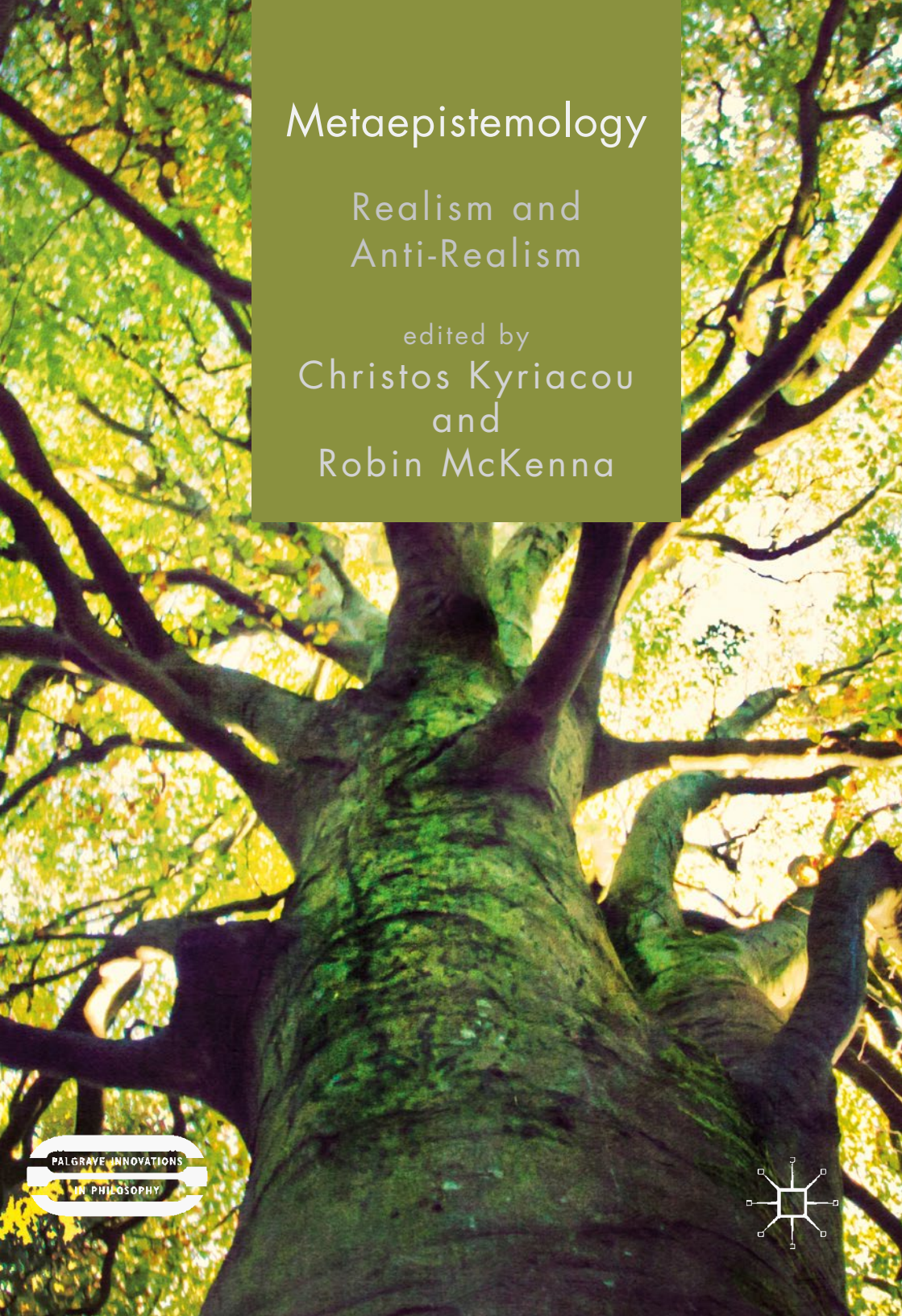
Metaepistemology

Realism and Anti-Realism

edited by
Christos Kyriacou
and
Robin McKenna



PALGRAVE INNOVATIONS
IN PHILOSOPHY

Palgrave Innovations in Philosophy

Series Editors
Vincent Hendricks
University of Copenhagen
Copenhagen, Denmark

Duncan Pritchard
University of Edinburgh
Edinburgh, UK

Palgrave Innovations in Philosophy is a new series of monographs. Each book in the series will constitute the 'new wave' of philosophy, both in terms of its topic and the research profile of the author. The books will be concerned with exciting new research topics of particular contemporary interest, and will include topics at the intersection of Philosophy and other research areas. They will be written by up-and-coming young philosophers who have already established a strong research profile and who are clearly going to be leading researchers of the future. Each monograph in this series will provide an overview of the research area in question while at the same time significantly advancing the debate on this topic and giving the reader a sense of where this debate might be heading next. The books in the series would be of interest to researchers and advanced students within philosophy and its neighboring scientific environments.

More information about this series at
http://www.palgrave.com/gp/series/14689

Christos Kyriacou • Robin McKenna
Editors

Metaepistemology

Realism and Anti-Realism

palgrave
macmillan

Editors
Christos Kyriacou
University of Cyprus
Nicosia, Cyprus

Robin McKenna
University of Vienna
Vienna, Austria

Palgrave Innovations in Philosophy
ISBN 978-3-319-93368-9 ISBN 978-3-319-93369-6 (eBook)
https://doi.org/10.1007/978-3-319-93369-6

Library of Congress Control Number: 2018951643

© The Editor(s) (if applicable) and The Author(s) 2018
This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Cover illustration: Image by Catherine MacBride https://www.gettyimages.co.uk/license/141377343

This Palgrave Macmillan imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Christos: To Sophocles and Helen
Robin: To Edith and Sophie

Acknowledgments

We would like to thank our contributors for providing us with an excellent set of essays, and for their assistance throughout the editing process. Thanks also to Brendan George and April James at Palgrave Macmillan for their prompt responses to numerous queries, and for supporting the project. Robin McKenna acknowledges support from the ERC Advanced Grant Project "The Emergence of Relativism" (Grant Number 339382).

Contents

Introduction 1
Christos Kyriacou and Robin McKenna

Part I Epistemic Realism 15

The Core Expressivist Manoeuvre 17
Terence Cuneo

Epistemic Reductionism and the Moral-Epistemic Disparity 45
Chris Heathwood

From *Moral* Fixed Points to *Epistemic* Fixed Points 71
Christos Kyriacou

Normative Reasons for Mentalism 97
Eva Schmidt

Epistemic Consequentialism: Haters Gonna Hate 121
Nathaniel Sharadin

Part II Epistemic Anti-realism 145

Knowledge, Reasons, and Errors About Error Theory 147
Charles Côté-Bouchard and Clayton Littlejohn

Constitutivism About Epistemic Normativity 173
Christopher Cowie and Alexander Greenberg

Correctness and Goodness 197
Allan Hazlett

The Genealogy of Relativism and Absolutism 217
Martin Kusch and Robin McKenna

Reasons Primitivism and Epistemic Expressivism 241
Teemu Toppinen

Part III Beyond the Realism/Anti-realism Divide 265

What Anti-realism About Hinges Could Possibly Be 267
Annalisa Coliva

Epistemic Schmagency? 289
A. K. Flowerree

Index 311

Notes on Contributors

Annalisa Coliva is Professor and Chair of the Department of Philosophy at the University of California, Irvine. She has published extensively on hinge epistemology, its historical origins and on self-knowledge. Her major monographs are *Extended Rationality. A Hinge Epistemology* (Palgrave 2016), *The Varieties of Self-Knowledge* (Palgrave 2015), *Moore and Wittgenstein. Scepticism, Certainty and Common Sense* (Palgrave 2010). Her book, *Relativism*, co-authored with Maria Baghramian, is forthcoming with Routledge (2018).

Charles Côté-Bouchard is a FRQSC postdoctoral fellow at Rutgers, the State University of New Jersey. He received his Ph.D. from King's College London in 2017. His current areas of research are epistemology and ethics. His articles concerning metaepistemology and epistemic normativity appear in *Synthese*, *Philosophical Studies*, the *International Journal of Philosophical Studies*, *Dialogue*, and *Manuscrito*.

Christopher Cowie is Assistant Professor at Durham University. He was previously a Research Fellow at the University of Cambridge. He works in metaethics and moral philosophy. He has published in journals including *The Philosophical Quarterly, Philosophical Studies, The Australasian Journal of Philosophy, Philosophy Compass* and *Synthese*.

Terence Cuneo is the Marsh Chair of Intellectual and Moral Philosophy at the University of Vermont. He is the author of *The Normative Web* (Oxford, 2007), *Speech and Morality* (Oxford, 2014), *Ritualized Faith* (Oxford, 2016), and (co-author of) *Understanding Morality* (Oxford, forthcoming).

A. K. Flowerree is a Postdoctoral Fellow at the Center for Contemporary Epistemology and the Kantian Tradition (CONCEPT) at the University of Cologne, and will be Assistant Professor at Texas Tech University beginning 2018. Her research focuses on the intersection of epistemology, ethics and metaethics.

Alexander Greenberg is a Lecturer in Philosophy at Birkbeck, University of London and a Stipendiary Lecturer in Philosophy at Lady Margaret Hall, University of Oxford. His research is at the intersection of epistemology and philosophy of mind, and focuses on epistemological, psychological, and normative questions about belief.

Allan Hazlett (PhD, Brown University, 2006) is Assistant Professor of Philosophy at Washington University in St Louis. Along with several papers, he is the author of two books: *A Luxury of the Understanding: On the Value of True Belief* (Oxford University Press, 2013) and *A Critical Introduction to Skepticism* (Bloomsbury Publishing, 2014). He works on various topics in social epistemology, ethics, and moral psychology. Among his recent papers are pieces on the "guise of the good" thesis, aesthetic testimony, and epistemic injustice.

Chris Heathwood is Associate Professor of Philosophy at the University of Colorado Boulder. He works mainly in theoretical ethics: in particular, in the philosophy of well-being, where he defends a desire-satisfaction theory, and in metaethics, where he explores the tenability of a robust moral realism. He also has interests in metaepistemology and metaphysics. His papers have appeared in *Noûs*, *Oxford Studies in Metaethics*, *Philosophical Studies*, the *Australasian Journal of Philosophy*, and other leading venues.

Martin Kusch is Professor for Philosophy of Science and Epistemology at the University of Vienna. His current work focuses on systematic and historical questions concerning epistemic relativism.

Christos Kyriacou is a Lecturer at the University of Cyprus. His main research interests lie in epistemology, metaethics and their intersection. He has published in *Synthese, Erkenntnis, Philosophical Psychology, Journal of Ethics and Social Philosophy, Journal of Philosophical Research, Philosophia* and elsewhere.

Clayton Littlejohn is Reader in Philosophy at King's College London. He is the author of *Justification and the Truth-Connection* (Cambridge University Press, 2012) and writes mainly on issues in epistemology and ethical theory.

Robin McKenna is a Postdoctoral Fellow at the University of Vienna. He has published several papers on a cluster of topics at the intersections of epistemology, philosophy of language and ethics in journals such as *Australasian Journal of Philosophy, Ethics* and *Synthese*. He is currently working on a book and a series of papers that argue that knowledge is a social kind.

Eva Schmidt is Assistant Professor at the University of Zürich. Her work focuses on practical and epistemic normativity, disjunctivist accounts of justification, and the nature of perception. In her current research project, she defends a reasons-first view of epistemic justification. Her book *Modest Nonconceptualism: Epistemology, Phenomenology, and Content* was published in 2015.

Nathaniel Sharadin is an Assistant Professor of Philosophy at The College of New Jersey. His research focuses on normative and meta-normative issues in both ethics and epistemology.

Teemu Toppinen is a postdoctoral fellow at the University of Helsinki. He is currently developing a novel, relational form of expressivism and exploring the prospects and implications of quasi-realist expressivism in relation to issues ranging from disagreement and deference to moral worth and transcendental arguments. While most of his published work deals in all sorts of things meta-normative (e.g., *Oxford Studies in Metaethics*), he has also dabbled in normative ethics (e.g., *The Philosophical Quarterly*).

List of Tables

Constitutivism About Epistemic Normativity
Table 1 Dispositional accounts 178
Table 2 TRUTH NORM 179



Introduction

Christos Kyriacou and Robin McKenna

1 Metaepistemology: Realism, Anti-realism, Constitutivism/Constructivism

Metaepistemology is, roughly, the branch of epistemology that asks questions about the existence, nature and authority of epistemic facts and reasons.[1] Typically, the metaepistemological landscape is divided into three parts on the basis of the ontological distinction between *epistemic realism*, *epistemic anti-realism* and *constitutivism\constructivism*. According to the standard reading of the distinction, epistemic realism is the position that asserts the existence of *mind-independent* epistemic facts and reasons for belief.[2] The nature of mind-independent epistemic facts and reasons for belief may be understood either in broadly naturalist (e.g. consequentialist) or non-naturalist ways, depending on the type of

C. Kyriacou (✉)
Department of Classics and Philosophy, University of Cyprus, Nicosia, Cyprus
e-mail: ckiriakou@gmail.com

R. McKenna
Department of Philosophy, University of Vienna, Vienna, Austria
e-mail: rbnmckenna@gmail.com

© The Author(s) 2018

C. Kyriacou, R. McKenna (eds.), *Metaepistemology*, Palgrave Innovations in Philosophy, https://doi.org/10.1007/978-3-319-93369-6_1

epistemic realism one adopts.[3] In any case, epistemic realism typically considers mind-independent epistemic facts as *categorically* reason-giving to rational agents.[4] That is, as facts that exert authority over us and constrain what we ought to believe if we are rational, independently of our desires, goals, wishes, intentions, whims, et cetera.

