition to contrary inclinations. We therefore want our moral judgments to be authoritative’ (idem, p. 43). Think of situations which involve a tragedy of the commons: to coordinate action by referring to common moral values might help to solve social predicaments.

Why should moral values be regarded as *objective* to serve this function? Followers of Mackie, such as Joyce (2006), have argued that objectifying moral demands has a motivational upshot: feeling an internal sense of moral obligation motivates us to behave as morality requires. Perhaps this is *part* of the explanation, but as Stanford (2018) has recently argued, it cannot be the full story. We are also motivated by *subjective* experiences, such as the experience of pain or hunger, which we do not cast in terms of having an external authority. What additional function is served by presenting and perceiving moral demands as *objective*?

Wright et al. (2013, 2014) have formulated a hypothesis about the psychosocial function of moral objectivism that may answer this question. Building on experimental findings which suggest that subjects’ attitudes towards regarding moral issues as objective is quite flexible and diverse, they suggest that a capacity to modulate the objectivity ascribed to moral issues helps to coordinate dialogue and action both within and between sociocultural groups:

Viewing a moral issue as objectively grounded removes it from the realm of legitimate personal/social negotiation (...). Viewing a moral issue as non-objectively grounded, on the other hand, allows people to acknowledge its moral significance (...), while at the same time maintaining room for choice, dialogue, and debate. (Wright et al. 2014, p. 31)

We contend that, at least in part, Wright et al.’s hypothesis is probably along the right lines: presenting a moral demand as objective typically signals an unwillingness of the agent making the demand to easily change her mind about it. But we disagree that objectifying moral issues removes them from the realm of legitimate negotiation. More plausibly, we think, an ascription of objectivity serves to signal on which topics agents are willing to stand their ground in moral discussion. This function, in turn, is at least partly modulated by agents’ personal interests, as our experimental findings suggest. Hence, modulating the objectivity ascribed to moral issues has a self-serving component: presenting an issue as objective typically signals that agents are willing to stand their ground in moral debate, often with regard to issues which they judge to be of specific personal importance.

7.4.2 Moral Objectification: A Signalling Hypothesis

Our signalling hypothesis contrasts with the view that the function of objectifying moral judgments serves the function of moral persuasion. As Mercier and Sperber (2017) argue, discussions typically do not serve to grasp truths but to persuade other parties. We grant that the same might typically be true for moral discussions, but we do *not* think that this is the main function of presenting moral issues as *objective*. Rather than dialectical effectiveness – i.e. an ability to efficaciously change other people’s minds – we think that objectifying moral issues typically serves to signal that agents care about an issue and are not easily prepared to change their minds about it. In doing so they give others a choice: on this issue you’re either with me, or against me. Hence, presenting moral issues as objective may be an effective means to form moral alliances and to enhance in-group cohesion.

Our signalling hypothesis also contrasts with the view that moral objectification is mostly due to conformity bias. We do not deny that conformity bias may play a role in moral decision-making, but hypothesize that when issues are presented as objective, this is typically not due to an agent seeking conformity. Rather than siding with the majority, presenting a moral issue as objective serves to signal agents’ personal commitment to the issue and their willingness to stand their ground, even in the face of contrary moral viewpoints and regardless of the consensus view. Perhaps signalling this commitment is specifically beneficial if a moral view is a minority position: it allows fellow activists to rally support for specific moral issues *irrespective* of what the majority says.

7.4.3 Choosing Sides

Two of the features of moral objectification outlined in the previous subsection – objectifying moral issues often instigates third parties to choose sides, and often does so on the basis of the contents of specific *issues* rather than majority consensus – are congruent with a recent evolutionary hypothesis by DeScioli and Kurzban (2013), who have argued in detail in support of the thesis that the main evolutionary function of moral judgment is choosing sides in a conflict. Choosing sides is a strategic interaction problem that involves a perpetrator, a victim and a third-party condemner. Different solutions are possible, some which involve decision procedures that would never be advisable in making individual choices but may be beneficial in multiplayer interactions.

One possible solution is to choose sides on the basis of pre-existing alliances. However, this strategy has the downside of escalating conflicts, resulting in greater fighting costs. A second solution is to choose sides on the basis of power: third parties side with the most powerful party in the conflict. However, this strategy has the downside of concentrating power in the hands of high-status individuals and may lead to authoritarianism and exploitation. A third solution is to choose sides by focusing on the *contents* of an action. Some actions may be flagged as *never-to-be-done* or *categorically wrong*; this public signal can subsequently coordinate third-party judgment. This solution is characteristic of moral conflicts, in which bystanders typically coordinate which side they choose not on the basis of the *identities* of the disputants but on the basis of their *actions*. Indeed, some characteristic features of morality – its impartiality, non-consequentialist characteristics and the categorical nature of moral demands – can be explained according to this strategic interaction framework. In short, moral decision-making is a strategy that can be used to avoid the costs of escalating conflicts, to avoid despotism and to choose on the basis of the acceptability of certain *actions*.

DeScioli and Kurzban suggest that this side-choosing strategy has played an important role in moral evolution. Indeed, it seems to fit well with the view that morality emerged in egalitarian societies of hunter-gatherers, with their characteristic anti-dominance hierarchies (Boehm 2012). But as DeScioli and Kurzban (2013, pp. 488-491) point out, once morality had evolved, new strategies emerged to use moral rules to the advantage of individuals or groups. In a later phase of moral evolution, the dominant side-choosing strategy seems to have altered: moral alliances were less often formed on the basis of judging specific actions, and more often on the basis of existing relations. Especially in times of increased rioting and warfare, factors such as group loyalty and self-interest are likely to have become more pervasive. During the Holocene, choosing sides on the basis of existing religious or political power structures may have turned into a dominant strategy, as suggested by the rise of authoritarian moral regimes (cf. Sterelny 2016).

Our hypothesis fits well with that of DeScioli and Kurzban (2013) and extends it to the metaethical domain. We think that objectifying moral issues is instrumental to the function of choosing sides. For parties engaged in conflict, presenting an issue as objective serves as an effective instrument to force third parties to choose sides in the debate. For third parties, presenting an issue as objective may serve to signal their commitment to defending this issue or to signal their allegiance to a given group. Plausibly, the latter strategy became increasingly important during a rather recent stage of the evolution of *Homo sapiens*, when competition between different religious and moral tribes increased (Sterelny 2016). In this context, presenting moral issues as objective may have been co-opted as a mechanism to delineate in-group/out-group boundaries and deepen the division between them.

7.5 A Psychological Challenge for Moral Realism?

In this chapter we have highlighted recent experimental findings from moral psychology, including those of our own, as well hypotheses from moral psychology concerning the psychosocial and evolutionary functions of moral objectivism. Let's assume that these findings are empirically robust and that these hypotheses are along the right lines. Do they have any implication for metaethics? More specifically, do they give rise to a challenge for moral realism, which does not face the same objections as Ruse's EDA (**subsection 7.1.2**)?

We should emphasize that the findings of the experiment we have conducted do not give rise to such a challenge. As noted, our data can be interpreted in realist-friendly terms: realists may argue that the best explanation of our findings is that respondents regard some issues as more important than others because they know that these issues reflect objective moral truths (cf. **subsections 7.3.4** and **7.3.5**).

That said, in the context of the broader findings and hypotheses discussed in this chapter, including recent hypotheses about the evolutionary origins of our perceived sense of moral objectivity, we do think that work in evolutionary moral psychology provides resources for metaethicists who seek to challenge moral realism – not by debunking realism but by challenging so-called presumptive arguments in support of it (as discussed in **chapter 6.5**).

One of the main arguments that realists have traditionally provided in support of their view comes from moral phenomenology and appeals to the *seeming* objectivity of moral demands (e.g. Brink 1989; Enoch 2014). But as noted, recent experimental findings do not straightforwardly vindicate the assumption that moral demands are conceived as objective in

the realist's sense. Moreover, to the extent that moral demands *are* conceived as strongly objective, there are reasons for thinking that, at least in part, this is an artefact of our evolved psychology (Joyce 2006; DeScioli and Kurzban 2013; Stanford 2018). This does not establish that the objectivity of moral demands is merely an illusion, as Ruse argues, but it does serve to undercut the presumptive support that our moral phenomenology bestows upon moral realism – which, as mentioned, is one of the main arguments that realists have traditionally provided in support of their view.

Differently put, psychological findings may serve to undercut the validity of the *desideratum* that metaethicists should vindicate the 'objective seemings' of our moral experience by arguing that the variety of moral experiences found by experimenters conflicts with this *desideratum*. Moreover, to the extent that there *is* such a *desideratum* nonetheless, debunkers might argue that our perceived sense of moral objectivity can be explained in evolutionary terms. This suggests that, whether or not moral realism is ultimately tenable, at least it need not be regarded as the default position in metaethics, as its defenders often suggest (e.g. Enoch 2014). While this result does not debunk realism, it does change the dialectic of the metaethical debate.

7.6 Conclusion

We have presented experimental findings which indicate that there is a positive relation between the personal interest that subjects have in a moral issue and the objectivity they bestow on this issue. In the ensuing discussion section, we have framed these findings in the context of a hypothesis about the psychosocial and evolutionary functions of moral objectivity. Expanding on work by DeScioli and Kurzban (2013), we propose that presenting moral demands as objective may serve the function of forcing third parties to choose sides in moral debate. Our experimental findings and this hypothesis are first and foremost relevant for moral psychology rather than metaethics. However, we do think that findings along the lines we have presented may have implications for the EDA debate: they may serve to undercut presumptive arguments in support of moral realism.

References

- Beebe, J.R. 2014. How Different Kinds of Disagreement Impact Folk Metaethical Judgments. In: J.C. Wright and H. Sarkissian, eds. *Advances in Experimental Moral Psychology: Affect, Character, and Commitments*. London: Continuum, 167-187.
- Bengson, J. 2013. Experimental Attacks on Intuitions and Answers. *Philosophy and Phenomenological Research* 86: 495-532.
- Bicchieri, C. 2006. *The Grammar of Society*. Cambridge: Cambridge University Press.
- Blackburn, S. 1993. *Essays in Quasi-Realism*. Oxford: Oxford University Press.
- Boehm, C. 2012. *Moral Origins: The Evolution of Virtue, Altruism, and Shame*. New York: Basic Books.
- Bonnefon, J., Shariff, A. and Rahwan, I. 2016. The Social Dilemma of Autonomous Vehicles. *Science* 352: 1573-1576.

- Brink, D. 1989. *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- DeScioli, P. 2016. The Side-Taking Hypothesis for Moral Judgment. *Current Opinion in Psychology* 7: 23-27.
- DeScioli, P. and Kurzban, R. 2009. Mysteries of Morality. *Cognition* 112: 281-299.
- DeScioli, P. and Kurzban, R. 2013. A Solution to the Mysteries of Morality. *Psychological Bulletin* 139: 477-496.
- Enoch, D. 2011. *Taking Morality Seriously: A Defense of Robust Realism*. Oxford: Oxford University Press.
- Enoch, D. 2014. Why I Am an Objectivist about Ethics (And Why You Are, Too). In: R. Shafer-Landau, ed. *The Ethical Life*. Oxford: Oxford University Press, 192-205.
- Fisher, M., Knobe, J., Strickland, B. and Keil, F.C. 2017. The Influence of Social Interaction on Intuitions of Objectivity and Subjectivity. *Cognitive Science* 41: 1119-1134.
- Goodwin, G.P. and Darley, J.M. 2008. The Psychology of Meta-Ethics: Exploring Objectivism. *Cognition* 106: 1339-1366.
- Goodwin, G.P. and Darley, J.M. 2012. Why Are some Moral Beliefs Perceived to be More Objective than Others? *Journal of Experimental Social Psychology* 48: 250-256.
- Habermas, J. 1991. *Erläuterungen zur Diskursethik*. Frankfurt am Main: Suhrkamp Taschenbuch Wissenschaft.
- Jackson, F. and Pettit, P. 1995. Moral Functionalism and Moral Motivation. *The Philosophical Quarterly* 45: 20-40.
- Joyce, R. 2006. *The Evolution of Morality*. Cambridge: MIT Press.
- Korsgaard, C. 1996. *The Sources of Normativity*. Cambridge: Cambridge University Press.
- Mackie, J.L. 1977. *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin Books.
- Mercier, H. and Sperber, D. 2017. *The Enigma of Reason*. Cambridge: Harvard University Press.
- Nichols, S. 2004. *Sentimental Rules*. New York: Oxford University Press.
- Nichols, S. 2014. Process Debunking and Ethics. *Ethics* 124: 727-749.
- Ruse, M. 1995 [1986]. Evolutionary Ethics: A Phoenix Arisen. In: P. Thompson, ed. *Issues in Evolutionary Ethics*. Buffalo: SUNY Press, 225-247.
- Ruse, M. 2010. The Biological Sciences Can Act as a Ground for Ethics. In: F. Ayala and R. Arp, eds. *Contemporary Debates in Philosophy of Biology*. Wiley-Blackwell, 297-315.
- Sarkissian, H., Parks, J., Tien, D., Wright, J.C. and Knobe, J. 2011. Folk Moral Relativism. *Mind & Language* 26: 428-505.
- Stanford, P.K. 2018. The Difference Between Ice Cream and Nazis: Moral Externalization and the Evolution of Human Cooperation. *Behavioral and Brain Sciences* 41: 1-57.
- Sterelny, K. 2016. Cooperation, Culture, Conflict. *British Journal for the Philosophy of Science* 67: 31-58.
- Street, S. 2006. A Darwinian Dilemma for Realist Theories of Value. *Philosophical Studies* 127: 109-166.