Epistemic anti-realism is the denial of epistemic realism. It is the position that denies the existence of mind-independent epistemic facts and reasons for belief and proposes instead that epistemic facts and reasons for belief are *mind-dependent*, that is, they are up to us in some sense (e.g. relativism, expressivism).[5] Typically, anti-realist theories are theories that sit well with a broadly naturalist picture because they ground epistemic facts and reasons for belief in natural facts about human psychology, our evolutionary origins and social predicament et cetera. Of course, different anti-realist theories can tell very different stories about epistemic facts and reasons for belief, depending on the contours of their commitments (expressivist, relativist etc.). In any case, epistemic anti-realism considers mind-dependent epistemic facts as *non-categorically* (or instrumentally) reason-giving to rational agents.[6] That is, as facts that exert some authority over us and constrain what we ought to believe, depending on our particular desires, goals, intentions, plans, et cetera. For epistemic anti-realism, there is no absolute fact of the matter about what one ought to believe if epistemically rational.

Some philosophers, however, have grown wary of the distinction between realism and anti-realism. On the one hand, realism seems beset by ontological and epistemological problems, such as whether such facts can plausibly exist, and, even if they exist, how we can have a plausible epistemology detailing how we can come to know such facts and have such reasons for belief. On the other hand, anti-realism seems beset by problems concerning objectivity and truth, such as how it could be that there is no absolute fact of the matter about what we ought to believe. For with its denial of mind-independent facts about what we ought to believe, anti-realism would seem to entail that *prima facie* normatively irrelevant factors (such as our social predicament and our evolutionary history) can ground epistemic normativity, or that "anything goes" in matters of ought to believe, or even worse, radical scepticism: there is no (objective) truth or knowledge to be had.

Concerned by these (and many other) problems of realism and anti-realism, some philosophers have attempted, in a Kantian spirit, to forge some middle ground between the two alternatives and ground epistemic normativity in facts constitutive of agency.[7] Accordingly, this approach is often called *constitutivism* (or sometimes constructivism). Constitutivism is a broadly Kantian project in the sense that the necessary preconditions constitutive of agency are searched for with the hope that they will ground normativity (moral, epistemic).[8] Given that constitutivism seeks to ground normativity in facts constitutive of agency, these facts are more naturally understood as *mind-dependent* facts. But although these facts are mind-dependent, they are *prima facie* not normatively irrelevant (or arbitrary) because they are necessary preconditions constitutive of agency. That is, insofar as one is an agent, one cannot help but be committed to these constitutive preconditions of agency. Accordingly, facts constitutive of agency can provide us, insofar as we are agents, with *categorical* reasons for belief (and without the realist's metaphysical baggage).

In the epistemic case, constitutivism has asked us to pay attention to belief, which is the mental state-vehicle of tracking the truth and coming to knowledge. Constitutivists have suggested that belief is a normative state that constitutively "aims at truth" and, therefore, truth must be its correctness condition. Besides, when we deliberate about what to believe, all other things equal, we seem to be looking for truth-related considerations that could justify our belief. Truth is thus the condition that should regulate what one ought to believe. Constitutivists have concluded that truth is the norm of correct belief (and assertion).

The constitutivist approach would seem to avoid the basic problems that beset realism and anti-realism. In regard to realism, there would be no difficult ontological and epistemological questions. Epistemic normativity is grounded in facts constitutive of agency and we can get to know such facts by paying close attention to the operation of doxastic deliberation in human agency. In regard to anti-realism, there would be epistemic objectivity and truth to be had. There would be things we ought to believe insofar as we are agents and objective truths to be had. Thus, it would not be the case that *prima facie* normatively irrelevant factors could ground epistemic normativity or that "anything goes" in doxastic matters or even that scepticism would threaten knowledge.

Now, the last 20 years or so has seen a growing interest in the aforementioned questions about the existence, nature and authority of epistemic facts and reasons. Realist, anti-realist and constitutivist approaches have been proposed and much discussion has ensued. On the one hand, this interest was facilitated by the general blossoming of the field of epistemology that has led to an ever-increasing sensitivity to metaepistemological questions.[9] On the other hand, this interest was facilitated by the widespread recognition that epistemology is as much a normative discipline as ethics is and that the two domains share structural similarities (reasons, facts, motivation, objectivity, truth etc.).[10]

The recognition that epistemology seems just as normative as ethics has led to various interesting proposals about the interconnections and analogies between the two domains. Philosophers, either coming from the direction of metaethics or the direction of epistemology, have explored various parity arguments (so-called "companions-in-guilt arguments") and analogies between the two domains. The interconnections and analogies explored in the literature are too varied to all be presented in a brief introduction, but here is a sample of these explorations.

Terence Cuneo (2007) has argued extensively in support of the moral-epistemic parity in the service of defending moral realism. His basic argument is that, given the normative structural analogies between epistemology and ethics, if there are epistemic facts then there are moral facts. But there are epistemic facts and, therefore, there are moral facts (and moral realism is true). Of course, not everyone has appealed to the moral-epistemic parity in order to support normative realism. Sharon Street (2006, 2009), Allan Gibbard (1990, 2003) and Matthew Chrisman (2007), for instance, have appealed to the moral-epistemic parity in order to support forms of epistemic anti-realism. They have argued that insofar as we have some reasons to think a form of moral anti-realism (such as relativism or expressivism) is a promising approach, we have equal reason to think that an analogous form of epistemic anti-realism is a promising approach to epistemic concepts (such as knowledge, rationality and justification).

Yet other times disparity arguments are run that stress differences between the moral and the epistemic and indicate that the two domains might not be on "a par" and might not share the same metanormative

fate. Chris Heathwood (2009), for instance, has argued—against Cuneo (2007)—that although we have reason to be robust moral realists, we have reason to be less-than-fully robust realists about epistemic reasonability. That is, we can be reductionists about epistemic reasonability.[11] In the same vein, Jonas Olson (2011, 2014) has argued that we can be moral error theorists but 'immanent' epistemic realists and Hallvard Lillehammer and Niklas Möller (2015) have argued that we can be coherentists and rationally believe moral error theory. Yet others inspired by moral consequentialism have appealed to epistemic consequentialism (e.g. Goldman 1979) or inspired by virtue ethics have appealed to virtue epistemologies, responsibilist and reliabilist (e.g. Sosa 2007; Zagzebski 1996). Even Timothy Williamson's (2000) "knowledge first" project about knowledge that suggested that knowledge is *sui generis* and indefinable seems to bear some resemblance to G.E. Moore's (1903) famous contention that goodness is *sui generis* and indefinable.

The current volume's goal is dual. On the one hand, to further stimulate and advance metaepistemological research and thereby help cement the emerging sub-discipline of metaepistemology as the epistemic analogue of metaethics. On the other hand, to further stimulate and advance the fruitful dialogue between the twin normative subjects of ethics and epistemology. To this effect, the volume brings together cutting-edge work in metaepistemology by both more established philosophers and more junior philosophers.

The papers cover all, or almost all, major positions in metaepistemology, realist, anti-realist as well as constitutivist. Some of the papers seek to establish new arguments in support of realism (Kyriacou, Schmidt) while others seek to establish new arguments in support of anti-realism (Kusch and McKenna, Toppinen). Yet others seek to undermine parity arguments and present notable differences between the moral and the epistemic (Côté-Bouchard and Littlejohn, Hazlett, Flowerree). Other papers present responses to objections to various theories, such as epistemic consequentialism and epistemic reductionism (respectively, Sharadin, Heathwood), raise new objections to theories such as constitutivism and expressivism (respectively, Cowie and Greenberg, Cuneo) or explore new constitutivist positions (Coliva).

2 Summary of Papers

For ease of exposition, we have divided papers in three parts: Part I covers realism, part II covers anti-realism and part III covers constitutivism\ constructivism. Let us briefly outline the contents of the papers in each of the three parts.

Part I (Epistemic Realism) includes the papers by Cuneo, Heathwood, Kyriacou, Schmidt and Sharadin.

Terence Cuneo in "The Core Expressivist Manoeuvre" critically engages with the "core expressivist manoeuvre," a methodological strategy proposed by Carter and Chrisman (2012) in order to enable expressivism to sidestep a host of objections levelled by Cuneo (2007) (and others) against the view.[12] Cuneo contends that extant versions of expressivism are poorly situated to execute this manoeuvre and, so, expressivists are presently not well-situated to appeal to the core manoeuvre in order to avoid certain types of theoretical burdens and deflect various types of criticisms of their view.

Chris Heathwood in "Epistemic Reductionism and the Moral-Epistemic Disparity" returns to his 2009 defence of epistemic reductionism about epistemic reasonability/justification in order to respond to objections by Cuneo and Kyriacou (forthcoming) and Rowland (2013), and to further refine his original proposal. In his 2009 work, Heathwood defended the following disparity between moral and epistemic facts: whereas moral facts are irreducibly normative, epistemic facts—facts such as that someone is epistemically justified in believing something—are reducible to facts from some other domain (such as facts about probabilities). This moral-epistemic disparity is significant because it undercuts an important kind of argument for robust moral realism.

Christos Kyriacou in "From *Moral* Fixed Points to *Epistemic* Fixed Points" seeks to extend to the epistemic domain an argument by Cuneo and Shafer-Landau (2014) for moral fixed points. Cuneo and

Shafer-Landau (2014) argued that there are moral conceptual truths that are substantive in content, what they called "moral fixed points." Kyriacou argues that insofar as we have some reason to postulate *moral* fixed points, we have equal reason to postulate *epistemic* fixed points (e.g. the factivity condition). To this effect, he shows that the two basic reasons Cuneo and Shafer-Landau (2014) offer in support of moral fixed points naturally carry over to epistemic fixed points. In particular, epistemic fixed points exhibit the four "marks" of conceptual truths that they identify and can be utilised to address important challenges to epistemic realism. He concludes that insofar as we have some reason to postulate moral fixed points, we have equal reason to postulate epistemic fixed points.

Eva Schmidt in "Normative Reasons for Mentalism" proposes to fruitfully transpose reasons-talk from ethics to epistemology, in particular to Conee and Feldman's mentalism about justification. She proceeds to examine in which sense mentalist epistemic reasons-talk should be best understood. She suggests subjective normative reasons will do the trick and responds to objections. Subjective normative reasons, in spite of the qualifier "subjective," are reasons as a realist would have them and Schmidt proceeds to discuss how exactly the resulting view should be elaborated. She concludes by suggesting a way in which this view ties in with epistemological disjunctivism, so as to ensure a close connection between justification and truth.

Nathaniel Sharadin in "Epistemic Consequentialism: Haters Gonna Hate" revisits the "epistemic separatedness objection" (cf. Berker 2013) to epistemic consequentialism. The objection suggests that there can be no epistemic trade-offs among propositions and therefore epistemic consequentialism, which that implies there can be such trade-offs, is mistaken. Sharadin unpacks the metaphor of the seperatedness of propositions (or better in his view, beliefs) and argues that, as in the moral analogue of the case, there is no absolute ban of epistemic trade-offs. There can be some trade-offs in some cases. He concludes with a methodological diagnosis of what might have led to Berker's mistake.

Part II (Epistemic Anti-realism) includes the papers by Côté-Bouchard and Littlejohn. Cowie and Greenberg, Hazlett, Kusch and McKenna and Toppinen.

Charles Côté-Bouchard and Clayton Littlejohn in "Knowledge, Reasons, and Errors About Error Theory" seek to undermine the moral-epistemic parity. One widely discussed objection to the moral error theory is that it cannot be true because it leads to an epistemic error theory. They argue that this objection is mistaken. Objectors may be right that the epistemic error theory is untenable. They agree with epistemic realists that our epistemological claims are not systematically in error. However, this is not because there are robustly normative facts, but rather because the truth of our epistemic claims doesn't turn on whether there are such facts. Epistemic facts, they argue, are not robustly or categorically normative. Moral error theorists can therefore respond to the objection by holding that their view does not commit them to the epistemic error theory.

Christopher Cowie and Alexander Greenberg in "Constitutivism About Epistemic Normativity" criticise two standard arguments for constitutivism. According to constitutivists about epistemic normativity, epistemic normativity is explained by the nature of belief. Specifically, it is explained by the fact that, as a matter of conceptual necessity, belief stands in a normative relation to truth. They examine and critique two arguments for this claim. The first is based on the transparency of belief. The second is based on Moore-paradoxical sentences. They develop a common objection to both. Both arguments rely on implausible claims about the relation between accepting a norm and being motivated by it. This point has sometimes been made in connection with the transparency of belief. But its development and application to Moore-paradoxical sentences is, they suggest, novel.