Chapter 7

Wright, J.C., Grandjean, P.T. and McWhite, C.B. 2013. The Meta-Ethical Grounding of our Moral Beliefs: Evidence for Metaethical Pluralism. *Philosophical Psychology* 26: 336-361.

Wright, J.C., McWhite, C.B. and Grandjean, P.T. 2014. The Cognitive Mechanisms of Intolerance: Do our Meta-Ethical Commitments Matter? In: T. Lombrozo, J. Knobe and S. Nichols, eds. *Oxford Studies in Experimental Philosophy, vol. 1*. Oxford: Oxford University Press, 28-61.

Appendix

ID	Category	Question
1	Factual	At sea level, water boils at 50 degree celcius or 100 degree Fahrenheit.
2	Moral	Cheating on an exam to obtain a job for which you are not qualified is wrong.
3	Taste	Classical music is better than rock music.
4	Taste	CNN provides better news coverage than does FoxNews.
5	Religion	Destroying religious symbols is unacceptable.
6	Moral	Discriminating against someone on the basis of race is unacceptable.
7	Moral	Disregarding the wellbeing of future generations is wrong.
8	Moral	Eating factory farmed animals is wrong.
9	Taste	Expensive wine is of better quality than cheap wine.
10	Moral	Falsely reporting on your taxable income is unacceptable.
11	Moral	Gay marriage is an acceptable arrangement.
12	Religion	God does not exist.
13	Moral	Having consensual sex with another close family member is unacceptable.
14	Moral	In hiring a new employee, it is acceptable to prefer the son of your friend to an unknown candidate, if their qualities are of equally good.
15	Moral	It is acceptable for a woman to have an abortion before the 3rd month of her pregnancy.
16	Moral	It is acceptable not to leave a tip at a restaurant, even when you are satisfied about the service.
17	Moral	It is acceptable to assist in the death of a terminally ill friend, who is in pain and wants to die.
18	Moral	It is acceptable to give less than 1 percent of your yearly income to cancer research.
19	Moral	It is acceptable to have friends only from your own ethnic background.
20	Moral	It is acceptable to tell the authorities that you have seen your friend watching child porn.
21	Moral	It is acceptable to use recreational drugs.
22	Moral	It is unacceptable to buy products that are the result of slave labour.
23	Moral	It is unacceptable to pay higher wages to men than to women for similar performances.
24	Moral	It is wrong for scientists to conduct research on embryonic human stem cells.
25	Moral	Killing innocent people is acceptable, if it serves a greater good.
26	Moral	Killing innocent people is unacceptable, even if it serves a greater good.
27	Moral	Knowingly overcharging someone for a service in order to make more money is unacceptable.
28	Moral	Marriage between two people of the same sex is acceptable.
29	Factual	New York is in the United States of America.
30	Moral	Sleeping next to a dead person is acceptable.
31	Moral	Smoking cigarettes openly in enclosed public places is acceptable.
32	Factual	The earth is flat.
33	Moral	Serving only vegetarian food at a party is acceptable.
34	Moral	To steal small office items, even if no one will notice, is unacceptable.
35	Moral	When you are on a business trip with a company credit card, it is unacceptable to buy the most expensive menu on the table.

8

Chapter 8

Conclusion

'A simple consideration of the course of evolution will (...) by no means suffice to inform us of the course we ought to pursue. We shall have to employ all the resources of a strictly ethical discussion to arrive at a correct valuation of the different results of evolution.'

– G.E. Moore (1903)

8.1 Synthesis of the Research Results

Let's return to the main question with which this dissertation began. What implications do recent empirical findings, particularly from the fields of evolutionary biology, anthropology, psychology and history, have for our understanding of moral objectivity? More specifically, can we procure a sufficiently strong account of moral objectivity in the face of morality's evolutionary origins?

I have argued that although recent findings from evolutionary biology and moral psychology do not undermine the objectivity of our moral beliefs, they do put pressure on realist views according to which moral objectivity is best understood in terms of stance-independence. The research articles that make up the chapters of this dissertation contain two original challenges to realism. The first challenge is related to Street's (2006) Darwinian Dilemma but expands on it in order to fix some shortcomings of Street's original argument. The second challenge builds on recent findings from experimental moral psychology and evolutionary psychology to debunk presumptive arguments in support of moral realism. The main positive claim of this dissertation is that moral objectivity is best understood in terms of stance-invariance. Let's succinctly review these three arguments.

8.1.1 The Genealogical Challenge for Moral Realism

Street's (2006) Darwinian Dilemma is generated by the empirical premise that the contents of our moral endorsements are saturated with evolutionary influence. This creates a dilemma for realists:

- either they hold that there is *no* relation between the evolutionary forces that have influenced the contents of our moral endorsements and the stance-independent moral facts or truths;
- or they hold that there *is* such a relation.

For realists who accept Street's empirical premise, both horns are problematic. Plausibly, realists who take up the first horn of the dilemma *after* having granted Street's empirical premise are committed to the view that assuming realism, the contents of moral truths could have been anything. Hence, they can only save their claim to epistemological optimism by positing an unlikely coincidence. However, unlikely coincidences cannot simply be posited

absent theoretical justification; hence, if no justification for positing a coincidence is forthcoming, optimistic realism fails. *Prima facie*, it seems more promising for realists to take up the dilemma's second horn. However, having granted the empirical premise of Street's dilemma, the realist's truth-tracking account is scientifically inferior to Street's adaptive-link account. Assuming that compatibility with our best science is regarded as a weighty metaethical *desideratum*, the dilemma discredits moral realism.

However, there is still an escape route for realists: they can deny the empirical premise that gives rise to the dilemma (e.g. Shafer-Landau 2012; FitzPatrick 2015; Huemer 2016). Indeed, there are good reasons to pursue this option: while evolutionary forces have certainly shaped the *basic contents* of our moral judgments, it is far less obvious that they have shaped our *full-fledged judgments* (cf. Buchanan and Powell 2015). As a result, the success of Street's Darwinian Dilemma seems to be limited: realists can legitimately withstand her evolutionary premise and pursue a divide-and-conquer strategy, submitting that while *some* full-fledged moral judgments are best explained in evolutionary terms, others are best explained on the assumption that we have tracked stance-independent moral truths.

Yet realists who pursue this strategy may be challenged on different genealogical grounds. In order to explain the correlation between our moral beliefs and the moral truths, realists are still committed to a tracking thesis, albeit in a slightly different context: over the course of *moral history* we have managed to track stance-independent moral truths. This is, once again, a scientific hypothesis, which has to compete with rival historical accounts that do not posit the existence of stance-independent moral truths. Taking into account all of the evidence, I have argued that the rival explanations that antirealists invoke are superior, in the following, related, respects: they are more parsimonious, more comprehensive, more clear, have a better fit with other disciplines, are more predictive and do not rely on a spurious moral epistemology. Indeed, they are even superior in shedding light on an *explanandum* – the origins of *inclusionary moral values* – that might have initially seemed to be particularly amenable to the realist's best case.

In sum, in order for the debunker's argument against realism to work, its scope should be expanded. Apart from evolutionary considerations, various *historical* considerations should be incorporated in order to provide a more comprehensive explanation of the contents of our full-fledged moral judgments. If evolutionary and historical findings allow us to come up with a sufficiently comprehensive genealogy of our moral endorsements, and if this genealogy is scientifically superior to the realist's explanation of our moral endorsements, then our best science gives us reason to prefer antirealism over realism. Since there are good reasons for thinking that we can indeed provide such a genealogy, it follows that, assuming that we should regard it as a *desideratum* for metaethical theories to be compatible with our best science, realism faces great difficulties.

Does this mean that science debunks moral realism? That depends on how the term debunking is understood and how broadly realism is defined. We may stipulate, for instance, that to debunk a position in metaethics is to show that the position fails to satisfy one or more critical metaethical *desiderata*, such that the overall balance of plausibility weighs heavily against it. Thus understood, scientific considerations *do* debunk realism. But it should be kept in mind that this debunking revolves around one distinct clause of realist positions: the stance-independence of moral truths and the resultant tracking account on which a robustly realist

success epistemology must rely. More moderate realists, who drop this clause, may be able to dodge the bullet fired at them by scientifically minded metaethicists.

8.1.2 The Psychological Challenge for Moral Realism

I have also advanced a second challenge for moral realism, building on recent findings about the psychosocial function of perceiving moral judgments as objective and about the evolutionary origins of such perceptions. For example, it has recently been hypothesized that both the impartiality typical of moral demands (DeScioli and Kurzban 2013) and their perceived authority-independence (Stanford 2018) can be explained in evolutionary terms. Does this new scientific picture have philosophical implications? More specifically, does it affect the metaethical doctrine that moral judgments (at least sometimes) make a claim to objectivity and that they (at least sometimes) succeed in being objective?

The answer depends on which view of moral objectivity one adheres to. For objectivists in the realist's sense, who posit the existence of a stance-independent realm of moral truths, the fact that our objectivist moral commitments can be evolutionarily explained is problematic in two ways. The first problem has to do with the foregoing genealogical challenge. These recent hypotheses from evolutionary psychology give further substance to evolutionary explanations of the contents of our moral endorsements and thereby increase the strength of antirealist explanations of these endorsements – that is, explanations that nowhere refer to stance-independent moral truths. If these explanations become more plausible, the realist's rival truth-tracking explanation becomes relatively less plausible.

But these recent findings also challenge realism in a different way. One of the main arguments that realists have traditionally provided in support of their view is an appeal to intuition, phenomenology and the seemingly objectivist features of moral discourse. These features, it has appeared to many metaethicists, are best explained on the assumption that moral realism is true. Indeed, in virtue of this appearance, metaethical debates often starts from the presumption that moral realism is the default metaethical view. However, the evolutionary account outlined above suggests that many objectivist features of moral discourse can be explained in purely scientific terms. In and by itself, this does not debunk moral realism – but it does undercut any presumed metaethical plausibility points that realist theories might have been thought to have in virtue of the fact that they vindicate morality's objective seemings. Indeed, these objective seemings may be little more than the culmination of various cognitive mechanisms, biases and illusions created for us by evolution, evolutionary psychologists suggests.

8.1.3 Explicating Moral Objectivity in terms of Stance-Invariance

Even if morality's objectivity in the realist's sense is illusory, this should not be taken to imply that morality's objectivity *as such* is illusory. Moral objectivity can also be understood in a different, antirealist sense, which is fully compatible with our evolutionary account understanding of morality's origins. This is the account – objectivity as *stance-invariance* – that I have defended in this dissertation. On this account, moral truths are objective if they can withstand argumentative scrutiny from a wide range of different evaluative standpoints, whereby an evaluative standpoint consists of the set of moral judgments that an informed and coherent moral agent would endorse upon reflection. In effect, a moral judgment is objective

if it can withstand scrutiny from the perspective of several different moral agents, who diverge with regard to their starting set of moral judgments. The greater the diversity of standpoints from which a moral judgment can withstand such scrutiny, the more objective we may take it to be. Paradigmatically objective moral judgments are able to withstand reflective scrutiny from the standpoints of all human beings: they apply to anyone, irrespective of his or her *specific* preferences. Hence, on this account, it is at least *possible* that universal moral demands can be vindicated – but whether this possibility is actualized depends on whether there are indeed moral demands that can withstand reflective scrutiny from the idealized standpoints of *all* human agents.