Alan Hazlett in "Correctness and Goodness" is concerned with whether we have any reason to prefer true belief and assertion to false belief and assertion. One familiar account of epistemic normativity tries to explain why we have reason to prefer true belief by appealing to the premise that truth is the aim of belief and assertion. On Hazlett's reading, this boils down to the claim that beliefs and assertions are correct if and only if they

are true. While Hazlett readily accepts this claim—indeed, it is hard to see why anyone would reject it—he argues that the fact that truth is the correctness condition for belief and assertion does not give us reason to prefer true belief and assertion to false belief and assertion. This argument amplifies a line of argument in Hazlett's (2013) book *A Luxury of the Understanding: On the Value of True Belief* (Oxford University Press), which argues for a form of anti-realism about epistemic normativity by considering and rejecting a series of realist accounts of epistemic normativity.

Martin Kusch and Robin McKenna in "The Genealogy of Relativism and Absolutism" bring together two strands in (meta)epistemological theorising that have lived separate lives: epistemic relativism and the genealogical method in epistemology. Kusch and McKenna's aims are twofold. First, they apply the genealogical method to the debate around epistemic relativism by charting the development of various views that either deny or accept a form of epistemic relativism. They argue that the genealogical method helps clarify the motivations underlying these forms of (anti-)relativism. Second, they situate the genealogical method with respect to the epistemic realism/anti-realism debate: does the genealogical method implicitly takes sides in this debate, and if so, which side? While they provide a range of reasons for thinking that the "genealogist" need not take sides, they ultimately conclude that the genealogical method sits better with epistemic anti-realism.

Teemu Toppinen in "Reasons Primitivism and Epistemic Expressivism" focuses on the position of reasons primitivism advocated by Derek Parfit (2011) and Thomas Scanlon (1998), among others. According to reasons primitivism, the concept of normative reason is unanalysable. Toppinen discusses Michael Smith's (2015, 2017) recent interesting case against the primitivist view. He starts by explaining Smith's argument and proceeds to suggest that reasons primitivists can deal with this argument if they accept a suitable *expressivist* view about judgements concerning normative reasons. He suggests that the expressivist could endorse (as a first stab) the following thesis about reasons for belief: A belief that the fact that p is a reason to believe that q consists in an inferential disposition to raise one's credence in q given p. Finally, he briefly notes that Smith's argument could also be given a metaphysical twist.

Part III (Beyond the Realism/Anti-realism Divide) includes the papers by Coliva and Floweree.

Annalisa Coliva in "What Anti-realism About Hinges Could Possibly Be" addresses the question of the metaphysical status of "heavyweight" Wittgensteinian hinges like "There is an external world." Coliva argues that the standard realist answer fails because the realist view invites the sceptical paradox: If we insist on a gap between mind and world, it should be no surprise that we then struggle to bridge it. But she doesn't think the standard anti-realist answer is attractive either. Even setting aside the worry that it amounts to a form of idealism, it seems to fair no better with the sceptical paradox, which concerns our warrant for asserting that there is an external world. Coliva proposes a third alternative, which she describes as a form of quietism. This alternative has several parts, of which two are particularly important. The first is the Kantian idea that hinges are conditions of the possibility for representation rather than representations of a mind-independent reality. The second is a version of pluralism about truth, on which some propositions (such as heavyweight hinges) are true in the minimal sense that they obey certain platitudes about the linguistic roles of the truth predicate. This view avoids the problems that beset both the standard realist and anti-realist accounts. It avoids the realist's problem because it doesn't require us to bridge the gap between mind and world, and it avoids the anti-realist's problem because it doesn't equate truth with ideal assertability.

Amy Flowerree in "Epistemic Shmagency?" engages with the powerful "schmagency objection" to constitutivism, raised by David Enoch (2006, 2011). While Enoch's objection has been widely discussed in the context of practical norms, no one has yet explored how the "schmagency objection" might undermine epistemic constitutivism. In this paper, she rectifies that gap. She develops the objection against a prominent form of epistemic constitutivism, belief constitutivism. Belief constitutivism is susceptible to the schmagency objection, she argues, because it locates the source of normativity in the *belief* rather than the *agent*. In the final section, she proposes a version of epistemic constitutivism that locates epistemic normativity as constitutive of agency and proposes that this version has the resources to respond to the schmagency objection.

Notes

1. See Kyriacou (2016a) for an overview of some basic metaepistemological topics.
2. For epistemic realism, see Boghossian (2006), Cuneo (2007), Heathwood (2009), Kim (1988) and Kyriacou (2016b).
3. For analytic naturalism, see Goldman (1979, 2015) and Heathwood (2009). For synthetic, Cornell-style naturalism see Jenkins (2007). For non-naturalism (or robust realism), see Boghossian (2006).
4. See Cuneo (2007) and Kelly (2003, 2007).
5. For epistemic relativism see MacFarlane (2014) and Rorty (1979). For epistemic expressivism see Chrisman (2007), Kappel (2010) and Kyriacou (2012).
6. For instrumentalism see Kornblith (1993), Leite (2007), Cowie (2014) and Sharadin (forthcoming).
7. See Shah (2003) and Shah and Velleman (2005). For criticism, see Enoch (2006, 2011) and Côté-Bouchard (2016).
8. See Korsgaard (1996, 2008). See also Street (2006, 2009) who is a constructivist critical of Korsgaard's Kantian constructivism.
9. Although some have talked of the "demise of epistemology" (Rorty 1979) this prediction has proven mistaken. Epistemology is thriving.
10. See Cuneo (2007) for a book-length defence of the moral-epistemic parity. See also Cuneo and Kyriacou (forthcoming) and Case (forthcoming).
11. For criticism of Heathwood (2009), see Rowland (2013), Cuneo and Kyriacou (forthcoming) and Kyriacou (forthcoming, sec. 5). For Heathwood's rejoinder see his contribution to this volume.
12. For another argument against "the core expressivist manoeuvre" that draws from cognitive psychology results, see Kyriacou (2017).

References

Berker, Selim. 2013. Epistemic Teleology and the Separateness of Propositions. *Philosophical Review* 122 (3): 337–393.

Boghossian, Paul. 2006. *Fear of Knowledge: Against Relativism and Constructivism*. Oxford: Oxford University Press.

Carter, J. Adam, and Matthew Chrisman. 2012. Is Epistemic Expressivism Incompatible with Inquiry? *Philosophical Studies* 159 (3): 323–339.

Case, Spencer. forthcoming. From Epistemic to Moral Realism. *Journal of Moral Philosophy*.

Chrisman, Matthew. 2007. From Epistemic Contextualism to Epistemic Expressivism. *Philosophical Studies* 135 (2): 225–254.

Côté-Bouchard, Charles. 2016. Can the Aim of Belief Ground Epistemic Normativity? *Philosophical Studies* 173 (12): 3181–3198.

Cowie, Christopher. 2014. In Defence of Instrumentalism About Epistemic Normativity. *Synthese* 191 (16): 4003–4017.

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

Cuneo, Terence, and Christos Kyriacou. forthcoming. Defending the Moral/Epistemic Parity. In *Metaepistemology*, ed. Conor McHugh, Jonathan Way, and Daniel Whiting. Oxford: Oxford University Press.

Cuneo, Terence, and Russ Shafer-Landau. 2014. The Moral Fixed Points: New Directions for Moral Nonnaturalism. *Philosophical Studies* 171 (3): 399–443.

Enoch, David. 2006. Agency, Shmagency: Why Normativity Won't Come from What is Constitutive of Action. *Philosophical Review* 115 (2): 169–198.

———. 2011. Shmagency Revisited. In *New Waves in Metaethics*, ed. Michael Brady, 208–233. Basingstoke: Palgrave Macmillan.

Gibbard, Allan. 1990. *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press.

———. 2003. *Thinking How to Live*. Cambridge, MA: Harvard University Press.

Goldman, Alvin. 1979. What is Justified Belief? In *Justification and Knowledge*, ed. George Pappas, 1–25. Dordrecht: Reidel.

———. 2015. Reliabilism, Veritism, and Epistemic Consequentialism. *Episteme* 12 (2): 131–143.

Hazlett, Allan. 2013. *A Luxury of the Understanding: On the Value of True Belief*. Oxford: Oxford University Press.

Heathwood, Chris. 2009. Moral and Epistemic Open-Question Arguments. *Philosophical Books* 50 (2): 83–98.

Jenkins, C.S. 2007. Epistemic Norms and Natural Facts. *American Philosophical Quarterly* 44 (3): 259–272.

Kappel, Klemens. 2010. Expressivism About Knowledge and the Value of Knowledge. *Acta Analytica* 25 (2): 175–194.

Kelly, Thomas. 2003. Epistemic Rationality as Instrumental Rationality: A Critique. *Philosophy and Phenomenological Research* 66 (3): 612–640.

———. 2007. Evidence and Normativity: Reply to Leite. *Philosophy and Phenomenological Research* 75 (2): 465–474.

Kim, Jaegwon. 1988. What is 'Naturalized Epistemology?'. *Philosophical Perspectives* 2: 381–405.

Kornblith, Hilary. 1993. Epistemic Normativity. *Synthese* 94 (3): 357–376.

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

———. 2008. *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*. Oxford: Oxford University Press.

Kyriacou, Christos. 2012. Habits-Expressivism About Epistemic Justification. *Philosophical Papers* 41 (2): 209–237.

———. 2016a. Metaepistemology. Edited by James Fieser and Bradley Dowden. *Internet Encyclopedia of Philosophy*. http://www.iep.utm.edu/meta-epi/.

———. 2016b. Ought to Believe, Evidential Understanding and the Pursuit of Wisdom. In *Epistemic Reasons, Norms and Goals*, ed. Martin Grajner and Pedro Schmechtig, 383–406. Berlin and Boston: DeGruyter.

———. 2017. Expressivism, Question Substitution and Evolutionary Debunking. *Philosophical Psychology* 30 (8): 1019–1042.

———. forthcoming. Evolutionary Debunking: The Milvian Bridge Destabilized. *Synthese*.

Leite, Adam. 2007. Epistemic Instrumentalism and Reasons for Belief: A Reply to Tom Kelly. *Philosophy and Phenomenological Research* 75 (2): 456–464.

Lillehammer, Hallvard, and Niklas Möller. 2015. We Can Believe the Error Theory. *Ethical Theory and Moral Practice* 18 (3): 453–459.

MacFarlane, John. 2014. *Assessment Sensitivity: Relative Truth and Its Applications*. Oxford: Oxford University Press.

Moore, G.E. 1903. *Principia Ethica*. Cambridge: Cambridge University Press.

Olson, Jonas. 2011. Error Theory and Reasons for Belief. In *Reasons for Belief*, ed. Andrew Reisner and Asbjørn Steglich-Petersen, 75–93. Cambridge: Cambridge University Press.

———. 2014. *Moral Error Theory: History, Critique, Defence*. Oxford: Oxford University Press.

Parfit, Derek. 2011. *On What Matters*. Oxford: Oxford University Press.

Rorty, Richard. 1979. *Philosophy and the Mirror of Nature*. Princeton: Princeton University Press.

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

Scanlon, Thomas M. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.

Shah, Nishi. 2003. How Truth Governs Belief. *Philosophical Review* 112 (4): 447–482.

Shah, Nishi, and David Velleman. 2005. Doxastic Deliberation. *Philosophical Review* 114 (4): 497–534.

Sharadin, Nathaniel. forthcoming. Epistemic Instrumentalism and the Reason to Believe in Accord with the Evidence. *Synthese*.

Smith, Michael. 2015. Evaluative Judgements, Judgements About Reasons, and Motivations. In *Motivational Internalism*, ed. Gunnar Björnsson, Caj Strandberg, Ragnar Francén Olinder, John Eriksson, and Frederik Björklund, 27–43. New York: Oxford University Press.

———. 2017. Parfit's Mistaken Metaethics. In *Does Anything Really Matter? Essays on Parfit and Objectivity*, ed. Peter Singer, 99–120. Oxford: Oxford University Press.

Sosa, Ernest. 2007. *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Volume I*. Oxford: Oxford University Press.

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

———. 2009. Evolution and the Normativity of Epistemic Reasons. *Canadian Journal of Philosophy* 35 (s1): 213–248.

Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press.

Zagzebski, Linda. 1996. *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge: Cambridge University Press.

Part I

Epistemic Realism



The Core Expressivist Manoeuvre

Terence Cuneo

A colleague of mine once described the experience of working through a prominent text defending expressivism as similar to that of "watching a shell game." Although he didn't elaborate on the comparison, I think I see what he meant. The experience was like that of trying to follow a series of clever and elaborate moves in which one catches only glimpses of what one is interested in seeing. I realise that the comparison is not flattering. But I suspect it voices an experience that some (and perhaps many) of us have had when engaging with the expressivist tradition: we find that the questions that we want to pursue about the existence and nature of normative reality deliberately hidden from view and exchanged for a series of different questions about the character of normative thought and discourse, which are themselves often pursued in a series of highly sophisticated manoeuvres.

Others have offered more optimistic appraisals of expressivism's methods. In their essay "Is Epistemic Expressivism Incompatible with

T. Cuneo (✉)
Department of Philosophy, University of Vermont, Burlington, VT, USA
e-mail: tcuneo@uvm.edu

© The Author(s) 2018
C. Kyriacou, R. McKenna (eds.), *Metaepistemology*, Palgrave Innovations in Philosophy, https://doi.org/10.1007/978-3-319-93369-6_2

Inquiry?," J. Adam Carter and Matthew Chrisman, for example, contend that expressivism's evasive character is not due to a lack of theoretical candour (Carter and Chrisman 2012) . Rather, it has both a principled rationale and a function whose aims include explaining why the view is immune to certain types of objections.

The rationale is what Carter and Chrisman call the "core expressivist manoeuvre." The core manoeuvre has three steps when applied to the normative domain. The first step is to switch the subject: rather than concern itself with traditional metanormative topics such as what normative properties or facts are (or would be), expressivism concerns itself with what it is to express a normative judgement or engage in normative discourse (of a given kind). The second step is to offer a distinctive, expressivist account of what normative thought and discourse are. At a first approximation, this account rejects the claim that normative thought and discourse have normative representational content, maintaining instead that they express attitudes of commendation and condemnation. The third step consists in expressivists concluding—after having defended their views about the character of normative thought and discourse—that they (in their role as theorists) can "just stop talking" about normative reality and, instead, issue first-order normative judgements.[1] In short, were it successfully executed, the core manoeuvre would enable expressivists to bracket prominent metanormative issues, such as those that concern the existence and nature of values, and focus on the topics that most interest them, namely, the character of normative thought and discourse. That is its function.

Let me add that, as it is presented, the core expressivist manoeuvre is not just one dialectical strategy among others. Carter and Chrisman present it as expressivism's *signature* dialectical strategy. Other philosophers such as Mark Schroeder appear to agree.[2] Were we to fail to appreciate the core manoeuvre's role and significance within the expressivist project, our grasp of this project would be seriously incomplete.[3]

In this essay, I propose to take a closer look at the core expressivist manoeuvre. The primary argument I am going to develop maintains that in order to successfully execute the core manoeuvre, expressivism must jointly satisfy two desiderata. The first is that it provides an account of normative thought and discourse that has sufficient expressive power.

Such an account must be capable of explaining the workings not only of some subset of normative thought and discourse (such as evaluations) but the full range thereof (and explain it in a unified way). I'll refer to this desideratum as *Expressive Power*. The second desideratum is that it provides an account of normative thought and discourse according to which they lack normative representational content. (I'll say more about this concept in Sect. 3. For present purposes, think of it as content that concerns or is about normative reality.) After all, if such thought and discourse were to have such content, it would make sense to inquire whether there are normative properties or facts that they represent, and what they are like. I will refer to this desideratum as *No Normative Representational Content* (or "No Normative Content," for short).

I shall assume that every satisfactory expressivist position must satisfy Expressive Power, given their aim of providing an account of ordinary normative thought and discourse. In contrast, I will not assume that every expressivist position must satisfy No Normative Content. It is a desideratum that must be fulfilled only by expressivist positions with certain aims, such as those that endeavor to execute the core manoeuvre. The central claim for which I will argue is that it is difficult to jointly satisfy these desiderata, as they pull in opposite directions. It might be that some expressivist views satisfy Expressive Power. But I'll argue that they are thereby poorly situated to satisfy No Normative Content. And it might be that some expressivist views satisfy No Normative Content. But, I'll contend, they are thereby not well-placed to satisfy Expressive Power. I will refer to this dynamic as the *central tension*.

The central tension bears upon the core manoeuvre in the following way. On the one hand, an expressivist view that satisfies Expressive Power will be well-positioned to execute the first two steps of the core manoeuvre. But such a view, I'll contend, will be poorly suited to move to the third step wherein expressivists can legitimately "stop talking" about normative reality. On the other hand, an expressivist view that satisfies No Normative Content will not be able to execute the core manoeuvre's second step, which is that of furnishing a satisfactory account of normative thought and discourse. And, so, such a view will also not be able to execute the core manoeuvre. The overarching conclusion at which I'll arrive is that expressivists are presently not well-situated to appeal to the core

manoeuvre in order to avoid certain types of theoretical burdens and deflect various types of criticisms of their view.

Let me emphasise two points at the outset. There is no agreed upon understanding of what expressivism is. In what follows, I will be working with what I take to be a widely accepted understanding of the view, although I am confident that some will reject it. Second, I will not be arguing that expressivism as such cannot execute the core manoeuvre. The history of expressivism is, after all, that of a position which has proven to be extraordinarily resourceful when responding to challenges. I expect that it will be equally resourceful in this case too. Nonetheless, I do contend that the most prominent versions of the view are not well-positioned to execute the manoeuvre. The conclusion I reach is that being positioned to execute the manoeuvre will require supplementing and perhaps developing expressivism in ways that involve modifying some of its central ambitions and commitments.

1 Section 1

A more precise account of the core expressivist manoeuvre requires some stage-setting. I'll begin by regimenting some terminology.

In what follows, I'll use the phrase "normative domains" to refer to domains such as ethics, epistemology, practical reason, and aesthetics. These domains are normative (in part) because they include thought and discourse that is evaluative ("x is good"), directive ("one ought to x," "one is required to x"), reason-expressing ("x is a reason for acting"), and fittingness-expressing ("x is apt or appropriate").[4] I'll use the term "metanormative inquiry" to concern inquiry into the character of what belongs to these domains—where that includes not only normative thought and discourse but also (perhaps) normative properties and facts.

We can distinguish two broad projects within metanormative inquiry, one concerning mind, the other concerning world. The first project concerns itself only with the nature of normative thought and discourse. This project asks questions such as: Are normative thought and discourse such as to represent normative reality? Or do they play some other role? Call this the *mental/linguistic project*. The second project concerns itself with

the character of normative properties or facts. This project concerns itself primarily with questions such as: Are there normative properties and facts of a given kind? And, if so, what are they like? Call this the *metaphysical project*.[5] According to the core expressivist manoeuvre, if we start with and ask the right questions when engaging in the mental/linguistic project, this would thereby "undermine" the metaphysical project (Carter and Chrisman 2012, 334).

When Carter and Chrisman maintain that executing the core expressivist manoeuvre undermines (what I am calling) the metaphysical project, I take it that they have the following in mind. Successfully executing the first two stages of the manoeuvre would imply that there is insufficient reason to pursue the types of questions that animate the metaphysical project. More specifically, I take the leading thought to be this: if engaging in the mental/linguistic project reveals that normative thought and discourse are not in the business of representing normative reality, then there would be insufficient reason to pursue the metaphysical project, asking whether normative reality exists and what it is like (and what relations we might bear to it). The idea is not that this implies that normative properties and facts do not exist. Carter and Chrisman are clear that, so far as epistemic facts go—this being the normative domain on which they focus—the epistemic expressivist can remain entirely neutral about whether such facts exist.[6] Nor is the idea that engaging in the metaphysical project is somehow "off limits" to theorists full-stop. After all, one might have reason to engage in the project because one is curious about, say, issues in moral ontology or finds thinking about them enriching. Rather, the idea is that, given the proper aim of metanormative inquiry, which consists in accurately characterising the normative domains, there is insufficient reason to engage in the metaphysical project.[7]

An analogy might help to illuminate the issues. Suppose one held the strange view that in thanking others, agents thereby represent "to-be-thanked-facts" (or "ought-to-be-thanked" facts). Under this unusual view, when you sincerely say "Thank you" to me, you thereby represent the fact *that I am to be thanked*. Such a view would make sense of engaging in the metaphysical project in which we endeavour to ascertain whether such facts exist and what they are like. But suppose you held the

sensible view that thanking itself does not represent to-be-thanked facts. Rather, it plays the broadly expressive role of expressing gratitude. If you held such a sensible view, then there would be little reason to hold that you ought to engage in the metaphysical project with regard to putative to-be-thanked-facts. Once again, you needn't deny that such facts exist. You might even want to think about what they would be like were they to exist. But they wouldn't be a proper object of such inquiry. Any position that took another to task for offering an unsatisfactory account of to-be-thanked facts would be off-target in that respect.

Let us suppose for the moment that the core manoeuvre captures how expressivists approach metanormative inquiry. Bringing this to light would have numerous benefits. It would help us to have a more adequate understanding of the expressivist project, and why its proponents make the types of methodological moves they do, directing their attention nearly exclusively to the mental/linguistic project. In addition, it would help us to see why some take expressivism to be highly attractive: it avoids altogether having to discharge certain kinds of theoretical burdens. These burdens, which normative realists must shoulder, include developing an account of normative reality and how we could grasp and refer to it. If the core manoeuvre were to undercut the metaphysical project, however, then there would be no theoretical pressure to develop an account of normative reality and how we could grasp and refer to it. That would be highly desirable, given how challenging it is to formulate satisfactory accounts regarding these issues.[8] Finally, having a clear understanding of the manoeuvre would help us to understand why expressivists hold that their view can entirely avoid certain types of criticisms. For example, critics such as Derek Parfit have accused expressivism of being close to a version of normative nihilism.[9] And a variety of theorists—myself included—have argued that epistemic expressivism is untenable because (in part) it commits itself to a range of implausible claims, some of which include the character of epistemic reality.[10] Successfully executing the core manoeuvre would enable expressivists to reply that their view is not vulnerable to many (and perhaps all) of these concerns, since it would imply that such facts are as theoretically extraneous as to-be-thanked facts.

The next two sections explore the issue of whether expressivism is well-placed to execute the core manoeuvre. I want to make two points before entering into that discussion.

The first is that, while one can find passages in which expressivists appear to endorse something like the core manoeuvre, it is by no means evident that they intentionally aim to engage in it when theorising.[11] There are, after all, a wide variety of passages in which expressivists such as Blackburn and Gibbard agree that there are moral facts and say informative things about what they take them to be like. Blackburn maintains that they are "shadows" of normative thought and discourse that do not "answer" to our normative thoughts (1993, 181, 206, 1999, 216). In some places, Gibbard suggests that they are not states of affairs that obtain but true thoughts; in others, he suggest that they are ordinary natural facts (2003, 182).[12] Since it can be very difficult to interpret these passages and harmonise them with the various things that expressivists say, it is difficult to know how much interpretive weight to place on them. Still, they are puzzling under an interpretation according to which expressivists take themselves to be engaging in the core manoeuvre, since its aim is to enable expressivists to avoid engaging in the metaphysical project altogether. However that may be, I shall assume that the core manoeuvre is of philosophical interest whether or not expressivists intentionally aim to engage in it. After all, it might be a strategy that they *should* try to implement (even if they fail to), given their fundamental aims.

The second point concerns what it would take for the core manoeuvre to succeed. Legitimately concluding that theorists can just "stop talking" about normative reality requires that there be insufficient reason to engage in the metaphysical project. It is worth asking why one would think that establishing that normative thought and discourse lack normative representational content would imply that there is insufficient reason to engage in the metaphysical project, as the core manoeuvre maintains. The thought, presumably, hinges on something like this assumption:

> (A) Establishing that normative thought and discourse lack normative representational content would thereby be to provide sufficient reason to hold that normative properties and facts would play no significant theoretical role.

(A) might be true because the only theoretical role normative properties and facts could play that would justify engaging in the metaphysical project is that of being the objects of the representational content of normative thought and discourse. Alternatively, (A) might be true because, while normative properties and facts might (in theory) play numerous such roles, establishing that they are not the objects of normative thought and discourse would undermine any reason to believe that they play these other roles.

(A) is highly contentious. A helpful way to see why is to return to the case of thanking. Our earlier discussion correctly assumed that, in the ordinary case, speech acts in which agents thank others are not a species of assertion (or anything assertion-like). Rather, thanking is a speech act that consists in agents committing themselves to being in the mental state of being grateful (or, perhaps, feeling gratitude). Thanking belongs to the category of speech acts that William Alston calls *expressives*.[13]

Ordinary normative discourse is, according to expressivism, more like thanking than asserting. It is more like thanking because such discourse also consists in agents committing themselves to being in mental states of certain kinds, such as commending or condemning acts of various types. On this occasion, I am not interested in exploring the issue of whether expressivism is correct about this. Instead, I wish to make the point that, even if it is, this would not license concluding that there is insufficient reason to engage in the metaphysical project.

The reason is that, like all speech act types, expressives have *aptness conditions* for their performance. Imagine that (under ordinary conditions) I were to utter the sentence "Thank you for this gift!" thereby thanking you. My speech act can be more or less apt depending on how things are in the world. Suppose, for example, the gift is intended for somebody else. Then I am liable to correction, and somebody in the know should inform me that the gift is not for me. Or suppose the wrapped box is not a gift at all but merely a prop in a theatre performance. Then it would be mistaken of me to thank anyone for it. Or suppose that the gift is a box of chocolates that you have laced with cyanide. Then it would be inapt of me to thank you for it, since it would make no sense for me (under ordinary conditions) to thank you for intending to poison me. Or suppose I know that you acquired the

gift wrongfully, perhaps stealing it from someone else. In this case, I might thank you for it, but I should qualify what I say, perhaps expressing the idea that I cannot accept such a gift however well intended you might be.