Is this view of objectivity sufficiently strong? I have argued that some considerations that have traditionally been regarded as weaknesses of moral antirealism – its apparent subjectivism or relativism – need not carry as much metaethical weight as has traditionally been placed on them. What account of moral objectivity should be vindicated depends, at least in part, on people's ordinary moral commitments. Experimental research – though still troubled by some methodological complications – is beginning to shed light on this question, and thus far it has given us little reason to think that people typically believe that moral demands are objective in the realist's strong sense. Hence, the antirealist's account of objectivity may not be as strong as that of the realist, but there are empirical grounds for thinking that it does not have to be.

Moreover, the antirealist's account of objectivity can in fact be quite strong, in a way that seems apt for the moral domain. Should moral facts or truths hold irrespectively of the beliefs or inclinations of any *specific* individual, in order to count as objective? Plausibly, this is indeed a *desideratum* for a credible account of moral objectivity, and antirealists can satisfy it, by framing moral objectivity in terms of stance-invariance. What sets apart an antirealist account of objectivity from a realist account is that moral truths can never obtain independently of *anyone's* desires or inclinations. Would an apparent moral transgression – harming an innocent person – still be morally wrong, even if no one thought it was, not even upon idealized reflection? The moral antirealist submits that *that* kind of moral objectivity is a pipe dream. To the extent that metaethicists are committed to it, evolutionary, psychological and historical findings show that in all likelihood they are mistaken.

8.2 Summary of Chapters

Chapter 1

In this introductory chapter I presented the aim of this dissertation: to assess the objectivity of morality's foundations in the light of morality's evolutionary origins, as well as other fields of empirical knowledge. What implications, if any, do empirical findings have for the standing of metaethical theories? I outlined my thesis that these findings detract from the plausibility of a realist account of moral objectivity but leave room for a more moderate, antirealist account of objectivity that is framed in terms of stance-invariance. Additionally, I delineated the topic of the research project, outlined the aims and methods of the dissertation, and demonstrated the coherence of the treatise.

Chapter 2

Can moral realism be debunked on purely philosophical grounds – that is, without appeal to empirical or scientific considerations? In this chapter I engaged with Street’s (2016a) practical/theoretical puzzle, which aspires to do just this. The puzzle is to explain the assumed coincidence between the set of normative judgments that causal forces have led us to believe and the set of normative judgments we take to be true – a puzzle, Street argues, that is troubling for normative realists. I discussed the promise and limitations of Street’s puzzle in the context of a general method for reasoning about coincidences, and concluded that while Street’s puzzle poses a challenge for normative non-naturalist realists such as Enoch (2010), it does not pose a *general* challenge for normative realism: realists who drop the clause that mind-independent moral facts are causally impotent can solve the puzzle. To pose a more general challenge to realist theories, Street’s epistemological challenge should be connected with the scientific challenge engrained in her Darwinian Dilemma.

Chapter 3

In this chapter I identified which empirical assumptions are relevant for EDAs against moral realism. I pointed out that Sharon Street’s and Michael Ruse’s respective EDAs both involve substantive claims about the evolution of our moral judgments. Street claims that the contents of our moral judgments have been indirectly shaped by natural selection; Ruse claims that our moral judgments are evolutionarily contingent. Combining Ruse’s and Street’s claims, I argued, yields an EDA that is more forceful than each of their arguments taken individually. Realists, however, may object to the evolutionary premise that underlies this argument, namely that evolutionary forces have heavily influenced the contents of our endorsed moral judgments. Debunkers, in turn, may counter this objection by drawing upon other sources of empirical knowledge to argue that the best scientific explanation of our moral endorsements involves no appeal to mind-independent truths.

Chapter 4

In this chapter I argued that the thesis that we have tracked mind-independent moral truths is indeed very implausible on historical grounds. I specifically focused on the claim made by Michael Huemer that over the course of human history there has been a global shift in moral values towards a broadly liberal orientation. According to Huemer, this shift better accords with a realist than an antirealist metaethics: it is best explained by the discovery of mind-independent truths through intuition. I argued, *contra* Huemer, that the relevant historical data are better explained by assuming the truth of moral antirealism. Realism does not fit the data as well as Huemer suggests, whereas antirealists have underappreciated resources to explain the relevant historical dynamics. These resources include an appeal to socialization, to technological and economical convergences, to lessons learned from history, to changes induced by consistency reasoning and to the social function of moral norms in overcoming some of the cooperation problems that globalizing societies face. I pointed out that the realist’s *explanans* has multiple shortcomings, that the antirealist’s *explanans* has several explanatory virtues, and concluded that the latter provides a superior account of the historical shift towards liberal values.

Chapter 5

Even if the position of moral antirealists is superior in terms of external accommodation, might realist positions not be superior in terms of internal accommodation? Some metaethicists, including Street (2016b), maintain that the ‘holy grail’ of metaethics is to vindicate the realist’s strong form of objectivism. In this chapter I questioned whether vindicating moral objectivity in the realist’s sense is indeed a metaethical *desideratum*, by analysing exactly which phenomena a metaethical account of objectivity should be able to capture. On the basis of recent experimental findings, I made a preliminary case for the thesis that antirealists have resources to vindicate an account of objectivity at just the right level of strength, given the commitments of ordinary moral thought and practice. While some antirealists have advanced an account of objectivity building on a Kantian version of constructivism, in this chapter I developed such an account departing from a Humean version of constructivism. I presented the thesis that to claim that a moral judgment is objective is to claim that it has the property of being stance-invariant: the judgment withstands scrutiny from the idealized standpoint of an evaluative agent, even under a (finite) range of counterfactual transformations of this evaluative standpoint. On this account, a moral demand can be objective in the sense that its binding force neither depends on the authority of the person making the demand, nor on the *specific* inclinations of the agent to whom the demand applies. However, the account cannot satisfy the realist’s more stringent condition of objectivity: that moral properties can be instantiated even if no one think they are. I concluded that while moral realism has traditionally been presented as a position with intuitive appeal, it is not obvious that ordinary moral thought and practice are indeed committed to this more stringent condition.

Chapter 6

This chapter engaged in more detail with recent experimental findings about the moral commitments of folk respondents and analysed the metaethical implications of these findings. Experiments suggest that people are metaethical pluralists: there is both inter- and intrapersonal variation with regard to people’s objectivist commitments. What metaethical implications do these findings have? I highlighted that, on the one hand, current research does not directly address traditional metaethical questions: the methods used and distinctions drawn by experimenters do not perfectly match those of metaethicists. But I went on to argue that in spite of this mismatch, the research findings are nonetheless metaethically relevant. For one, they extend the field of moral psychology with new data and hypotheses. But they also provide tentative evidence that touches on the adequacy of theses in moral semantics and moral metaphysics. With regard to the tenability of moral realism, they help to put pressure on so-called presumptive arguments in support of moral realism.

Chapter 7

Presumptive arguments in support of moral realism can be criticized on the basis of experimental psychology, but also on the basis of evolutionary psychology. As pointed out in the previous chapter, even if we grant the phenomenological *datum* that moral demands are typically experienced as being objective, this need not give support to moral realism. After all, there might be an alternative explanation for this phenomenological *datum*: for instance, our experience of morality’s apparent objectivity can perhaps be explained in evolutionary terms. Ruse (1995) proposed a general evolutionary explanation along these lines, and recently

Stanford (2018) has advanced a more detailed hypothesis, driven by theoretical as well as empirical considerations. In this chapter, based on collaborative work with Michael Klenk, we advanced a slightly different hypothesis about the evolutionary function of perceiving moral demands as objective, expanding on DeScioli and Kurzban's (2013) proposal. We presented experimental evidence indicating that one of the factors which correlates with increased attributions of objectivity is subjects' personal interest: the more important subjects find a moral statement, the more objective they take it to be. While aware of some of the limitations of our study, we argued that the overall picture we presented lends support to the view that perceiving moral demands as objective serves a psychosocial function. Moreover, we argued that morality's objective appearance be explained, at least in part, in evolutionary terms, which serves to undercut the realist's presumption that our moral phenomenology provides tentative support for moral realism.

8.3 Realism versus Antirealism: The Overall Balance of Plausibility

It has previously been argued that metaethicists face a trade-off between vindicating a scientifically respectable theory and a theory that adequately captures the appearances of moral discourse – a trade-off between external and internal accommodation (Finlay 2007). If the arguments presented in this dissertation have been along the right lines, this diagnosis is incorrect. Realism is not obviously superior in terms of internal accommodation. Both realism and antirealism are able to capture core aspects of people's ordinary moral commitments, including the claim that many moral judgments are at least somewhat objective. Arguably, antirealists are even better placed than realists to explain the observed diversity in folk moral commitments.

But even if we grant that considerations regarding internal accommodation are ultimately inconclusive in settling the debate between moral realists and antirealists, considerations regarding external accommodation are not: moral antirealism has a superior fit with our best science. Therefore, if we take science seriously, we should conclude that claims to moral objectivity are best understood in antirealist terms.

References

- Buchanan, A. and Powell, R. 2015. The Limits of Evolutionary Explanations of Morality and Their Implications for Moral Progress. *Ethics* 126: 37-67.
- DeScioli, P. and Kurzban, R. 2013. A Solution to the Mysteries of Morality. *Psychological Bulletin* 139: 477-496.
- Enoch, D. 2010. The Epistemological Challenge to Metanormative Realism: How Best to Understand It, and How to Cope with It. *Philosophical Studies* 148: 413-438.
- Finlay, S. 2007. Four Faces of Moral Realism. *Philosophy Compass* 6: 820-849.
- FitzPatrick, W. 2015. Debunking Evolutionary Debunking of Ethical Realism. *Philosophical Studies* 172: 883-904.
- Huemer, M. 2016. A Liberal Realist Answer to Debunking Skeptics: The Empirical Case for Realism. *Philosophical Studies* 173: 1983-2010.

- Moore, G.E. 1993 [1903]. *Principia Ethica*. Cambridge: Cambridge University Press.
- Ruse, M. 1995. Evolutionary Ethics: A Phoenix Arisen. In: P. Thompson, ed. *Issues in Evolutionary Ethics*. Buffalo: SUNY Press, 225-247.
- Shafer-Landau, R. 2012. Evolutionary Debunking, Moral Realism and Moral Knowledge. *Journal of Ethics and Social Philosophy*, www.jesp.org, 7(1): 1-37.
- Street, S. 2006. A Darwinian Dilemma for Realist Theories of Value. *Philosophical Studies* 127: 109-166.
- Street, S. 2016a. Objectivity and Truth: You'd Better Rethink It. In: R. Shafer-Landau, ed. *Oxford Studies in Metaethics*, vol. 11. Oxford: Clarendon Press, 293-334.
- Street, S. 2016b. Constructivism in Ethics and the Problem of Attachment and Loss. *Aristotelian Society Supplementary Volume xc*: 161-189.
- Stanford, P.K. 2018. The Difference Between Ice Cream and Nazis: Moral Externalization and the Evolution of Human Cooperation. *Behavioral and Brain Sciences* 41: 1-57.

Bibliography

- Alexander, R.D. 1987. *The Biology of Moral Systems*. New York: Aldine de Gruyter.
- Anderson, E. 2016. The Social Epistemology of Morality: Learning from the Forgotten History of the Abolition of Slavery. In: M.S. Brady and M. Fricke, eds. *Contribution to The Epistemic Life of Groups*. Oxford: Oxford University Press, 75-94.
- Artiga, M. 2015. Rescuing Tracking Theories of Morality. *Philosophical Studies* 172: 3357-3374.
- Bagnoli, C. 2015. Moral Objectivity: A Kantian Illusion? *Journal of Value Inquiry* 49: 31-45.
- Baumard, N., André, J.B., Sperber, D. 2013. A Mutualistic Approach to Morality: The Evolution of Fairness by Partner Choice. *Behavioral and Brain Sciences* 36: 59-122.
- Beebe, J.R. 2014. How Different Kinds of Disagreement Impact Folk Metaethical Judgments. In: J.C. Wright and H. Sarkissian, eds. *Advances in Experimental Moral Psychology: Affect, Character, and Commitments*. London: Continuum, 167-187.
- Beebe, J.R. 2015. The Empirical Study of Folk Metaethics. *Etyka* 15: 11-28.
- Beebe, J.R. and Sackris, D. 2016. Moral Objectivism across the Lifespan. *Philosophical Psychology* 29: 912-929.
- Bengson, J. 2013. Experimental Attacks on Intuitions and Answers. *Philosophy and Phenomenological Research* 86: 495-532.
- Berker, S. 2014. Does Evolutionary Psychology Show that Normativity is Mind-Dependent? In: J. D'Arms and D. Jacobson, eds. *Moral Psychology and Human Agency: Essays on the Science of Ethics*. New York: Oxford University Press, 215-252.
- Berker, S. 2018. The Unity of Grounding. *Mind* 127: 729-777.
- Berker, S. Forthcoming. The Explanatory Character of Moral Principles. *Noûs*.
- Bicchieri, C. 2006. *The Grammar of Society*. Cambridge: Cambridge University Press.
- Blackburn, S. 1984. *Spreading the Word: Groundings in the Philosophy of Language*. New York: Oxford University Press.
- Blackburn, S. 1993. *Essays in Quasi-Realism*. Oxford: Oxford University Press.
- Blackburn, S. Unpublished Manuscript. Sharon Street on the Independent Normative Truth as such. www2.phil.cam.ac.uk/~swb24/PAPERS/Meanstreet.htm
- Boehm, C. 2012. *Moral Origins: The Evolution of Virtue, Altruism, and Shame*. New York: Basic Books.
- Bogardus, T. 2016. Only all Naturalists Should Worry about Only One Evolutionary Debunking Argument. *Ethics* 126: 636-661.

- Bonnefon, J., Shariff, A. and Rahwan, I. 2016. The Social Dilemma of Autonomous Vehicles. *Science* 352: 1573-1576.
- Bourget, D. and Chalmers, D.J. 2014. What Do Philosophers Believe? *Philosophical Studies* 170: 465-500.
- Braddock, M. 2016. Evolutionary Debunking: Can Moral Realists Explain the Reliability of Our Moral Judgments? *Philosophical Psychology* 29: 844-857.
- Bratman, M. 2012. Constructivism, Agency, and The Problem of Alignment. In: J. Lenman and Y. Shemmer, eds. *Constructivism in Practical Philosophy*. Oxford: Oxford University, 81-98.
- Brink, D. 1989. *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- Brosnan, K. 2011. Do the Evolutionary Origins of our Moral Beliefs undermine Moral Knowledge? *Biology & Philosophy* 26: 51-64.
- Brosnan, S.F and de Waal, F. 2003. Monkeys Reject Unequal Pay. *Nature* 425: 297-299.
- Buchanan, A. and Powell, R. 2015. The Limits of Evolutionary Explanations of Morality and Their Implications for Moral Progress. *Ethics* 126: 37-67.
- Burge, T. 2010. *Origins of Objectivity*. Oxford: Clarendon Press.
- Campbell, R. and Kumar, V. 2012. Moral Reasoning on the Ground. *Ethics* 122: 273-312.
- Carnap, R. 1947. On the Application of Inductive Logic. *Philosophy and Phenomenological Research* 8: 133-148.
- Carnap, R. 1956. *Meaning and Necessity*. Chicago: University of Chicago Press.
- Chudek, M. and Henrich, J. 2011. Culture–Gene Coevolution, Normpsychology and the Emergence of Human Prosociality. *Trends in Cognitive Sciences* 15: 218-226.
- Clarke-Doane, J. 2012. Morality and Mathematics: The Evolutionary Challenge. *Ethics* 122: 313-340.
- Clarke-Doane, J. 2014. Moral Epistemology: The Mathematics Analogy. *Noûs* 48: 238-255.
- Clarke-Doane, J. 2016. Debunking and Dispensability. In: U.D. Leibowitz and N. Sinclair, eds. *Explanation in Ethics and Mathematics: Debunking and Dispensability*. Oxford: Oxford University Press, 23-36.
- Copp, D. 2008. Darwinian Skepticism about Moral Realism. *Philosophical Issues* 18: 186-206.
- Crow, D. 2016. Causal Impotence and Evolutionary Influence: Epistemological Challenges for Non-Naturalism. *Ethical Theory and Moral Practice* 19: 379-395.
- Cuneo, T. and Shafer-Landau, R. 2014. The Moral Fixed Points. New Directions for Moral Nonnaturalism. *Philosophical Studies* 171: 399-443.

Bibliography

- Curry, O. 2016. *Morality as Cooperation: A Problem-Centred Approach*. In T. K. Shackelford and R. D. Hansen eds. *The Evolution of Morality*. Heidelberg: Springer International Publishing, 27-51.
- Darwin, C. 2013 [1871]. *The Descent of Man and Selection in Relation to Sex*. Ware: Wordsworth Classics.
- Das, R. 2015. Evolutionary Debunking of Morality: Epistemological or Metaphysical? *Philosophical Studies* 173: 417-435.
- Diaconis, P. and Mosteller, F. 1989. Methods for Studying Coincidences. *Journal of the American Statistical Association* 84: 853-861.
- Deem, M.J. 2016. Dehorning the Darwinian Dilemma for Normative Realism. *Biology & Philosophy* 31: 727-746.
- DeScioli, P. 2016. The Side-Taking Hypothesis for Moral Judgment. *Current Opinion in Psychology* 7: 23-27.
- DeScioli, P. and Kurzban, R. 2009. Mysteries of Morality. *Cognition* 112: 281-299.
- DeScioli, P. and Kurzban, R. 2013. A Solution to the Mysteries of Morality. *Psychological Bulletin* 139: 477-496.
- Dewey, J. and Tufts, J.H. 1908. *Ethics*. New York: Henry Holt.
- Doris, J. and Plakias, A. 2008. How to Argue about Disagreement: Evaluative Diversity and Moral Realism. In: W. Sinnott-Armstrong, ed. *Moral psychology, vol. 2*. Cambridge: MIT Press, 303-331.
- Doyle, A.C. 2011 [1913]. *The Adventure of the Dying Detective*. Portland (OR): The Floating Press.
- Dworkin, R. 1996. Objectivity and Truth: You'd Better Believe It. *Philosophy & Public Affairs* 25: 87-139.
- Dworkin, R. 2011. *Justice for Hedgehogs*. Cambridge: Harvard University Press.
- Enoch, D. 2007. An Outline of an Argument for Robust Metanormative Realism. In: R. Shafer-Landau, ed. *Oxford Studies in Metaethics 2*. New York: Oxford University Press, 21-50.
- Enoch, D. 2009. How is Moral Disagreement a Problem for Realism? *Journal of Ethics* 13: 15-50.
- Enoch, D. 2010. The Epistemological Challenge to Metanormative Realism: How Best to Understand it, and How to Cope with It. *Philosophical Studies* 148: 413-438.
- Enoch, D. 2014. Why I Am an Objectivist about Ethics (And Why You Are, Too). In: R. Shafer-Landau, ed. *The Ethical Life*. Oxford: Oxford University Press, 192-205.

- Enoch, D. 2016. Dispensability Arguments in Metaethics: Even better than Mathematics? In: U.D. Leibowitz and N. Sinclair, eds. *Explanation in Ethics and Mathematics: Debunking and Dispensability*. Oxford: Oxford University Press, 236-254.
- Enoch, D. 2018. Non-Naturalistic Realism in Meta-Ethics. In: T. McPherson and D. Plunkett, eds. *The Routledge Handbook of Metaethics*. New York: Routledge, 29-42.
- Falk, R. 1989. Judgment of Coincidences: Mine versus Yours. *The American Journal of Psychology* 102: 477-493.
- Feltz, A. and Cokely, E.T. 2008. The Fragmented Folk: More Evidence of Stable Individual Differences in Moral Judgments and Folk Intuitions. In: B.C. Love, K. McRae and V.M. Sloutsky, eds. *Proceedings of the 30th Annual Conference of the Cognitive Science Society*. Austin: Cognitive Science Society, 1771-1776.
- Fine, K. 2012. Guide to Ground. In: F. Correia and B. Schnieder eds. *Metaphysical Grounding: Understanding the Structure of Reality*. Cambridge: Cambridge University Press, 37-80.
- Finlay, S. 2007. Four Faces of Moral Realism. *Philosophy Compass* 6: 820-849.
- Fisher, M., Knobe, J., Strickland, B. and Keil, F.C. 2017. The Influence of Social Interaction on Intuitions of Objectivity and Subjectivity. *Cognitive Science* 41: 1119-1134.
- Fitzpatrick, S. 2014. Moral Realism, Moral Disagreement, and Moral Psychology. *Philosophical Papers* 43: 161-190.
- FitzPatrick, W. 2008. Robust Ethical Realism, Non-Naturalism, and Normativity. In: R. Shafer-Landau, ed. *Oxford Studies in Metaethics, vol. 3*. Oxford: Clarendon Press, 159-205.
- FitzPatrick, W. 2012. Review of Philip Kitcher. The Ethical Project. *Ethics* 123: 167-174.
- FitzPatrick, W. 2014. Why There Is no Darwinian Dilemma for Ethical Naturalism. In: M. Bergmann and P. Kain, eds. *Challenges to Moral and Religious Belief: Disagreement and Evolution*. Oxford: Oxford University Press, 237-255.
- FitzPatrick, W. 2015. Debunking Evolutionary Debunking of Ethical Realism. *Philosophical Studies* 172: 883-904.
- Flanagan, O., Ansell, A., Martin, S. and Steenbergen, G. 2014. Empiricism and Normative Ethics: What Do the Biology and the Psychology of Morality Have To Do with Ethics? *Behaviour* 151: 209-228.
- Fraser, B. 2014. Moral Error Theories and Folk Metaethics. *Philosophical Psychology* 27: 789-806.
- Fraser, B. and Sterelny, K. 2017. Evolution and Moral Realism. *The British Journal for the Philosophy of Science* 68: 981-1006.
- Gibbard, A. 1990. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Oxford: Clarendon Press.
- Gibbard, A. 1999. Morality as Consistency in Living: Korsgaard's Kantian Lectures. *Ethics*

Bibliography

110: 140-164.

Gill, M.B. 2009. Indeterminacy and Variability in Meta-Ethics. *Philosophical Studies* 145: 215-234.

Goodwin, G.P. and Darley, J.M. 2008. The Psychology of Meta-Ethics: Exploring Objectivism. *Cognition* 106: 1339-1366.

Goodwin, G.P. and Darley, J.M. 2010. The Perceived Objectivity of Ethical Beliefs: Psychological Findings and Implications for Public Policy. *Review of Philosophy and Psychology*. 1: 161-188.

Goodwin, G.P. and Darley, J.M. 2012. Why Are some Moral Beliefs Perceived to be More Objective than Others? *Journal of Experimental Social Psychology* 48: 250-256.

Gould, S.J. and Lewontin, R.C. 1979. The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proceedings of the Royal Society of London B* 205: 581-598.

Greene, J. 2014. Beyond Point-and-Shoot Morality: Why Cognitive (Neuro)Science Matters for Ethics. *Ethics* 124: 695-726.

Griffiths, T.L and Tenenbaum, J.B. 2007. From Mere Coincidences to Meaningful Discoveries. *Cognition* 103: 180-226.

Habermas, J. 1991. *Erläuterungen zur Diskursethik*. Frankfurt am Main: Suhrkamp Taschenbuch Wissenschaft.

Habermas, J. 1995. Reconciliation through the Public Use of Reason: Remarks on John Rawls's Political Liberalism. *The Journal of Philosophy* 92: 109-131.

Haidt, J. and Bjorklund, F. 2008. Social Intuitionists Answer Six Questions about Moral Psychology. In: W. Sinnott-Armstrong, ed. *Moral psychology, vol. 2*. Cambridge: MIT Press.

Haidt, J., Koller, S.H. and Dias, M.G. 1993. Affect, Culture and Morality, or Is It Wrong to Eat Your Dog? *Attitudes and Social Cognition* 65: 613-628.

Hamilton, W.D. 1963. The Evolution of Altruistic Behavior. *American Naturalist* 97: 354-356.

Hamlin, K. 2014. The Origins of Human Morality: Complex Socio-Moral Evaluations by Preverbal Infants. *New Frontiers Social Neuroscience* 21: 165-188.