What I am calling "aptness conditions," then, cover a fairly wide range of ways in which the performance of a speech act can be fitting or unfitting. What our examples illustrate is that, by all appearances, they involve agents bearing normative relations of various sorts against one another. If I am mistaken about who has given me the gift, or whether the wrapped box is a gift, for example, then you are entitled to correct me.

It is a good question what the character of these normative relations is (whether for example, they include moral and epistemic relations).[14] Rather than explore this issue on this occasion, I'll note this: any satisfactory account of speech must take into account the aptness conditions for its performance. Moreover, that there are such conditions for the performance of expressives is compatible with expressivism offering us a more or less correct account of normative thought and discourse. If that is so, normative reality might play interesting explanatory roles, such as constituting the aptness conditions for speech, even if the aim of such thought and discourse is not to represent normative reality.[15] It follows that undercutting the metaphysical project is not as straightforward as successfully defending an expressivist account of normative thought and discourse. To accomplish that, expressivists would also have to provide a satisfactory defense of (A). I am, however, aware of no passages that attempt such a defense. While this strikes me as a gap in their project, nothing in the argument that follows will hinge on it.

2 Section 2

I turn now to the primary argument regarding the core manoeuvre that I want to develop. This argument, recall, consists in identifying two desiderata that an expressivist position that engages in the core manoeuvre must satisfy. On the one hand, it must satisfy Expressive Power, presenting an account of normative thought and discourse that has sufficient expressive power. On the other, it must satisfy No Normative Content,

implying that such thought and discourse lack normative representational content. The argument maintains that satisfying these desiderata pulls expressivism in opposite directions. That is what I've called the *central tension*. Its bearing on the core manoeuvre, recall, is this. If expressivism satisfies Expressive Power, then it is not positioned to undercut the metaphysical project. But if it satisfies No Normative Content, then it is poorly positioned to move beyond the second step of the core manoeuvre. Either way—or so runs the argument—the prospects for executing the core manoeuvre are not promising.

In this section, I'll develop the first half of the argument. I'll do so by presenting a version of expressivism that promises to satisfy Expressive Power. I'll then ask whether this view is positioned to execute the core manoeuvre.

By far and away, the version of expressivism best positioned to satisfy Expressive Power is what I'll call *standards expressivism*. This position has two primary commitments. First, it identifies a mental state of a certain kind that is the attitudinal component of normative judgement. This attitude is helpfully thought of as being one of endorsement or acceptance. But it is not normative belief, since it is not directed toward predicative normative representational contents. Second, standards expressivism identifies the contents toward which these attitudinal states are directed. As its name indicates, the view is naturally understood to maintain that these are standards or "norms." These standards or norms, in turn, are characterised as directives or imperatives. Or better yet, they are characterized as *imperatival contents*, which have the form <Do ϕ!> or <Do not ϕ!>, where "ϕ" stands for some action.[16] In its paradigm form, then, standards expressivism tells us that normative judgements consist in an agent taking some non-doxastic attitude, such as accepting or endorsing, toward an imperatival content.

Perhaps the best-known version of standards expressivism is the "norm expressivist" view developed in Allan Gibbard's *Wise Choices, Apt Feelings* (Gibbard 1990).[17] When developing his view, Gibbard observes that normative judgements are best thought of as being directed at not individual standards or norms but "systems of norms." (1990, 86–90) Gibbard continues by claiming that we think of a

> system of norms as a system of permissions and requirements, applicable to a wide range of actual and hypothetical circumstances …We can characterize any system *N* of norms by a family of basic predicates '*N*-forbidden', '*N*-optional', and '*N*-required'… These predicates are descriptive rather than normative: whether a thing, say, is N-permitted will be a matter of fact. It might be N-permitted without being rational, for the system N might have little recommend it. (1990, 87)

I'll return later to Gibbard's remark that these deontic predicates are "descriptive" rather than normative. For now, let me draw attention to how Gibbard presents his view.

In the passage just quoted, Gibbard maintains that systems of norms are best understood to incorporate deontic modals. They are, as I'll say, *deontically laden*.[18] Although Gibbard himself does not draw attention to the issue, there is good reason to characterise imperatival contents in this way. After all, an unqualified imperatival content such as <Do ϕ!> is multiply ambiguous. It might tell us that ϕing is required. Or it might tell us that ϕing is merely advisable. We need a way to resolve this ambiguity, specifying its exact deontic status. The most obvious way—the way that Gibbard appears to endorse—is to understand such norms as being deontically laden, including or being qualified by deontic modals, such as being permissible, advisable, required, forbidden, or inadvisable. To which I'll add that qualifying imperatival contents in this way is what (in part) positions standards expressivism to satisfy Expressive Power. By incorporating such modals into its account of the contents of normative thought and discourse, the view is capable of providing a sufficiently fine-grained and nuanced characterisation of such thought and discourse.

Every presentation of expressivism of which I'm aware characterises the view as one according to which normative thought is not descriptive. Jamie Dreier, for example, writes that "Expressivists are distinguished by their claim that there is nothing to making a normative judgement over and above being in a state that plays a certain 'non-cognitive' psychological role, a role more like desire than it is like factual belief." (2004, 39) Michael Ridge maintains that the view offers "a distinctively nonrepresentational theory of the states of mind for the target discourse." (2015, 7)[19] What is left implicit in these passages is a qualification. It is not that

normative judgements fail to have representational content full stop, failing to represent such things as actions and persons. It is rather that they fail to have *normative* representational content, failing to represent normative properties of such actions and persons.

There is, however, a distinction to mark between different kinds of normative representational content that is important for our purposes. The distinction is that between *predicative* and *non-predicative* normative representational content. (Recall that, according to our earlier gloss, such content represents or is about normative properties or facts.[20]) To appreciate the distinction, consider some examples.

Imagine that you form the intention to investigate whether the UN's enacting economic sanctions would be epistemically irresponsible. Under a natural interpretation, your intention has non-predicative normative representational content. It has normative representational content since it is about the normative property *being epistemically irresponsible*. But it is not predicative, since it does not predicate epistemic irresponsibility of the sanctions (or, alternatively, it does not represent epistemic irresponsibility as instantiated by these sanctions).

Imagine, further, that upon learning of your investigation, a colleague were to issue a warning: "Do not wrongfully obtain any evidence!" This speech act also has normative representational content, since it concerns the property *being wrongful*. But it also lacks predicative normative representational content, since it does not predicate of your actions the property *being wrongful* (or, alternatively, it does not represent *being wrongful* as instantiated by these actions).

Finally, imagine that you were to report the findings of your investigation by sincerely uttering the sentence: "The UN's enacting economic sanctions is epistemically irresponsible." Under a natural (broadly cognitivist) interpretation, your speech act has normative representational content, since it concerns the epistemic property *being epistemically irresponsible*. But in this case the normative representational content is predicative, since it predicates *being epistemically irresponsible* of the imposition of the sanctions (or, alternatively, it represents *being epistemically irresponsible* as instantiated by the imposition of the sanctions).

Here is the reason why marking this distinction is important. Standards expressivism denies that normative thought and discourse have predicative

normative representational content. Arguably, moreover, it provides a characterisation of such thought and discourse that secures this result, holding that it consists in the endorsement of imperatival contents. This characterisation, however, is compatible with normative thought and discourse having *non-predicative* normative representational content. Indeed, under a very natural reading of standards expressivism, according to which imperatival contents are deontically laden, the view implies that normative judgements have normative representational content. These judgements, after all, concern imperatival contents that are deontically laden, which include not only what is permissible, required, or forbidden, but also what one ought to do on the whole. But if standards expressivism is compatible with or implies:

> Normative thought and discourse have normative representational content,
>
> then it is compatible with or implies that such thought and discourse represent normative properties. But if such thought and discourse is compatible with or implies this, then standards expressivism fails to satisfy No Normative Content, since this claim maintains that such thought and discourse lack normative representational content. Far from undercutting the metaphysical project, then, standards expressivism appears to provide either no reason to abstain from it or a clear rationale for engaging in it.[21]

It is worth briefly expanding upon the main claim here. Our earlier discussion indicated that non-assertoric speech acts such as thanking have aptness conditions. Some speech act types, such as imperatives, are such that their aptness conditions include *correctness* conditions, which are (*inter alia*) important for making sense of logically valid inferences that involve their contents.[22] While most agree that the correctness conditions for imperatives are not truth-conditions, there is disagreement about what such correctness conditions consist in. In his discussion of these issues, Chrisman maintains that correctness conditions for imperatival contents evaluates them as "legitimate *relative to a norm*," which is "an operation on a set of possible circumstances an agent might face telling us of at each circumstance what is required/recommended, forbidden, and optional." (2016, 144)[23] Chrisman's view regarding the

correctness conditions for imperatives mirrors Gibbard's view regarding their content. This makes sense given the plausible assumption that there must be an isomorphism between contents and their correctness conditions. But the point I wish to emphasise is that, although standards expressivists such as Gibbard say nothing regarding the correctness conditions for imperatival contents, any view that appeals to such contents must address this issue. And when it does, the questions that animate the metaphysical project wait to be answered.

Earlier I said that my strategy in this section would be two-fold. It would involve, first, presenting a version of expressivism that promises to satisfy Expressive Power and, second, asking whether the view is positioned to execute the core manoeuvre. I have chosen to focus on standards expressivism not only because it is probably the most heavily worked out extant version of expressivism, but also because it is a position that explicitly endeavours to satisfy Expressive Power. In this respect, it stands apart from many and perhaps all other expressivist positions. But, if what I've argued is on the mark, it is precisely because standards expressivism is positioned to satisfy Expressive Power that it is thereby poorly suited to execute the core manoeuvre, as it fails to satisfy No Normative Content. This conclusion is, I believe, both surprising and instructive.

It is surprising because expressivists present their views as being sharply different from broadly "factualist" positions such as realism and constructivism. There are indeed important differences. Typically, for example, factualist positions do not identify normative representational content with deontically laden imperatival content.[24] And, typically, factualists have not concerned themselves with the issue of whether normative judgements involve an element of endorsement. Still, if our discussion has been on the mark, the differences between these views appear to be not nearly as pronounced as we might have assumed. Under a natural reading, *both* types of view hold that deontic normative judgements are directed at deontically laden representational contents regarding actions that are permissible, obligatory, and forbidden. As such, under a natural reading, both types of view appear as if they must address questions internal to the metaphysical project, some of which concern whether the

contents of these judgements satisfy correctness conditions, and what the character of these conditions is.

The conclusion is instructive because, once we get clear on what standards expressivism commits itself to, it is worth asking whether the view is a version of expressivism at all. When introducing his position in *Thinking How to Live*, Gibbard writes that "from a basis that excludes normative facts and treats humanity as part of the natural world, I explain why we would have normative concepts that act much as normative realists proclaim." (2003, xii) Ridge maintains that the

> expressivist gambit is to reject the cognitivist's representationalist order of explanation. On the expressivist approach, we do not start with normative states of the world and explain normative judgements as cognizing them. Rather, we start with normative judgement, and its distinctively practical functional role. Normative judgements function to settle the "thing to do"…. In quasi-realist fashion, we then "earn the right" to the realist-sounding things that ordinary folks say by explaining how this expressivist picture is compatible with literal talk of normative truth, the logical validity of normative inferences, and the like. (2014, 101–02)

Ignore for the moment that Gibbard is probably not best read as claiming that the basis of his view excludes normative facts, since he says elsewhere that his view is noncommittal on this issue (2003, x, 88). And bracket Ridge's claim that realists typically explain how normative concepts work by reference to normative facts, since they often do not.[25] Furthermore, place to the side Ridge's claim that cognitivism starts with normative states of the world and endeavors to explain how we represent them, since that, too, is a highly controversial characterisation of cognitivism. (It would imply either that error-theories are not cognitivists or employ a fundamentally different approach to the mental/linguistic project than their realist and constructivist rivals.) Instead, note that realists typically explain how normative concepts work, or what normative terms mean, by reference to their role in representing normative reality (*if* any such reality there be). Looked at this way, there are differences between standards expressivism and its rivals. But, again, these differences do not

appear to be deep: by all appearances, both types of view appeal to normative representational content in order to explain (at least in part) the character of normative judgements. Whatever deep differences there are between these views, they do not appear to lie here.

3 Section 3

The first half of what I've called the central tension states that any view that satisfies Expressive Power is poorly situated to satisfy No Normative Content and, so, execute the core manoeuvre. While I have employed standards expressivism to illustrate the point, the conclusion is not one that concerns standards expressivism alone. For, arguably, any position that satisfies Expressive Power will have to avail itself of deontically laden content and, so, offer a characterisation of the correctness conditions thereof. These commitments, I claim, make undercutting the metaphysical project difficult.

In this section, I turn to the second half of the central tension. Recall that this half of the tension states that, when views satisfy No Normative Content, they are not well-placed to satisfy Expressive Power. And, so, these views are also not positioned to execute the core manoeuvre. My strategy, again, is to look at extant expressivist views—in this case, those that appear to satisfy No Normative Content.

Begin with the position developed in Gibbard's *Thinking How to Live* (Gibbard 2003). This position offers an approach to normative thought and discourse that emphasises not the acceptance of deontically laden imperatival contents, but being in "planning states." When simplified, the leading idea is that to think that one ought on the whole to act in some way is to decide to act in that way (2003, 17).[26] However, in his development of this view, Gibbard almost entirely limits his attention to thought and discourse concerning the so-called overall ought. As such, Gibbard's proposal concerns only a small fragment of normative thought and discourse. (Indeed, Gibbard is explicit that his aim is not to offer an explanation of how actual, ordinary normative thought and discourse operate but an idealised model that illustrates how such thought and discourse might work. See Gibbard 2003, introduction.) There are

passages in which Gibbard seems to suggest that his plan-laden approach is the key to understanding normative concepts as such (2003, 21). And there is a section in which Gibbard attempts to extend the approach to other ostensibly normative concepts, such as KNOWLEDGE. But Gibbard does not indicate whether he believes the approach can be extended to the whole range of normative thought and discourse. And it is not easy to see how it could be. After all, to judge that it is (morally) good that an event happened several thousand years ago, that an agent has a (non-decisive) epistemic reason to believe a proposition, or that it is (aesthetically) fitting to feel joy when listening to excellent music needn't involve being in a planning state, let alone deciding to enact a plan. For a planning state approach to have a chance at satisfying Expressive Power, it would need to be supplemented considerably, perhaps in ways the incorporate the commitments of standards expressivism.[27]

Something similar is true, I believe of Simon Blackburn's views. While having witnessed various permutations and being wide-ranging in a variety of respects, Blackburn's views do not represent an attempt to develop a version of expressivism that concerns the whole range of normative thought and discourse. Rather, their primary aim has been to provide expressivist models of normative thought and discourse that have the right sort of *structure* to address the Frege-Geach problem.

These models make no claim to capturing the breadth and nuances of ordinary normative thought and discourse. In fact, they tend to be extremely crude. In his early *Spreading the Word*, Blackburn works with a "Boo! Hurrah!" version of a natural language "containing no evaluative predicates." (1980, 193) In his later "Attitudes and Contents" (chap. 10 of Blackburn 1993), Blackburn combines this toy "Boo! Hurrah!" language with some insights from deontic logic. This new language does not concern the "more usual obligations and permissions" that deontic logic typically does. But, Blackburn writes, it has the advantage of providing a "logical apparatus" that "should apply wherever we have the idea of a goal or aim and the corresponding idea of something to be avoided or not to be avoided." (1993, 189) The move is striking when compared to Gibbard's early position. In *Wise Choices, Apt Feelings* (1990), Gibbard takes directives to concern deontically laden imperatival contents precisely *because* doing so enables him to develop a model of normative

thought and discourse with expressive power. Yet Blackburn writes as if it is an advantage of the modified "Boo! Hurrah!" language that it does not take such things as their content, even though it thereby sacrifices the expressive power that a position such as (early) Gibbard's enjoys. However that may be, the point is that Blackburn's views might satisfy No Normative Content. But they have not endeavoured to satisfy Expressive Power. And without a great deal of supplementation, which may involve introducing deontically laden imperatival contents, it is not apparent how they could. They are not well-suited to execute the core manoeuvre.

Perhaps the most promising candidate for an expressivist position that finds itself positioned to satisfy both Expressive Power and No Normative Content is the view developed in Mark Schroeder's *Being For* (2008a). Those familiar with Schroeder's discussion know that its primary conceptual move is to reject the assumption—endorsed by the expressivist views with which he engages—that "to each normative predicate there corresponds a distinct attitude." (2008a, 57) That assumption, according to Schroeder, makes it impossible for the view to address Frege-Geach. Instead, Schroeder proposes that expressivists locate a general noncognitive attitude that corresponds to normative predicates. Taking a cue from *Wise Choices, Apt Feelings* (Gibbard 1990), Schroeder's proposes the attitude *being for*. Under this proposal, "wrong" corresponds to *being for blaming for*. To judge that murder is wrong, according to this position, is *to be for blaming for murder* (2008a, 58).

There is no discussion of expressivism that is more sophisticated than Schroeder's. Still, by its own admission, the view Schroeder develops is not designed to satisfy Expressive Power (even if it might find itself positioned to do so). In fact, at one point Schroeder acknowledges that *being for blaming* is the wrong sort of analysis of "wrong," but replies that nothing "turns on which analysis we choose." (2008a, 58 n3). Given certain theoretical aims, this seems right; the first step that an expressivist view should take is probably to locate a view with the right type of structure to address Frege-Geach. But given other theoretical aims, such as that of executing the core manoeuvre, this answer is unsatisfactory. For a view that satisfactorily addresses Frege-Geach might not satisfy either Expressive Power or No Normative Content.

In fact, there is reason to believe that Schroeder's proposal does not. For one thing, *being for blaming for* seems not to be a necessary condition or component of judging that an act is wrong: I may judge that an action is wrong and not be for blaming the agent who performed it because blaming would exacerbate issues, or the wrong isn't sufficiently weighty to warrant blame, or that there are conditions that excuse the agent from blame. Most consequentialist positions, for example, would agree that to judge that an act is wrong because it fails to maximise value is not perforce to be for blaming anything. For another, *being for blaming* is not a sufficient condition or component of judging that an act is wrong: I could be for blaming violations of the law, etiquette, or religious standards. But I may not judge that such violations are *wrong*.

All this suggests that a more nearly satisfactory account would need to make sense of the connection between judging that an act is wrong and being for blaming that act. Standards appear to fit the bill. For they can make sense of why some behavior *warrants* blame or disapproval. If I am for blaming because an act violates an acceptable standard, then that would make sense of why I am for blaming, provided countervailing conditions do not hold. Moreover, standards can supply an explanation at the right level of description. When we judge something is wrong, we typically judge that it is so in virtue of some constellation of properties it has: it causes gratuitous suffering, or is an invasion of privacy, or is insulting. Standards are helpfully thought of as offering general verdicts about acts of certain types that agents can consult, even when agents are ignorant about or have false views about why these standards hold. A standard that expresses the norm that *one should not kill innocents* offers such a verdict, even if there are different explanations as to why it is wrong to kill innocents and agents hold false beliefs about the matter.

I have not attempted to canvass the entire range of extant expressivist positions that attempt to satisfy No Normative Content. But even without doing so, we can see a pattern. The most prominent versions of expressivism that satisfy No Normative Content make no attempt to satisfy Expressive Power. Moreover, it is difficult to see how they could do so without very considerable supplementation in which they take the form of something akin to standards expressivism. For all I've said, these supplementations might be available (although I am unsure what form they

could take). Nonetheless, the extant views that attempt to satisfy No Normative Content do not make it beyond the core manoeuvre's second step. As such, they are not positioned to execute it.

4 Section 4

The primary question with which I've been concerned is whether expressivist views are well-positioned to execute the core manoeuvre. The answer for which I've argued is that the most prominent versions of the view are not. This answer implies that it would be premature for expressivists to appeal to the core manoeuvre in order to reap its potential theoretical benefits, such as not rendering their view liable to objections of certain kinds, including those levelled against epistemic expressivism. At best, executing the core manoeuvre remains an ideal, which may or may not be realisable.

There are, however, two important issues that remain unaddressed. The first concerns a passage quoted earlier from Gibbard. In this passage, I noted that Gibbard writes that standards or norms are not normative but factual (and, hence, in his view, natural). However, I have proceeded under the assumption that deontically laden imperatival contents—and, more generally, standards—are normative in the sense that they (*inter alia*) concern the permissible, obligatory, and forbidden. This raises the worry that my discussion fails to make contact with Gibbard's position (and perhaps those of other expressivists). For Gibbard apparently does not understand deontic concepts and features to fall under the designation "normative." If he is right about that, his view might be positioned to execute the core manoeuvre. The thinking would be that deontically laden imperatival contents are not normative representational contents. The questions we might raise about them, such as whether they satisfy correctness conditions, are not ones concerning the normative domain.

My view is that Gibbard and his factualist rivals are often talking past one another. The factualists typically have one thing in mind by normativity, while expressivists have another thing in mind. According to the factualists, the normative concerns or involves the evaluative, the directive, the favouring, and the fitting (or some subset thereof). These are

said to be properties that things might have. In contrast, according to Gibbard, normativity is an "element" of thought and language:

> What, though, of the special element that makes normative thought and language normative? There is such an element, I am claiming, and it involves a kind of endorsement—an endorsement that any descriptivist analysis treats inadequately. (1990, 33; cp. 9–10)

What makes a state of mind normative, according to Gibbard, is that it involves endorsing something.

Let me place to the side Gibbard's rationale for thinking that deontic concepts and features do not fall under the characterisation "normative," since it strikes me as unhelpful.[28] Does this difference in views between Gibbard and factualists regarding the normative imply that positions such as standards expressivism are in fact positioned to execute the core manoeuvre?

I doubt it. Distinguish two types of phenomena that the term "normative" might denote. Let's use the term "normative$_F$" ("F" for "factualist") to denote that with which factualists concern themselves: the directive, the evaluative, the favouring, and the fitting. And let's employ the term "normative$_E$" ("E" for "expressivist") to denote that with which expressivists concern themselves, namely, endorsement. By all appearances, factualists and standards expressivists agree that there are entities that are denoted by the phrase "normative$_F$," including such things as deontically laden standards. Factualists and expressivists also agree that there are entities picked out by the term "normative$_E$," namely, states of endorsement. The disagreement concerns which phenomenon is of deep theoretical interest, at least in metanormative inquiry. Factualists maintain that it is the referent of "normative$_F$" that is of primary theoretical interest, while expressivists hold that it is the referent of "normative$_E$" that is of primary theoretical interest. (There might be some who believe that they are of equal theoretical interest.)

We needn't settle who is right about this to see the following point. Suppose a theory were committed to the existence of the Fs (whatever they might be). Proponents of this theory would not undercut the metaphysical project with regard to the Fs simply because they fail to

categorise the Fs in some way that rival theories do. For example, if your theory commits itself to the existence of propositions, it does not thereby undercut the metaphysical project with regard to them simply because it does categorise propositions as the objects of thought. For, even if propositions are not categorised as objects of thought, there are many questions to address regarding their theoretical roles and ontological status. The same is true regarding deontically laden imperatival contents and their correctness conditions. Even if they do not fall under the designation "normative$_F$," there are a wide variety of questions to address regarding their theoretical roles and ontological status—questions that those in the factualist tradition in fact address. It follows that proponents of expressivism do not undercut the metaphysical project because they fail to categorise the deontic realm as normative (while their opponents do).

The second issue I wish to raise concerns a caveat that I've employed throughout our discussion. At various points, I've noted that "under a natural interpretation," standards expressivism implies that normative judgements have deontically laden imperatival content. But advocates of the view might maintain that the view does not imply normative thought and discourse have deontically laden normative representational content. While there are a variety of ways of working this idea out, probably the most obvious alternative is to offer an expressivist (or more broadly non-descriptivist) account of the entire range of deontic modals themselves.[29] In that case, deontic judgements will have representational content but not normative representational content, since deontic modals would not play a representational role with regard to normative reality.

To my knowledge, expressivists have not attempted to provide such an account. Perhaps it could be done. Rather than attempt to assess the prospects of such a project, I will close by noting these two things. First, in order to satisfy Expressive Power and, so, execute the core manoeuvre, such an account would have to be provided. It would be illegitimate to appeal to the core manoeuvre with a mere promissory note that such an account will be developed. And, second, even if such an account were provided, expressivists would need to take the further step of satisfactorily arguing that the correctness conditions for deontically laden

imperatival contents are not themselves normative entities, such as standards with deontically laden content. These are not trivial explanatory projects.[30]

Notes

1. Here is a representative passage: "As we see things, the core expressivist maneuver involves changing the question from one about the nature of epistemic facts or values to one about the nature of epistemic evaluations. Once an expressivist has given his answer to the new question by insisting on a disanalogy between epistemic thought and discourse and descriptive thought and discourse, and once he has argued that this answer undermines the first question…[which is] a question about the nature of some feature of the world whose existence is disputed: epistemic facts or values … he can just stop talking (at least in the capacity of stating his theory)" (Carter and Chrisman 2012, 334).
2. See Schroeder (2008a, 151, b, 8). Cp. also what Dreier (2015) calls the "Expressivist Sidestep".
3. According to Carter and Chrisman, the objectors to whom they are responding "each misunderstand what is crucial to epistemic expressivism" (2012, 334).
4. I am working, then, with a fairly broad characterisation of the normative. But it does mirror that offered by expressivists such as Ridge (2014, chap. 1). If one wishes, one can think of the normative more narrowly as some subset of the normative, such as what is *decisively* reason-giving. In what follows, I remain neutral regarding the relations these different dimensions of the normative bear to one another.
5. In principle, one could understand this project broadly (so that it concerns any attempt to answer the questions above) or narrowly (so that it concerns only certain kinds of attempts to answer these questions, such as those that concern properties and facts that are "substantive"). I understand the project broadly. My reason for doing so is that, even under a narrow understanding of the project, expressivists will face questions that appear to belong to the metaphysical project, and which they seem interested in side-stepping.
6. Carter and Chrisman (2012, 334–35). Although see Chrisman (2011) for a somewhat different approach.