Handfield, T. 2016. Genealogical Explanations of Chance and Morals. In: U.D. Leibowitz and N. Sinclair, eds. *Explanation in Ethics and Mathematics: Debunking and Dispensability*. Oxford: Oxford University Press, 58-82.

Hannon, M. Forthcoming. Intuitions, Reflective Judgments, and Experimental Philosophy. *Synthese*.

Hanson, L. 2016. The Real Problem with Evolutionary Debunking Arguments. *The Philosophical Quarterly* 67: 508-533.

- Heiphetz, L. and Young, L.L. 2017. Can only one Person be Right? The Development of Objectivism and Social Preferences regarding Widely Shared and Controversial Moral Beliefs. *Cognition* 167: 78-90.
- Henrich, J. 2015. *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton (NJ): Princeton University Press.
- Hopster, J.K.G. 2017. Two Accounts of Moral Objectivity: From Attitude-Independence to Standpoint-Invariance. *Ethical Theory & Moral Practice* 20: 763-780.
- Horgan, T. and Timmons, M. 2006. Expressivism, yes! Relativism, no! In: R. Shafer-Landau, ed. *Oxford Studies in Metaethics, vol. 1*. Oxford: Clarendon Press, 73-98.
- Hrdy S.B. 2009. *Mothers and Others: The Evolutionary Origins of Mutual Understanding*. Cambridge: Harvard University Press.
- Huemer, M. 2005. *Ethical Intuitionism*. New York: Palgrave Macmillan.
- Huemer, M. 2008. Revisionary Intuitionism. *Social Philosophy and Policy* 25: 368-392.
- Huemer, M. 2016. A Liberal Realist Answer to Debunking Skeptics: The Empirical Case for Realism. *Philosophical Studies* 173: 1983-2010.
- Hume, D. 1738. *A Treatise of Human Nature*, Book III, part I, section I.
- Jackson, F. and Pettit, P. 1995. Moral Functionalism and Moral Motivation. *The Philosophical Quarterly* 45: 20-40.
- Jaeger, G. 2015. The Wisdom of Hindsight and the Limits of Humean Constructivism. *Philosophical Explorations* 18: 45-57.
- Joyce, R. 2006. *The Evolution of Morality*. Cambridge: MIT Press.
- Joyce, R. 2006. Metaethics and the Empirical Sciences. *Philosophical Explorations* 9: 133-148.
- Joyce, R. 2015. Moral Anti-realism versus Realism: Intuitions. *Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/moral-anti-realism/moral-realism-intuitive.html>
- Joyce, R. 2016. *Essays in Moral Skepticism*. Oxford: Oxford University Press.
- Joyce, R. 2016. Reply: Confessions of a Modest Debunker. In: U.D. Leibowitz and N. Sinclair eds. *Explanation in Ethics and Mathematics*. Oxford: Oxford University Press, 124-148.
- Judd, C. M., Westfall J. and Kenny, D.A. 2012. Treating Stimuli as a Random Factor in Social Psychology: A New and Comprehensive Solution to a Pervasive but Largely Ignored Problem. *Journal of Personality and Social Psychology* 103: 54-69.
- Kahane, G. 2011. Evolutionary Debunking Arguments. *Noûs* 45: 103-125.
- Kauppinen, A. 2007. The Rise and Fall of Experimental Philosophy. *Philosophical*

Bibliography

Explorations 10: 95-118.

Khoo, J. and Knobe, J. 2018. Moral Disagreement and Moral Semantics. *Noûs* 52: 109-143.

Kirchin, S. 2003. Ethical Phenomenology and Metaethics. *Ethical Theory and Moral Practice* 6: 241-264.

Kitcher, P. 1998. Psychological Altruism, Evolutionary Origins, and Moral Rules. *Philosophical Studies* 89: 283-316.

Kitcher, P. 2011. *The Ethical Project*. Cambridge: Harvard University Press.

Klenk, M. 2017. Old Wine in New Bottles: Evolutionary Debunking Arguments and the Benacerraf-Field Challenge. *Ethical Theory & Moral Practice* 20: 781-795.

Knobe, J. 2016. Experimental Philosophy Is Cognitive Science. In: J. Sytsma and W. Buckwalter, *A Companion to Experimental Philosophy*.

Korsgaard, C. 1996. *The Sources of Normativity*. Cambridge: Cambridge University Press.

Kumar, V. Forthcoming. The Ethical Significance of Cognitive Science. In: S.-J. Leslie and S. Cullen, eds. *Current Controversies in Philosophy of Cognitive Science*. New York: Routledge.

Laskowski N. and Finlay, S. 2017. Conceptual Analysis in Metaethics. In: T. McPherson and D. Plunkett, eds. *Routledge Handbook of Metaethics*. New York: Routledge, 536-551.

Leibowitz, U.D. and Sinclair, N. eds. 2016. *Explanation in Ethics and Mathematics: Debunking and Dispensability*. Oxford: Oxford University Press.

Lillehammer, H. 2003. Debunking Morality: Evolutionary Naturalism and Moral Error Theory. *Biology & Philosophy* 18: 567-581.

Loeb, D. 2007. The Argument from Moral Experience. *Ethical Theory and Moral Practice* 10: 469-484.

Loeb, D. 2008. Moral Incoherentism: How to Pull a Metaphysical Rabbit Out of a Semantic Hat. In: W. Sinnott-Armstrong, ed. *Moral Psychology: The Cognitive Science of Morality, vol. 2*. Cambridge: MIT Press, 355-385.

Lott, M. 2018. Must Realists be Skeptics? An Aristotelian Reply to a Darwinian Dilemma. *Philosophical Studies* 175: 71-96.

Ludwig, K. 2010. Intuitions and Relativity. *Philosophical Psychology* 23: 427-445.

Lutz, M. 2017. Forthcoming. What Makes Evolution a Defeater? *Erkenntnis*.

Mackie, J.L. 1977. *Ethics: Inventing Right and Wrong*. New York: Penguin.

Mercier, H. and Sperber, D. 2017. *The Enigma of Reason*. Cambridge: Harvard University Press.

Mitchell-Yellin, B. 2015. Aligning with the Good. *Journal of Ethics and Social Philosophy*, 1-8.

- Mogensen, A.L. 2015. Evolutionary Debunking Arguments and the Proximate/Ulimate Distinction. *Analysis* 75: 196-203.
- Mogensen, A. Forthcoming. Ethics, Evolution, and the Coincidence Problem: A Skeptical Appraisal. <http://andreamogensen.com/wp-content/uploads/2014/10/Coincidence-Problem-v5.pdf>
- Moore, G.E. 1993 [1903]. *Principia Ethica*. Cambridge: Cambridge University Press.
- Nagel, T. 1986. *The View from Nowhere*. New York: Oxford University Press.
- Nagel, T. 2012. *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature is almost certainly False*. Oxford: Oxford University Press.
- Nichols, S. 2004. *Sentimental Rules*. New York: Oxford University Press.
- Nichols, S. 2004. After Objectivity: An Empirical Study of Moral Judgment. *Philosophical Psychology* 17: 5-28.
- Nichols, S. 2014. Process Debunking and Ethics. *Ethics* 124: 727-749.
- Nowak, M. 2012. *Supercooperators: Altruism, Evolution, and Why We Need Each Other to Succeed*. New York: Free Press.
- Nozick, R. 2001. *Invariances: The Structure of the Objective World*. Cambridge: Belknap Press of Harvard University Press.
- Owens, D.J. 1992. *Causes and Coincidences*. Cambridge: Cambridge University Press.
- Pinker, S. 2011. *The Better Angels of Our Nature: Why Violence Has Declined*. New York: Viking.
- Pleasant, N. 2010. Moral Argument Is Not Enough: The Persistence of Slavery and the Emergence of Abolition. *Philosophical Topics* 38: 159-180.
- Pözlner, T. 2017. Are Moral Judgements Adaptations? Three Reasons Why It Is So Difficult to Tell. *South African Journal of Philosophy* 36: 425-439.
- Pözlner, T. 2017. Revisiting Folk Moral Realism. *Review of Philosophy and Psychology* 8: 455-476.
- Pözlner, T. 2018. *Moral Reality and the Empirical Sciences*. New York: Routledge.
- Pözlner, T. Forthcoming. Can the Empirical Sciences Contribute to the Realism / Antirealism Debate? *Synthese*.
- Putnam, H. 2004. *Ethics without Ontology*. Cambridge: Harvard University Press.
- Rawls, J. 1971. *A Theory of Justice*. Cambridge: Harvard University Press.
- Richerson, P.J. and Boyd, R. 2005. *Not by Genes Alone: How Culture Transformed Human Evolution*. Chicago: The University of Chicago Press.

Bibliography

- Ruse, M. 1995. Evolutionary Ethics: A Phoenix Arisen. In: P. Thompson, ed. *Issues in Evolutionary Ethics*. Buffalo (NY): SUNY Press, 225-247.
- Ruse, M. 2010. The Biological Sciences Can Act as a Ground for Ethics. In: F. Ayala and R. Arp, eds. *Contemporary Debates in Philosophy of Biology*. Wiley-Blackwell, 297-315.
- Sarkissian, H. 2016. Aspects of Folk Morality: Objectivism and Relativism. In: W. Buckwalter and J. Sytsma, eds. *A Companion to Experimental Philosophy*. London: Blackwell, 212-224.
- Sarkissian, H. 2017. Folk Platitudes as the Explananda of Philosophical Metaethics: Are They Accurate? And Do They Help or Hinder Inquiry? *Journal of Indian Council of Philosophical Research* 34: 565-575.
- Sarkissian, H., Parks, J., Tien, D., Wright, J.C. and Knobe, J. 2011. Folk Moral Relativism. *Mind & Language* 26: 428-505.
- Scanlon, T.M. 1998. *What We Owe to Each Other*. Cambridge: Harvard University Press.
- Schechter, J. 2018. Explanatory Challenges in Metaethics. In: T. McPherson and D. Plunkett, eds. *The Routledge Handbook of Metaethics*. New York: Routledge, 443-459.
- Schloss, J. 2014. Evolutionary Accounts of Morality: Accounting for the Normal but not the Normative. In: H. Putnam, S. Neiman and J. Schloss, eds. *Understanding Moral Sentiments: Darwinian Perspectives*. New Brunswick: Transaction Publishers, 81-121.
- Schwitzgebel, E. and Cusman, F. 2012. Expertise in Moral Reasoning? Order Effects on Moral Judgment in Professional Philosophers and Non-Philosophers. *Mind & Language* 27: 135-153.
- Sen, A.K. 1993. Positional Objectivity. *Philosophy & Public Affairs* 22: 126-145.
- Shafer-Landau, R. 2003. *Moral Realism: A Defence*. Oxford: Oxford University Press.
- Shafer-Landau, R. 2012. Evolutionary Debunking, Moral Realism and Moral Knowledge. *Journal of Ethics and Social Philosophy*, www.jesp.org, 7(1): 1-37.
- Sinclair, N. 2012. Moral Realism, Face-Values and Presumptions. *Analytic Philosophy* 53: 158-179.
- Skarsaune, K.O. 2011. Darwin and Moral Realism: Survival of the Iffiest. *Philosophical Studies* 152: 229-243.
- Sober, E. 2012. Coincidences and How to Reason about Them. In: H.W. de Regt et al., eds. *EPSA Philosophy of Science, Amsterdam 2009*. Springer, 355-374.
- Stanford, P.K. 2018. The Difference Between Ice Cream and Nazis: Moral Externalization and the Evolution of Human Cooperation. *Behavioral and Brain Sciences* 41: 1-57.
- Street, S. 2006. A Darwinian Dilemma for Realist Theories of Value. *Philosophical Studies* 127: 109-166.
- Street, S. 2008. Constructivism about Reasons. *Oxford Studies in Metaethics*, vol. 3: 207-245.

- Street, S. 2009. In Defense of Future Tuesday Indifference. Ideally Coherent Eccentrics and the Contingency of What Matters. *Philosophical Issues* 19: 273-298.
- Street, S. 2010. What is Constructivism in Ethics and Metaethics? *Philosophy Compass* 5: 363-384.
- Street, S. 2011. Mind-Independence without the Mystery: Why Quasi-Realists Can't Have It Both Ways. In: R. Shafer-Landau, ed. *Oxford Studies in Metaethics, vol. 6*. Oxford: Clarendon Press, 1-32.
- Street, S. 2012. Coming to terms with Contingency: Humean Constructivism about Practical Reason. In: J. Lenman and Y. Shemmer, eds. *Constructivism in Practical Philosophy*. Oxford: Oxford University Press, 40-59.
- Street, S. 2016. Objectivity and Truth: You'd Better Rethink It. In: R. Shafer-Landau, ed. *Oxford Studies in Metaethics, vol. 11*. Oxford: Clarendon Press, 293-334.
- Street, S. 2016. Constructivism in Ethics and the Problem of Attachment and Loss. *Aristotelian Society Supplementary Volume xc*: 161-189.
- Street, S. 2017. Nothing 'Really' Matters, but That's Not What Matters. In: P. Singer, ed. *Does Anything Really Matter: Parfit on Objectivity*. Oxford: Oxford University Press, 121-148.
- Sterelny, K. 2016. Cooperation, Culture, Conflict. *British Journal for the Philosophy of Science* 67: 31-58.
- Tersman, F. 2017. Moral Skepticism and the Benacerraf Challenge. In: D. Machuca, ed. *Moral Skepticism: New Essays*. New York: Routledge, 183-198.
- Theriault, J., Waytz, A., Heiphetz, L., and Young, L. 2017. Examining Overlap in Behavioral and Neural Representations of Morals, Facts, and Preferences. *Journal of Experimental Psychology* 146: 1586-1605.
- Timmons, M. 2010. Objectivity in Moral Discourse. In: A. Barber and R. Stainton, eds. *Concise Encyclopedia of Philosophy of Language and Linguistics*. Oxford: Elsevier Ltd., 539-545.
- Tomasello, M. 2016. *A Natural History of Human Morality*. Cambridge: Harvard University Press.
- Trivers, R.L. 1971. The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology* 46: 35-57.
- Tropman, E. 2014. Evolutionary Debunking Arguments: Moral Realism, Constructivism, and Explaining Moral knowledge. *Philosophical Explorations* 17: 126-140.
- Turiel, E. 1983. *The Development of Social Knowledge: Morality and Convention*. Cambridge: MIT Press.
- Vavova, K. 2014. Debunking Evolutionary Debunking. In: R. Shafer-Landau, ed. *Oxford Studies in Metaethics, vol. 9*: 76-101.

Bibliography

- Waal de, F. 2006. *Primates and Philosophers: How Morality Evolved*. Princeton: Princeton University Press.
- Warneken F. and Tomasello, M. 2011. The Roots of Human Altruism. *British Journal of Psychology* 100: 455-471.
- Warneken, F. 2018. How Children Solve the Two Challenges of Cooperation. *Annual Review of Psychology* 69: 205-229.
- Weinberg, J.M., Alexander, J., Gonnerman, C. and Reuter, S. 2012. Restrictionism and Reflection: Challenge Deflected, or Simply Redirected? *The Monist* 95: 200-222.
- Westermarck, E.A. 1906. *The Origin and Development of the Moral Ideas*. London: Macmillan.
- Wielenberg, E. 2010. On the Evolutionary Debunking of Morality. *Ethics* 120: 441-464.
- Wielenberg, E. 2016. Ethics and Evolutionary Theory. *Analysis* 76: 502-515.
- Wilson, D.S. and Sober, E. 1994. Reintroducing Group Selection to the Human Behavioural Sciences. *Behavioural and Brain Sciences* 17: 585-654.
- Wright, J.C., Grandjean, P.T. and McWhite, C.B. 2013. The Meta-Ethical Grounding of Our Moral Beliefs: Evidence for Metaethical Pluralism. *Philosophical Psychology* 26: 336-361.
- Wright, J.C., McWhite, C.B. and Grandjean, P.T. 2014. The Cognitive Mechanisms of Intolerance: Do our Meta-Ethical Commitments Matter? In: T. Lombrozo, J. Knobe and S. Nichols, eds. *Oxford Studies in Experimental Philosophy, vol. 1*. Oxford: Oxford University Press, 28-61.

Samenvatting in het Nederlands

Dit proefschrift bestaat uit zes artikelen, geschreven voor publicatie in filosofische vaktijdschriften, samen met een inleiding en een conclusie. Daarin belicht ik recente empirische inzichten over de oorsprong van de menselijke moraal, en analyseer in hoeverre die gevolgen hebben voor de metaethische grondslagen van morele objectiviteit. Meer specifiek ga ik in op de vraag of we morele objectiviteit beter op basis van een ‘realistische’ of ‘antirealistische’ metaethiek kunnen benaderen. Ik betoog dat een antirealistische benadering metaethisch gezien superieur is, met name omdat zij ons, beter dan een realistische benadering, in staat stelt om ons filosofische begrip van morele objectiviteit te verzoenen met onze beste wetenschappelijke kennis.

Oorsprong

Verscheidene wetenschappen dragen bij aan ons inzicht over het ontstaan van de moraal. Met name vanuit evolutionaire hoek is de laatste jaren veelvuldig onderzoek gedaan naar de oorsprong van onder meer morele emoties, tribale morele instincten, ons rechtvaardigheidsgevoel en normbesef. Nog lang niet alle details zijn opgehelderd; over tal van specifieke vragen die betrekking hebben op onze morele evolutie bestaan rivaliserende wetenschappelijke opvattingen. Niettemin bestaat er tegenwoordig aanzienlijke consensus over enkele centrale inzichten – bijvoorbeeld over het belang van samenwerking in de evolutie van de Hominini, en de functionele rol die (proto-)morele normen vermoedelijk hebben gespeeld in het stimuleren daarvan. Gold het verklaren van moreel gedrag – meer specifiek van moreel altruïsme – in Darwins tijd nog als één van de grote uitdagingen voor de evolutietheorie, tegenwoordig is er een overdaad aan potentiële verklaringen voorhanden. De vraag is niet meer óf het ontstaan van moreel altruïsme evolutionair verklaard kan worden, maar welke van de verschillende theoretische mogelijkheden daartoe het meest geschikt is.

Overigens hebben hedendaagse verklaringen van de moraal niet alleen betrekking op moreel altruïsme. De moraal is een uiterst breed verschijnsel, met tal van facetten die zich kunnen lenen voor evolutionaire verklaringswijzen. Bovendien komen verklaringen niet alleen uit evolutionaire hoek. Hoewel het ontstaan van basale morele motivaties goed in evolutionaire termen kunnen worden uitgelegd, is het een open vraag hoeveel inzicht evolutionaire verklaringen geven als het gaat om onze meer verfijnde morele oordelen. Een belangrijke filosofische distinctie hierbij is of verklaringen betrekking hebben op de *capaciteit* tot moreel gedrag, of op de *inhoud* van de moraal. Het is goed mogelijk dat de menselijke capaciteit om morele oordelen te vellen grotendeels in evolutionaire termen kan worden verklaard, maar dat een adequate verklaring van de inhoud van die oordelen sterk moet leunen op inzichten uit andere disciplines, zoals de antropologie, sociologie, psychologie en geschiedschrijving.

Grondslagen

De filosofische metaethiek is van oudsher niet gericht op het verklaren van de moraal, maar op het rechtvaardigen ervan – en vooral op de aard van die rechtvaardiging. Drukt een moreel oordeel een waarheid uit? Is die waarheid objectief? En zo ja, hoe moeten we die claim van objectiviteit begrijpen? Hoewel metaethici zich ook bezighouden met tal van andere vragen die

betrekking hebben op de aard van de moraal, beperkt het huidige proefschrift zich tot een analyse van morele objectiviteit.

Meer specifiek is de vraag die ik centraal stel wat een moreel oordeel objectief maakt, ervan uitgaande dat we sommige morele oordelen inderdaad als objectief kunnen beschouwen. Zoals ik in het inleidende **hoofdstuk 1** uiteenzet kunnen we het metaethische landschap in tweeën delen, als het gaat om de grondslagen van morele objectiviteit. Aan de ene kant staan realistische theorieën, die stellen dat objectieve morele uitspraken standpuntonafhankelijk zijn. Een manier om deze claim van standpuntonafhankelijkheid te begrijpen is om ons een wereld voor te stellen, waarin een specifiek moreel oordeel (bijvoorbeeld het oordeel dat plezier goed is en pijn slecht) door niemand wordt onderschreven: zelfs in die wereld, stelt de realist, zou dit oordeel waar zijn (in de vooronderstelling dat dit een objectieve morele waarheid betreft).

Aan de andere kant staan antirealistische benaderingen, die een alternatieve kijk op morele objectiviteit propageren. Volgens één van die benaderingen – de positie die ik in dit proefschrift verdedig – zijn objectieve morele oordelen niet onafhankelijk van menselijke standpunten als zodanig, maar beschikken die oordelen over een andere eigenschap: zij kunnen de toets van (geïdealiseerde) kritische reflectie weerstaan vanuit een verscheidenheid aan menselijke standpunten. Anders gesteld: als de waarheid van een moreel oordeel invariant is met betrekking tot het standpunt van verscheidene morele actoren, omdat het de toets van beredeneerde kritiek vanuit al die standpunten kan weerstaan, dan kunnen we die waarheid als objectief beschouwen.

Genealogische kritiek

Moreel realisme en moreel antirealisme gaan gepaard met verschillende epistemologische theorieën – theorieën over hoe we morele kennis kunnen verwerven en rechtvaardigen. Een centrale these van dit proefschrift is dat een realistische morele epistemologie problematisch is. Die these is niet nieuw; onder meer Sharon Street heeft haar in recente jaren verdedigd. Het werk van Street vormt het beginpunt van veel metaethische beschouwingen in de eerste helft van dit proefschrift. Soms verdedig ik haar positie, soms bekritiseer ik die, en vaak probeer ik haar argumenten verder uit te bouwen en te perfectioneren.

Stel dat we, met de realist, vooronderstellen dat de waarheid van morele oordelen onafhankelijk is van menselijke standpunten. Wat garandeert dan dat onze morele oordelen met die onafhankelijke waarheden samenvallen? Eén van de epistemologische kritiekpunten is dat realisten die garantie niet kunnen geven: onze morele oordelen zijn, als we uitgaan van een realistische epistemologie, onbetrouwbaar. Realisten kunnen slechts vasthouden aan de vooronderstelling dat onze morele oordelen samenvallen met morele waarheden, als zij zich beroepen op een onwaarschijnlijke coïncidentie (zie **hoofdstuk 2**).

Realisten kunnen zich tegen deze kritiek verweren door te stellen dat onze morele oordelen niet *toevallig* samenvallen met morele waarheden, maar dat onze oordelen die waarheden *traceren*. In haar ‘Darwiniaanse Dilemma’ betoogt Street echter dat deze strategie realisten committeert aan een onhoudbare evolutionaire hypothese. Street concludeert dat een antirealistische metaethiek dit probleem kan omzeilen, en daarom de voorkeur verdient. In **hoofdstuk 3** laat ik zien dat sommige realisten Streets argument kunnen blokkeren, door te ontkennen dat de inhoud van onze morele oordelen sterk gekleurd is door evolutionaire invloeden. Maar vervolgens betoog ik dat realisten die deze strategie volgen nog steeds op een

onwaarschijnlijke hypothese leunen, zij het dat hun hypothese in dit geval niet onwaarschijnlijk is op evolutionaire, maar op historische gronden.

Dit laatste punt illustreer ik in **hoofdstuk 4**, door kritisch in te gaan op een recent argument van de realist Michael Huemer. Volgens Huemer is er gedurende de menselijke geschiedenis een sterke convergentie opgetreden in de richting van dezelfde morele waarden, en geeft deze trend inductieve steun aan de positie van realisten. Ik betoog, tegen Huemer, dat de historische trend die hij beschrijft juist beter kan worden verklaard op antirealistische gronden. Deze casus illustreert een van de originele inzichten die ik in dit proefschrift ontwikkel: niet-evolutionaire verklaringswijzen zijn essentieel voor het succes van pogingen om de plausibiliteit van moreel realisme op genealogische gronden te ondermijnen.

Psychologische kritiek

Behalve genealogische kritiek op moreel realisme, beroep ik me ook op recente inzichten uit de experimentele morele psychologie om te betogen dat een antirealistische metaethiek meer plausibel is dan haar realistische tegenhanger. Eén de voornaamste argumenten die van oudsher is ingebracht ter verdediging van moreel realisme is dat die positie het beste strookt met onze gebruikelijke morele intuïties. De intuïtie dat martelen objectief verkeerd is, bijvoorbeeld, drukt volgens realisten een onveranderlijke morele waarheid uit – een waarheid waar niet aan te tornen valt, wat wij er verder ook van denken. Die positie strookt met onze intuïties, stellen realisten: zij geeft de beste uitleg van ons gangbare morele spreken en denken.