7. As for the "proper aims" of metanormative inquiry, I'll remain neutral regarding exactly what they are and how to substantiate them.
8. As Schroeder (2008a, 151, b, 8) emphasises.
9. Parfit (2011, 408–10, 2017, 191) retracts this charge.
10. In Cuneo (2007, chap. 6). Carter and Chrisman (2012) engage with one argument I offer against expressivism.
11. The two passages in which something akin to the core maneuver is most explicitly voiced are Gibbard (2003, 6) and Blackburn (1999, 50).
12. Carter and Chrisman (2012, 334 n8) offer a reading of what expressivists have in mind when they talk about "deflated" normative facts.
13. I borrow the terminology from Alston (1999, chap. 4). To commit oneself to being in a mental state does not imply that one is in such a state. There can be commitments that are insincere or mistaken.
14. In Cuneo (2014), I develop an argument for the conclusion that normative facts are among those that explain how it is that we could speak. An implication of that argument is that, even if expressivism were to offer a correct description of how normative thought and discourse typically functions, normative facts play crucial explanatory roles; see Cuneo (2014, chap. 6).
15. They might play other interesting explanatory roles too, as Sturgeon (1985, 1992) contends.
16. I assume this is what Gibbard means by "imperatives." Chrisman (2016, 131) includes an illuminating discussion regarding such contents.
17. Ridge (2014) is the most developed version of the view. I find no indications in Ridge's discussion that he endorses the core maneuver.
18. In his development of standards expressivism, Ridge (2014, 27–28) offers an even more nuanced account of directives, maintaining (*inter alia*) that the normative modal "must" adverts to context-specific standards that require, while the normative modal "ought" adverts to context-specific standards that recommend. So, under Ridge's view, it is no surprise that we can move fluidly between sentences such as "You are required to ϕ" and "You must ϕ!".
19. This denial concerns the deep structure of normative thought and discourse. See Blackburn (1993, 184–85). Expressivism's advocates sometimes point out that it would not be inaccurate to say that such thought and discourse have predicative normative representational content, provided that we are speaking only of the surface structure of such thought and discourse or, alternatively, that we have some sufficiently deflated

sense of predication or representation in mind. For the point about surface structure, see Blackburn (1993, 157, 185); for the point about deflationism, see Blackburn (1999, chap. 3).

20. Some content *Φ*'s being about *x* does not entail that *x* exists, is actual, or is instantiated. For example, thoughts *about* unicorns represent unicorns, though there are no such creatures.
21. So, when Gibbard writes that, given his norm expressivism, the "main thing to be explained is not what a norm is, but what 'accepting a norm is'," (1990, 46) it is natural to be puzzled. Given his views about imperatival contents, it makes perfect sense to ask questions about the ontological status of these norms and what they represent, if they have some source, what it would be for one set to be correct, and how they could be "natural," as Gibbard claims. These are the very sorts of questions at issue in the metaphysical project. Their answers are open to a variety of concerns and objections.
22. See the references supplied by Chrisman (2016, sec. 5.3). A crucial component of the expressivist semantics that Schroeder develops are *mistake conditions* (2008a, chap. 10). When assessing this proposal, Schroeder writes that any expressivist view will have to explain what mistakes are in the relevant sense (2008a, 168).
23. Interestingly, when Chrisman offers an example of such an account, he does so by appealing to mere imperatives devoid of deontic modals.
24. Copp (1995, 2007) are exceptions. In his (2007), Copp writes that standards are the "semantic contents expressed by imperatives, and so they are not believed, nor do they represent the world as being one way or another." (14) I am not sure why Copp maintains that imperatives fail to represent the world as being one way or another.
25. See, for example, Wedgwood's (2007, pt. 1) discussion of normative concepts and Cuneo and Shafer-Landau's (2014) discussion of moral concepts.
26. There is a limited effort to extend the account to the category of reasons; see Gibbard's (2003, 188–91) discussion of reasons and Ridge's (2014, chap. 1) critique.
27. To be clear, Gibbard takes the planning-state approach and standards expressivism to be complementary. But, as best I can tell, he does not explain how to mesh the two approaches. The worry is that any attempt to combine them would lack the requisite theoretical unity.

28. Recall it is this: an act "might be *N*-permitted without being rational, for the system *N* might have little recommend it." (1990, 87) But the same is true of endorsements. Many have little to recommend them. So, if this is Gibbard's rationale for rejecting the factualist account of normativity, it is also a rationale for rejecting the account he offers.
29. Chrisman (2016) offers an inferentialist proposal regarding "ought." As best I can tell, there is no discussion of whether the approach could be generalised to all normative modals.
30. Thanks to Adam Carter, Matthew Chrisman, Tyler Doggett, and the UVM Ethics Reading Group for their comments on a draft of this essay.

References

Alston, William. 1999. *Illocutionary Acts and Sentence Meaning*. Ithaca, NY: Cornell University Press.

Blackburn, Simon. 1980. *Spreading the Word*. Oxford: Oxford University Press.

———. 1993. Moral Realism. In *Essays in Quasi-Realism*, ed. Simon Blackburn, 111–129. Oxford: Oxford University Press.

———. 1999. *Ruling Passions*. Oxford: Oxford University Press.

Carter, J. Adam, and Matthew Chrisman. 2012. Is Epistemic Expressivism Incompatible with Inquiry? *Philosophical Studies* 159 (3): 323–339.

Chrisman, Matthew. 2011. Ethical Expressivism. In *The Continuum Companion to Ethics*, ed. Christian Miller, 29–54. New York: Continuum.

———. 2016. *The Meaning of "Ought": Beyond Descriptivism and Expressivism in Metaethics*. Oxford: Oxford University Press.

Copp, David. 1995. *Morality, Normativity, and Society*. Oxford: Oxford University Press.

———. 2007. *Morality in a Natural World: Selected Essays in Metaethics*. Cambridge: Cambridge University Press.

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

———. 2014. *Speech and Morality: On the Metaethical Implications of Speaking*. Oxford: Oxford University Press.

Cuneo, Terence, and Russ Shafer-Landau. 2014. The Moral Fixed Points: New Directions for Moral Nonnaturalism. *Philosophical Studies* 171 (3): 399–443.

Dreier, James. 2004. Metaethics and the Problem of Creeping Minimalism. *Philosophical Perspectives* 18 (1): 23–44.

———. 2015. Explaining the Quasi-Real. In *Oxford Studies in Metaethics*, ed. Russ Shafer-Landau, vol. 10, 273–298. Oxford: Oxford University Press.
Gibbard, Allan. 1990. *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press.
———. 2003. *Thinking How to Live*. Cambridge, MA: Harvard University Press.
Parfit, Derek. 2017. *On What Matters Vol. 3*. Oxford: Oxford University Press.
Ridge, Michael. 2014. *Impassioned Belief*. Oxford: Oxford University Press.
———. 2015. I Might Be Fundamentally Mistaken. *Journal of Ethics and Social Philosophy* 9 (3): 1–21.
Schroeder, Mark. 2008a. *Being For*. Oxford: Oxford University Press.
———. 2008b. *Noncognitivism in Ethics*. Abingdon: Routledge.
Sturgeon, Nicholas. 1985. Moral Explanations. In *Morality, Reason and Truth*, ed. David Copp and David Zimmerman, 49–78. Totowa, NJ: Rowman and Allanheld.
———. 1992. Nonmoral Explanations. In *Philosophical Perspectives 6: Ethics*, ed. James Tomberlin, 97–117. Atascadero, CA: Ridgeview Press.
Wedgwood, Ralph. 2007. *The Nature of Normativity*. Oxford: Oxford University Press.

Epistemic Reductionism and the Moral-Epistemic Disparity

Chris Heathwood

Reductionist doctrines about normative and evaluative phenomena enjoy serious advantages, such as in explaining how we can come to know about normative reality, in explaining why the normative depends on the non-normative, and in avoiding the spectre of Ockham's razor.[1] Unfortunately, some evaluative phenomena resist reduction. This is true, in my view, of moral and axiological facts. When we say that people ought to be more kind, or that things would be better if they were, it does not appear that we could report these same facts using non-normative, non-evaluative language. But things are different, I believe, when it comes to epistemic facts. When we say that someone is justified in believing something, we can report that same fact using non-normative language. Reductionism in metaepistemology is more plausible than reductionism in metaethics.

This moral-epistemic disparity is not only interesting in its own right, it shows that an important and popular line of argument for a robust

C. Heathwood (✉)
Department of Philosophy, University of Colorado Boulder, Boulder, CO, USA
e-mail: heathwood@colorado.edu

© The Author(s) 2018
C. Kyriacou, R. McKenna (eds.), *Metaepistemology*, Palgrave Innovations in Philosophy, https://doi.org/10.1007/978-3-319-93369-6_3

moral realism does not succeed. According to this argument, there is a kind of moral-epistemic parity: epistemic facts and moral facts share deep similarities, so that whatever account is most plausible of the nature of one will be most plausible of the nature of the other. Furthermore, the argument continues, a robust realist account of epistemic facts is most plausible. Consequently, we have good reason to accept robust moral realism. In my view, this argument founders on the fact that whereas we are forced to reject a reductive naturalist account of moral facts, we are not so forced when it comes to epistemic facts. Committed as I am to robust moral realism, I don't welcome this result. But I believe that this is where the arguments lead.

I have defended epistemic reductionism and the moral-epistemic disparity in earlier work (Heathwood 2009). The purpose of the present essay is to clarify and refine this position and to defend it against recent criticisms by Richard Rowland (2013) and Terence Cuneo and Christos Kyriacou (2018).[2]

1 Preliminaries

I feel quite confident in certain ethical claims, some to the effect that some things that people do are not ok, others to the effect that some states of affairs are good in themselves. Consequently, I believe that there are such properties as those of being morally wrong and being intrinsically good. But what sort of properties are they?

According to *reductionism* in metaethics, moral facts and properties are identical to facts and properties from some other domain—facts and properties that can be adequately captured, expressed, or stated using wholly non-moral vocabulary. Most reductionist theories in ethics reduce ethical facts and properties to natural—that is, empirical—facts and properties. Psychological properties are the most common. Hume's famous remark that "when you pronounce any action or character to be vicious, you mean nothing, but that from the constitution of your nature you have a feeling or sentiment of blame from the contemplation of it" (1975, 469) is suggestive of a reductive naturalist account of the moral property of viciousness. A theory according to which the property of

being intrinsically good just is the property of being a state of pleasure is another example of reductive naturalism in metaethics.

Non-naturalists in ethics, by contrast, hold that moral properties are their own kind of thing, that they cannot be reduced to or identified with any natural or descriptive property.[3] Each moral property is what it is and not another thing. Crucially for our purposes, this is not to deny that there are necessary connections, and indeed explanatory ones, between moral and non-moral facts. Virtually everyone believes in these. So, for example, some non-naturalists about intrinsic value are hedonists. They hold that whenever a thing is intrinsically good, this is in virtue of the fact that that thing is a state of pleasure. They hold that the property of being intrinsically good is necessarily coextensive with the property of being a state of pleasure.[4] But they hold that this connection is a synthesis of two distinct phenomena rather than, as reductionists hold, a single phenomenon with two different, equally legitimate characterisations, one evaluative and the other descriptive.[5]

I should also make clear how I understand the moral realism/anti-realism distinction. As I use these terms, moral realism is the doctrine that there are some moral facts and that at least some of the moral properties involved in these facts are "stance-independent."[6] A property is stance-independent just in case whether a thing has the property does not depend upon the attitudes or practices that certain (possibly hypothetical) observers take up towards that thing. Thus, the Humean view above is not a realist view, since it makes viciousness stance-dependent, while the hedonist view above is a form of realism.[7] Although the moral realism/anti-realism distinction cuts across the moral reductionism/non-reductionism distinction, virtually all non-reductionists are realists. I use the term "robust moral realism" to refer non-reductionist moral realism of a certain sort: non-naturalist moral realism.

Corresponding to these categories in metaethics are analogous categories in metaepistemology. Just as there are properties like that of an act's being morally wrong or a state of affairs' being intrinsically good, there are the properties of having a reason to believe something or of a belief's being epistemically justified or reasonable. About these epistemic properties, the same questions arise. Can they be reduced to properties from some other domain? Are they stance-independent?

Some of us are occasionally gripped by moral scepticism. We are visited by genuine doubts about whether anything anyone does really is forbidden or whether any states of the universe really are any better or worse than any others. Such doubts are rare, but even rarer are doubts about whether any *beliefs* are any more or less *reasonable* than any others. For this reason, if there is moral-epistemic parity, the near irresistibility of the idea that there are epistemic facts can be used to quell doubts over the existence of moral facts.