In **hoofdstuk 5** plaats ik, deels op basis van de resultaten van psychologisch onderzoek, vraagtekens bij deze stelling. Tevens expliciteer ik in dit hoofdstuk de bovengenoemde antirealistische notie van morele objectiviteit: objectiviteit als ‘standpunt-invariantie’. Ik verdedig deze antirealistische theorie tegenover drie tegenargumenten, en betoog, op basis van psychologische studies, dat zij onze gebruikelijke intuïties over morele objectiviteit minstens zo goed verklaart als realistische theorieën. Bijgevolg komt een van de voornaamste argumenten ter verdediging van moreel realisme op losse schroeven te staan.

Deze psychologische kritiek op moreel realisme berust ten dele op de vooronderstelling dat resultaten uit de experimentele psychologie metaethische implicaties kunnen hebben. Maar wat is precies de metaethische relevantie van psychologische experimenten? In **hoofdstuk 6** geef ik een overzicht van recent onderzoek naar de vraag in hoeverre willekeurige respondenten morele objectivisten zijn, en welke metaethische lessen we daaruit kunnen trekken. Ik benadruk de methodologische haken en ogen van dit corpus van onderzoek, en onderstreep dat haar metaethische implicaties in meerdere opzichten niet vanzelfsprekend zijn, maar filosofische interpretatie en argumentatie behoeven. Maar ik benadruk ook dat veel bestaande metaethische argumenten in ieder geval ten dele op empirische aannames berusten, en dat hun deugdelijkheid daarom empirisch kan worden getoetst. Dat geldt ook voor de zogenaamde ‘presumptieve argumenten’ waar realisten zich van oudsher op beroepen. Ik concludeer dat de resultaten van psychologisch onderzoek weliswaar niet volstaan om moreel realisme geheel te ontkrachten, maar wel om één van de belangrijkste argumenten ten faveure van realisme te weerleggen.

In **hoofdstuk 7** voeg ik nog een verdere overweging toe die bijdraagt aan de weerlegging van ‘presumptieve argumenten’ ten faveure van moreel realisme. Onze objectivistische neigingen kunnen ook vanuit een evolutionair perspectief worden verklaard,

zonder ons op moreel realisme te beroepen. Sterker nog, volgens recente theorieën van enkele prominente evolutionair psychologen is het geloof in de objectiviteit van morele oordelen doorgaans adaptief, en heeft dat geloof een belangrijke rol gespeeld in de evolutie van sociaal menselijk gedrag. Niet alleen de oorsprong van de moraal als zodanig, zo suggereren deze theorieën, maar ook de oorsprong van *morele objectiviteit* kan, tot op zekere hoogte, in evolutionaire termen worden verklaard.

Metaethiek en wetenschap

Dit proefschrift gebruikt het analytisch instrumentarium van de metaethiek, maar beroept zich ook op inzichten uit de de experimentele morele psychologie, de geschiedschrijving, de evolutionaire biologie, de evolutionaire antropologie en de evolutionaire psychologie. Tevens bevat het proefschrift een klein element van experimenteel werk dat ik zelf heb verricht, in samenwerking met Michael Klenk (zie **hoofdstuk 7**). Dit beroep op empirische bronnen is in de hedendaagse metaethiek geen orthodoxie. Maar het is wel wenselijk, ook voor metaethici, zo hoop ik dat dit proefschrift duidelijk maakt. We verwerven een rijker filosofisch inzicht door metaethische theorieën te verbinden met onze wetenschappelijke kennis.

Voor elke metaethische theorie geldt dan ook het *desideratum* dat zij op zijn minst *verenigbaar* moet zijn met onze beste wetenschap. Dat *desideratum* is, uiteindelijk, doorslaggevend in de metaethische evaluatie van moreel realisme tegenover moreel antirealisme. Van oudsher is gedacht dat realistische theorieën beter in staat zijn om onze morele intuïties te waarborgen, terwijl antirealisten beter in staat zijn om hun theorieën wetenschappelijk te staven. Mijn genealogische kritiek op moreel realisme onderschrijft dit laatste punt, terwijl mijn psychologische kritiek vraagtekens plaatst bij de stelling dat moreel realisme beter aansluit bij onze intuïties. Aangezien het belangrijkste argument om moreel realisme te prefereren ten opzichte van moreel antirealisme geen standhoudt, zijn de wetenschappelijke tekortkomingen van de realistische positie funest voor haar verdedigers, en is een antirealistische metaethiek superieur.

Curriculum Vitae

I was born in Amsterdam (6 March 1987) and completed my high school education at the Stedelijk Gymnasium Haarlem (2005). My university education started with a year of liberal arts studies at Lewis & Clarke College, Portland (OR). In 2006 I began to study at the University of Amsterdam (UvA), where I obtained Bachelor degrees in both history and philosophy (2010, both *cum laude*). I also pursued my Master's degree in philosophy (2012, *cum laude*) at the University of Amsterdam, with exchange semesters at the University of Edinburgh (2010) and the University of Melbourne (2011). Subsequently, the Free University of Amsterdam (VU) generously funded me to pursue an extra research year (2012-2013) under the guidance of now Emeritus Professor Hans Radder. In 2014 I started PhD research at Utrecht University (UU). From the second half of 2017 until early 2018 I was a visiting fellow at Harvard University.

As a university teacher, I have been a teaching assistant on a course on the Philosophy of Psychology at the Free University of Amsterdam, have taught multiple introductory courses in Ethics and Political Philosophy at the University of Amsterdam and have co-taught a seminar on Moral Psychology for Master's students at Utrecht University.

During my PhD studies I have co-organized various symposia and conferences. From 2014 until 2017 I was the convener of the monthly OZSW Research Seminar for Analytic Philosophy. Additionally, since 2017 I have been an editor of *Wijzgerig Perspectief*, where I have been the main editor for two special issues: 'Toeval' (#2, 2017) and 'Technologie en Ideologie' (#2, 2018).

Since 2012 I have been involved in philosophical writing for a greater public, mostly by working for *Filosofie Magazine* – first as an intern, subsequently as an editor and later as a freelance journalist and columnist. When I started my PhD project at Utrecht University in 2014, I was allowed to retain one weekday to work as an editor and writer. This has enabled me, over the course of the last five years, to publish three philosophy books geared towards a broader audience. In 2014 I co-authored *Boeddhisme voor Denkers* (Ten Have), together with Han de Wit, which has subsequently been published in French as *Bouddha Philosophie*. In 2018 I published two books, *De andere afslag: Hoe had het leven anders kunnen lopen?* (Amsterdam University Press) and *Food for Thought: Wereldfilosofie aan de keukentafel* (Ten Have).

My PhD research is featured on the OZSW website *The Young Philosophers*: www.theyoungphilosophers.org/interview/jeroen-hopster-1987.

Additional Scholarly Publications (2014–2018)

These are publications on topics related to the research project, which have not been incorporated in the chapters of the dissertation.

Book

Hopster, J.K.G. 2018. *De Andere Afslag: Hoe had het leven anders kunnen lopen?* Amsterdam: Amsterdam University Press.

Book Chapters

Hopster, J.K.G. 2016. Zijn jouw intuïties de mijne? In: M. Janssens, L. Bauwens, S. van Balen, T. Tettero, F. Philips, eds. *Brainwash: De belangrijkste vragen van nu*. Amsterdam: Nijgh & van Ditmar.

Hopster, J.K.G. 2015. Seculier Zelfvertrouwen. In: F. Bolkestein, P. Cliteur, M. Fennema, eds. *De succesvolle mislukking van Europa*. Leiden: Leiden Publications.

Reviews

Hopster, J.K.G. 2017. Evolutionary Chance and Contingency: In Search for Systematics. *Metascience* 26, 481-485.

Hopster, J.K.G. 2017. Review of P. Singer, *Ethics in the Real World: 82 Brief Essays on Things that Matter*. *Erasmus Journal for Philosophy and Economics* 10, 120-123.

Hopster, J.K.G. 2017. Review of U. Leibowitz and N. Sinclair eds. *Explanation in Ethics and Mathematics*. *Ethical Perspectives* 24, 318-320.

Hopster, J.K.G. 2017. Review of Robert Kretsinger, *History and Philosophy of Biology*. *The Quarterly Review of Biology* 92, 84.

Hopster, J.K.G. 2016. Review of J. Harris, *How to be Good: The Possibility of Moral Enhancement*. *Ethical Perspectives* 23, 757-759.

Hopster, J.K.G. 2016. Review of T. Heams, P. Huneman, G. Lecointre and M. Silberstein eds., *Handbook of Evolutionary Thinking in the Sciences*. *The Quarterly Review of Biology* 91, 362-363.

Hopster, J.K.G. 2016. De genengekte voorbij? *Filosofie Magazine* 24, nr. 6.

Hopster, J.K.G. 2016. Het brein als katapult. *Filosofie Magazine* 24, nr. 5.

Hopster, J.K.G. 2015. Darwinisme van de gaten. *Filosofie Magazine* 23, nr 1.

Popular Essays

Hopster, J.K.G. 2016. Grensgevallen. *Filosofie Magazine* 24, nr 4.

Additional Scholarly Publications (2014–2018)

Hopster, J.K.G. 2015. De nieuwe erfelijkheidsleer. Hoe onze levensstijl de genen van het nageslacht verandert. *Filosofie Magazine* 23, nr. 12.

Hopster, J.K.G. 2015. De schaduwkant van empathie.
www.filosofie.nl/nl/artikel/42300/index.html

Hopster, J.K.G. 2014. Foley's hoofd en de morele onderbuik.
www.filosofie.nl/nl/artikel/41385/index.html

Quaestiones Infinitae

PUBLICATIONS OF THE DEPARTMENT OF PHILOSOPHY AND RELIGIOUS STUDIES

- VOLUME 21. D. VAN DALEN, *Torens en Fundamenten* (valedictory lecture), 1997.
- VOLUME 22. J.A. BERGSTRA, W.J. FOKKINK, W.M.T. MENNEN, S.F.M. VAN VLIJMEN, *Spoorweglogica via EURIS*, 1997.
- VOLUME 23. I.M. CROESE, *Simplicius on Continuous and Instantaneous Change* (dissertation), 1998.
- VOLUME 24. M.J. HOLLENBERG, *Logic and Bisimulation* (dissertation), 1998.
- VOLUME 25. C.H. LEIJENHORST, *Hobbes and the Aristotelians* (dissertation), 1998.
- VOLUME 26. S.F.M. VAN VLIJMEN, *Algebraic Specification in Action* (dissertation), 1998.
- VOLUME 27. M.F. VERWEIJ, *Preventive Medicine Between Obligation and Aspiration* (dissertation), 1998.
- VOLUME 28. J.A. BERGSTRA, S.F.M. VAN VLIJMEN, *Theoretische Software-Engineering: kenmerken, faseringen en classificaties*, 1998.
- VOLUME 29. A.G. WOUTERS, *Explanation Without A Cause* (dissertation), 1999.
- VOLUME 30. M.M.S.K. SIE, *Responsibility, Blameworthy Action & Normative Disagreements* (dissertation), 1999.
- VOLUME 31. M.S.P.R. VAN ATTEN, *Phenomenology of choice sequences* (dissertation), 1999.
- VOLUME 32. V.N. STEBLETSOVA, *Algebras, Relations and Geometries (an equational perspective)* (dissertation), 2000.
- VOLUME 33. A. VISSER, *Het Tekst Continuüm* (inaugural lecture), 2000.
- VOLUME 34. H. ISHIGURO, *Can we speak about what cannot be said?* (public lecture), 2000.
- VOLUME 35. W. HAAS, *Haltlosigkeit; Zwischen Sprache und Erfahrung* (dissertation), 2001.
- VOLUME 36. R. POLI, *ALWIS: Ontology for knowledge engineers* (dissertation), 2001.
- VOLUME 37. J. MANSFELD, *Platonische Briefschrijverij* (valedictory lecture), 2001.
- VOLUME 37A. E.J. BOS, *The Correspondence between Descartes and Henricus Regius* (dissertation), 2002.
- VOLUME 38. M. VAN OTEGEM, *A Bibliography of the Works of Descartes (1637-1704)* (dissertation), 2002.
- VOLUME 39. B.E.K.J. GOOSSENS, *Edmund Husserl: Einleitung in die Philosophie: Vorlesungen 1922/23* (dissertation), 2003.
- VOLUME 40. H.J.M. BROEKHUIJSE, *Het einde van de sociaaldemocratie* (dissertation), 2002.
- VOLUME 41. P. RAVALLI, *Husserls Phänomenologie der Intersubjektivität in den Göttinger Jahren: Eine kritisch-historische Darstellung* (dissertation), 2003.
- VOLUME 42. B. ALMOND, *The Midas Touch: Ethics, Science and our Human Future* (inaugural lecture), 2003.
- VOLUME 43. M. DÜWELL, *Morele kennis: over de mogelijkheden van toegepaste ethiek* (inaugural lecture), 2003.
- VOLUME 44. R.D.A. HENDRIKS, *Metamathematics in Coq* (dissertation), 2003.
- VOLUME 45. TH. VERBEEK, E.J. BOS, J.M.M. VAN DE VEN, *The Correspondence of René Descartes: 1643*, 2003.
- VOLUME 46. J.J.C. KUIPER, *Ideas and Explorations: Brouwer's Road to Intuitionism* (dissertation), 2004.
- VOLUME 47. C.M. BEKKER, *Rechtvaardigheid, Onpartijdigheid, Gender en Sociale*

- Diversiteit; Feministische filosofen over recht doen aan vrouwen en hun onderlinge verschillen* (dissertation), 2004.
- VOLUME 48. A.A. LONG, *Epicetus on understanding and managing emotions* (public lecture), 2004.
- VOLUME 49. J.J. JOOSTEN, *Interpretability formalized* (dissertation), 2004.
- VOLUME 50. J.G. SIJMONS, *Phänomenologie und Idealismus: Analyse der Struktur und Methode der Philosophie Rudolf Steiners* (dissertation), 2005.
- VOLUME 51. J.H. HOOGSTAD, *Time tracks* (dissertation), 2005.
- VOLUME 52. M.A. VAN DEN HOVEN, *A Claim for Reasonable Morality* (dissertation), 2006.
- VOLUME 53. C. VERMEULEN, *René Descartes, Specimina philosophiae: Introduction and Critical Edition* (dissertation), 2007.
- VOLUME 54. R.G. MILLIKAN, *Learning Language without having a theory of mind* (inaugural lecture), 2007.
- VOLUME 55. R.J.G. CLAASSEN, *The Market's Place in the Provision of Goods* (dissertation), 2008.
- VOLUME 56. H.J.S. BRUGGINK, *Equivalence of Reductions in Higher-Order Rewriting* (dissertation), 2008.
- VOLUME 57. A. KALIS, *Failures of agency* (dissertation), 2009.
- VOLUME 58. S. GRAUMANN, *Assistierte Freiheit* (dissertation), 2009.
- VOLUME 59. M. AALDERINK, *Philosophy, Scientific Knowledge, and Concept Formation in Geulincx and Descartes* (dissertation), 2010.
- VOLUME 60. I.M. CONRADIE, *Seneca in his cultural and literary context: Selected moral letters on the body* (dissertation), 2010.
- VOLUME 61. C. VAN SIJL, *Stoic Philosophy and the Exegesis of Myth* (dissertation), 2010.
- VOLUME 62. J.M.I.M. LEO, *The Logical Structure of Relations* (dissertation), 2010.
- VOLUME 63. M.S.A. VAN HOUTE, *Seneca's theology in its philosophical context* (dissertation), 2010.
- VOLUME 64. F.A. BAKKER, *Three Studies in Epicurean Cosmology* (dissertation), 2010.
- VOLUME 65. T. FOSSEN, *Political legitimacy and the pragmatic turn* (dissertation), 2011.
- VOLUME 66. T. VISAK, *Killing happy animals. Explorations in utilitarian ethics.* (dissertation), 2011.
- VOLUME 67. A. JOOSSE, *Why we need others: Platonic and Stoic models of friendship and self-understanding* (dissertation), 2011.
- VOLUME 68. N. M. NIJSINGH, *Expanding newborn screening programmes and strengthening informed consent* (dissertation), 2012.
- VOLUME 69. R. PEELS, *Believing Responsibly: Intellectual Obligations and Doxastic Excuses* (dissertation), 2012.
- VOLUME 70. S. LUTZ, *Criteria of Empirical Significance* (dissertation), 2012
- VOLUME 70A. G.H. BOS, *Agential Self-consciousness, beyond conscious agency* (dissertation), 2013.
- VOLUME 71. F.E. KALDEWAI, *The animal in morality: Justifying duties to animals in Kantian moral philosophy* (dissertation), 2013.
- VOLUME 72. R.O. BUNING, *Henricus Reneri (1593-1639): Descartes' Quartermaster in Aristotelian Territory* (dissertation), 2013.
- VOLUME 73. I.S. LÖWISCH, *Genealogy Composition in Response to Trauma: Gender and Memory in 1 Chronicles 1-9 and the Documentary Film 'My Life Part 2'* (dissertation), 2013.
- VOLUME 74. A. EL KHAIRAT, *Contesting Boundaries: Satire in Contemporary Morocco* (dissertation), 2013.

- VOLUME 75 A. KROM, *Not to be sneezed at. On the possibility of justifying infectious disease control by appealing to a mid-level harm principle* (dissertation), 2014.
- VOLUME 76 Z. PALL, *Salafism in Lebanon: local and transnational resources* (dissertation), 2014.
- VOLUME 77 D. WAHID, *Nurturing the Salafi Manhaj: A Study of Salafi Pesantrens in Contemporary Indonesia* (dissertation), 2014.
- VOLUME 78 B.W.P VAN DEN BERG, *Speelruimte voor dialoog en verbeelding. Basisschoolleerlingen maken kennis met religieuze verhalen* (dissertation), 2014.
- VOLUME 79 J.T. BERGHUIJS, *New Spirituality and Social Engagement* (dissertation), 2014.
- VOLUME 80 A. WETTER, *Judging By Her. Reconfiguring Israel in Ruth, Esther and Judith* (dissertation), 2014.
- VOLUME 81 J.M. MULDER, *Conceptual Realism. The Structure of Metaphysical Thought* (dissertation), 2014.
- VOLUME 82 L.W.C. VAN LIT, *Eschatology and the World of Image in Suhrawardī and His Commentators* (dissertation), 2014.
- VOLUME 83 P.L. LAMBERTZ, *Divisive matters. Aesthetic difference and authority in a Congolese spiritual movement 'from Japan'* (dissertation), 2015.
- VOLUME 84 J.P. GOUDSMIT, *Intuitionistic Rules: Admissible Rules of Intermediate Logics* (dissertation), 2015.
- VOLUME 85 E.T. FEIKEMA, *Still not at Ease: Corruption and Conflict of Interest in Hybrid Political Orders* (dissertation), 2015.
- VOLUME 86 N. VAN MILTENBURG, *Freedom in Action* (dissertation), 2015.
- VOLUME 86A P. COPPENS, *Seeing God in This world and the Otherworld: Crossing Boundaries in Sufi Commentaries on the Qur'ān* (dissertation), 2015.
- VOLUME 87 D.H.J. JETHRO, *Aesthetics of Power: Heritage Formation and the Senses in Post-apartheid South Africa* (dissertation), 2015.
- VOLUME 88 C.E. HARNACKE, *From Human Nature to Moral Judgement: Reframing Debates about Disability and Enhancement* (dissertation), 2015.
- VOLUME 89 X. WANG, *Human Rights and Internet Access: A Philosophical Investigation* (dissertation), 2016.
- VOLUME 90 R. VAN BROEKHOVEN, *De Bewakers Bewaakt: Journalistiek en leiderschap in een gemediatiseerde democratie* (dissertation), 2016.
- VOLUME 91 A. SCHLATMANN, *Shi'i Muslim youth in the Netherlands: Negotiating Shi'i fatwas and rituals in the Dutch context* (dissertation), 2016.
- VOLUME 92 M.L. VAN WIJNGAARDEN, *Schitterende getuigen. Nederlands luthers avondmaalsgerei als indenteitsdrager van een godsdienstige minderheid* (dissertation), 2016.
- VOLUME 93 S. COENRADIE, *Vicarious substitution in the literary work of Shūsaku Endō. On fools, animals, objects and doubles* (dissertation), 2016.
- VOLUME 94 J. RAJAIHAH, *Dalit Humanization. A quest based on M.M. Thomas' theology of salvation and humanization* (dissertation), 2016.
- VOLUME 95 D.L.A. OMETTO, *Freedom & Self-knowledge* (dissertation), 2016.
- VOLUME 96 Y. YALDIZ, *The Afterlife in Mind: Piety and Renunciatory Practice in the 2nd/8th- and early 3rd/9th-Century Books of Renunciation (Kutub al-Zuhd)* (dissertation), 2016.
- VOLUME 97 M.F. BYSKOV, *Between experts and locals. Towards an inclusive framework for a development agenda* (dissertation), 2016.
- VOLUME 98 A. RUMBERG, *Transitions toward a Semantics for Real Possibility*

- (dissertation), 2016.
- VOLUME 99 S. DE MAAGT, *Constructing Morality: Transcendental Arguments in Ethics* (dissertation), 2017.
- VOLUME 100 S. BINDER, *Total Atheism* (dissertation), 2017.
- VOLUME 101 T. GIESBERS, *The Wall or the Door: German Realism around 1800*, (dissertation), 2017.
- VOLUME 102 P. SPERBER, *Kantian Psychologism* (dissertation), 2017.
- VOLUME 103 J.M. HAMER, *Agential Pluralism: A Philosophy of Fundamental Rights* (dissertation), 2017.
- VOLUME 104 M. IBRAHIM, *Sensational Piety: Practices of Mediation in Christ Embassy and Nasfat* (dissertation), 2017.
- VOLUME 105 R.A.J. MEES, *Sustainable Action, Perspectives for Individuals, Institutions, and Humanity* (dissertation), 2017.
- VOLUME 106 A.A.J. POST, *The Journey of a Taymiyyan Sufi: Sufism Through the Eyes of Imād al-Dīn Aḥmad al-Wāsiṭī (d. 711/1311)* (dissertation), 2017.
- VOLUME 107 F.A. FOGUE KUATE, *Médias et coexistence entre Musulmans et Chrétiens au Nord-Cameroun: de la période coloniale Française au début du XXIème siècle* (dissertation), 2017.
- VOLUME 108 J. KROESBERGEN-KAMPS, *Speaking of Satan in Zambia. The persuasiveness of contemporary narratives about Satanism* (dissertation), 2018.
- VOLUME 109 F. TENG, *Moral Responsibilities to Future Generations. A Comparative Study on Human Rights Theory and Confucianism* (dissertation), 2018.
- VOLUME 110 H.W.A. DUIJF, *Let's Do It! Collective Responsibility, Joint Action, and Participation* (dissertation), 2018.
- VOLUME 111 R.A. CALVERT, *Pilgrims in the port. Migrant Christian communities in Rotterdam* (dissertation), 2018.
- VOLUME 112 W.P.J.L. VAN SAANE, *Protestant Mission Partnerships: The Concept of Partnership in the History of the Netherlands Missionary Council in the Twentieth Century* (dissertation), 2018.
- VOLUME 113 D.K. DÜRING, *Of Dragons and Owls. Rethinking Chinese and Western narratives of modernity* (dissertation), 2018.
- VOLUME 114 H. ARENTSHORST, *Perspectives on freedom. Normative and political views on the preconditions of a free democratic society* (dissertation), 2018.
- VOLUME 115 M.B.O.T. KLENK, *Survival of Defeat – Evolution, Moral Objectivity, and Undercutting* (dissertation), 2018.
- VOLUME 116 J.H. HOEKJEN, *Pars melior nostri. The Structure of Spinoza's Intellect* (dissertation), 2018.
- VOLUME 117 C.J. MUDDE, *Rouwen in de marge. De materiële rouwcultuur van de katholieke geloofsgemeenschap in vroegmodern Nederland* (dissertation), 2018.
- VOLUME 118 K. GRIT, *"Christians by Faith, Pakistani by Citizenship". Negotiating Christian Identity in Pakistan* (dissertation), 2019.
- VOLUME 119 J.K.G. HOPSTER, *Moral Objectivity: Origins and Foundations* (dissertation), 2019.