2 The Moral Open Question Argument

It seems to me that moral properties and facts resist analysis in terms of natural facts and properties, or any other facts and properties that can be adequately captured in non-moral terms. My belief is based on open-question-type considerations. The Open Question Argument has been discussed, defended, criticised, reformulated, and re-criticised for generations; I will content myself with laying out how I understand the argument in the moral case, which will put us in a position to compare it to an analogous argument in the epistemic case.[8]

Reductive naturalists about some normative or evaluative property identify it with some natural property. This view can take one of two forms. On *analytic reductionism*, the natural-normative identity claim is analytic; on *synthetic reductionism*, it is synthetic. The Open Question Argument makes trouble for analytic reductionism.[9] Analytic reductionism will be our focus here since it is, in my view, the crux of the moral-epistemic disparity: an analytic version of epistemic reductionism does not fall prey to the Open Question Argument, whereas all versions of analytic ethical reductionism do fall prey.

The trouble with analytic reductionism in ethics is that whatever natural property is put forth in the reductionist theory as the one that is identical to a certain ethical property, it will be an "open question" whether the things that have the natural property also have the evaluative property. That is to say, of a thing that is known to have the natural property, neither answer ("Yes" or "No") to the question of whether it also has the evaluative property would suggest that the person giving that answer fails

to grasp the concept of that evaluative property; nor would either answer suggest that the person simply means something peculiar when they use the corresponding evaluative term. Relatedly, neither answer seems incoherent, even if one of them is known to be false.

To illustrate, suppose that, after working on the topic for some time, considering all the many known theories and arguments and counterexamples, you come to the view that hedonism is the correct axiological theory. You believe that

> all states of pleasure are good in themselves and only states of pleasure are good in themselves.

Thus, you believe that the properties of being a state of pleasure and being intrinsically good are necessarily coextensive. You then encounter a fellow philosopher who, having scrutinised the same data, disagrees. This philosopher thinks that malicious pleasures are of no intrinsic value. You tell him that of course you agree that a malicious pleasure may have bad effects, and that malicious pleasures are a sign of a vicious character. Is he sure that he isn't thinking of those things when he refrains from evaluating malicious pleasure positively? Yes, he says, he is sure; he has *intrinsic* value in mind. And on the question of the intrinsic value of malicious pleasure, it continues to seem to him that malicious pleasure is of no intrinsic value.

Committed hedonist that you are, you have no choice but to think that your colleague has simply got this one wrong. His view is mistaken. But will you think, further, that his view betrays a conceptual confusion, or a failure to grasp the very question of whether something is intrinsically good, or is an incoherent view? I don't think so. As much as we might enjoy being able to accuse a staunch philosophical opponent of these things, such an accusation is not plausible in this case. Nor would you think that your colleague just must mean something different by "intrinsically good" than you do. Instead, how things seem is not that you and your colleague are talking past each other but that you simply have different beliefs about a single topic.

This finding has metaethical implications for you. You are convinced that *being intrinsically good* and *being a state of pleasure* are two properties

that necessarily go together. But since it does not betray a conceptual confusion to say, of a thing that is known to be a state of pleasure, that it is not intrinsically good, you conclude that you cannot analytically reduce intrinsic goodness to the property of being a state of pleasure, or regard "intrinsic goodness" and "being a state of pleasure" as two names for a single notion. Rather, to say of a thing that it is an episode of pleasure is one thing, and to say of a thing that it is intrinsically good is another. The Open Question Argument thus convinces you that analytical reductionism in metaethics is mistaken.

The point here is not just one about the particular identification of the concepts of being intrinsically good and being a state of pleasure. The same line of reasoning would go through no matter which natural or non-evaluative concepts we thought to be necessarily co-extensive with the concept of intrinsic goodness. Those who affirmed evaluative judgments in conflict with this necessary connection would not thereby seem to us to be conceptually confused or to be talking about some other topic. We would simply take them to be speaking falsely about our shared topic.

Some people object that this sort of argument proves too much. They think that it could be used to refute reductive analyses on any topic, or at least any reductive analysis that is less than obvious. I agree that there are some non-obvious analytical reductions, but I don't think this argument conflicts with that fact.

Before considering a case of a non-obvious reduction, it will be useful to begin with an example of an obvious one. Consider unclehood. After briefly reflecting on what it is to be an uncle, you come to the view that

> to be an uncle is to be a male who has a sibling who has an offspring.

You then encounter a fellow inquirer who, having considered the matter as well, disagrees. He insists, let's suppose, that the offspring has to be a male for the person to be an uncle. Before us is Joe. Joe, we all agree, has a sibling with an offspring. But is Joe an uncle? Our interlocutor answers, "No" (because the offspring is a girl). Committed as you are to the view that to be an uncle is to be a male who has a sibling who has an offspring, you have no choice but to think that your fellow inquirer has got this one wrong. His view about Joe is mistaken. Will you think, further, that his

view betrays a conceptual confusion, or a failure to grasp the very question of whether someone is an uncle? *Yes*. Your fellow inquirer simply fails to grasp unclehood. He means something different from the rest of us when he uses the term "uncle" (he also means something different from what the term actually means). Consequently, the question of whether Joe is an uncle is *not* an open question. Thus, the Open Question Argument, when applied to our analysis of unclehood, does not generate a false positive in this case of an obviously correct reduction.

Nor does it in the case of non-obvious reductions. Consider bachelorhood. Suppose that you are thinking about the extension of this property. After considering the known theories and arguments and counterexamples, you come to the conclusion that what we can call "the simple view of bachelorhood" is right:

> all unmarried men are bachelors and all bachelors are unmarried men.

You then encounter a fellow inquirer who, having scrutinised the same data, disagrees. She thinks that the pope is not a bachelor. You tell her that of course you agree that the pope is not an *eligible* bachelor, and you note that often when we are interested to know whether some man is a bachelor, we are interested in his eligibility for marriage. Is she sure that she isn't just thinking of eligibility for marriage? Yes, she says, she is sure; she is talking about plain old bachelorhood, and it seems to her that the pope simply does not qualify as a bachelor.

Committed defender of the simple view that you are, you have no choice but to think that your colleague has got this one wrong. Her view is mistaken. Will you think, further, that her view betrays a failure to understand the concept of bachelorhood? Yes; you *should* think this. You should think that she just does not get what it is to be a bachelor. Though it is more understandable in this case than in the unclehood case, since it is less obvious what the correct analysis of bachelorhood is. In the hedonism case, you don't think that your interlocutor fails to get *what it is* to be intrinsically good; he simply fails to see which things have this property. But in the bachelorhood case, if you are convinced of the simple view of bachelorhood, you should think that your interlocutor just fails to get what it is to be a bachelor. She must mean something different

when she uses the term "bachelor," something different from what you think the term means. Since you think that *in fact* the right property to associate with the term "bachelor" is given by the simple view, you will think that she is linguistically or conceptually confused. Thus, an Open Question Argument designed to show that we should be non-reductionists about bachelorhood would fail. And it would fail even though it is not obvious what the correct analysis of bachelorhood is.

When we encounter someone who holds a belief that conflicts with our hedonistic theory of intrinsic value, we judge, perhaps after some initial questioning, that they are working with the same concept as us—the concept of intrinsic value. We just think that they hold different beliefs about when the concept applies. This is the normal case of disagreement, as when people disagree about, say, whether alligators eat their young. The parties to such disagreements share the same concepts. But when we encounter someone who holds a belief that conflicts with our simple view of bachelorhood, we conclude that they in fact have a different *concept* of bachelorhood and that the meaning that they associate with the word "bachelor" is different from what the word actually means. The Open Question Argument is a good argument against analytic reductionism in metaethics, but it would not be a good argument against any analytic reductionist view about bachelorhood.

This should go without saying, but, as with any philosophical argument, it is always possible for a stubborn advocate of the view being attacked simply to reject the key premise on which the argument against their view is based. Thus, the committed analytic reductive hedonist can respond to the colleague who is convinced that malicious pleasure is not good by saying that the colleague just must not get the concept of intrinsic goodness; they must not get what we are saying about something when we say that it is intrinsically good; they must just mean something different by "intrinsically good" from what the word actually means. If someone is not prepared to reconsider their theory, then this is probably what they should say. Our job, then, is to evaluate the plausibility of saying this. For my part, this imagined reply strikes me as implausible, as something that only someone in the grip of a theory would say. Thus, it seems to me that, all things considered, the Open Question Argument

casts doubt on analytic reductionism in ethics, though an analogous argument does not cast doubt on analytic reductionism about unclehood or bachelorhood.

3 The Epistemic Open Question Argument

So it is with ethics and uncles and bachelors. What about epistemology? I believe that when it comes to analytic reductionism about epistemic properties, the Open Question Argument is not compelling. It is like the Open Question Argument against analytic reductionism about bachelorhood or unclehood. Since the Open Question Argument is the main obstacle standing in the way of analytic reductionist doctrines about normative topics, and since it fails in the epistemic case, we should accept analytic epistemic reductionism. And, interestingly, we should do so even if we aren't sure just what the correct reductive analyses of the epistemic notions are.

To explain why I think the epistemic Open Question Argument fails, we can begin with a sample analysis of epistemic reasons or reasonableness in non-normative terms. But I would like to emphasize at the outset that it does not matter whether the sample analysis is perfectly extensionally adequate. For our question is the question of whether epistemic reductionism as a general thesis about the nature of epistemic facts, abstracted away from any particular proposed reduction, is true. My method for answering this abstract question is as follows. First, we identify an analysis that is at least not obviously extensionally hopeless; *prima facie*, it at least has a shot at being extensionally adequate. Next, we *suppose* that this analysis is in fact extensionally adequate (we can do this even if we suspect or even know that it is not). Then we ask, within the scope of this supposition, how plausible it would be to hold that not only does the condition on the right-hand side of the analysis necessarily covary with that on the left-hand side, but they are in fact just one condition under two labels. And the method for answering this is the open-question test (concerning conceptual confusion, idiosyncratic meanings, and incoherence) that we have been working with.

Here is a kind of first-pass theory of the concept of reasonable belief that is at least not obviously extensionally inadequate on its face:

> a claim is reasonable for a subject to believe just in case the claim is likely to be true, given the subject's evidence.

To illustrate, suppose Sally looks out her window first thing in the morning and sees that the streets are wet. Relative to Sally's prior background beliefs together with this new piece of information, it is very likely *that it recently rained*. Moreover, intuitively, given the case as described, it is in fact reasonable for Sally to believe that it recently rained. Thus, we have a confirming instance of our account of epistemic reasonability in terms of probable truth.

Suppose we find that the theory in question is in fact extensionally adequate.[10] There are no counterexamples, or cases in which a belief is, intuitively, epistemically reasonable for some subject but not probable relative to their evidence (or vice versa). Our first-order epistemological investigation into the conditions under which a belief is reasonable would be complete (at least in the absence of new data on the matter). But a metaepistemological question of central importance would remain. When we say that a claim is reasonable for a subject to believe, *is all we are saying* that that fact is likely to be true given the subject's evidence? Or, alternatively, are a claim's being reasonable for a person to believe and its being probable relative to their evidence two distinct states of affairs—one irreducibly normative and one natural—that necessarily go together? In other words, is epistemic reductionism true or is epistemic nonreductionism true?

To answer this question, we can apply our open-question reasoning as earlier described. We can suppose that, although we are convinced of and committed to the extensional adequacy of our epistemological theory (the biconditional on display above), we encounter, just as in the earlier cases, someone who holds a belief in conflict with it. Let's suppose that it is Sally, about her own case above. Sally agrees that, given her information, it is likely that it recently rained. But Sally, to our surprise, denies that it is reasonable for her to believe that it rained. Here is what she says: "I see that the streets are wet. And I realise that, given this, and given

what else I know, it is very likely that it rained recently. Nonetheless, I don't think that it would be reasonable for me to believe that it rained recently."

If you are like me, these remarks of Sally's will strike you as very odd indeed. She is saying roughly this: "Yes, that claim is, from my perspective, very probably true, but it is not reasonable for me to believe it." These remarks seem more similar to the words of the person who thinks that, for you to qualify as an uncle, your sibling needs a male child than they are to the remarks of the person who thinks that malicious pleasure does not qualify as intrinsically good. Sally sounds conceptually confused, like she doesn't understand what she is saying.

Note that her claim is worse than Moore-paradoxical. Moore-paradoxical claims, such as of the form, "that claim is true, but I don't believe it," are at least logically consistent, even if seriously infelicitous. But Sally's remark sounds incoherent. At a minimum, we do not have a positive intuition that the claim *is* self-consistent, as we do in the case of the claim, relative to the supposition that hedonism is true, that malicious pleasure is not intrinsically good. To build a case for metaepistemological non-naturalism on the basis of the idea that Sally's view is coherent would be to place quite a bold thesis on quite a frail reed.

It is possible to get Sally's claim to sound positively coherent, but only if we interpret the term "reasonable" in the wrong way, as signifying some other sort of normative evaluation, such as a prudential one. Pascal's position was something like this: "given all our evidence, theism is just as likely to be false as it is to be true, but, still, it is reasonable for us to believe that it is true." A starker variation on Pascal is this: "given all our evidence, theism is very unlikely to be true but, still, it is reasonable for us to believe that it is true." These Pascalian claims are indeed coherent. But that is only because we are interpreting "reasonable" to mean *prudentially* reasonable. That such claims are coherent is of no help in supporting the sort of claim that defenders of the epistemic Open Question Argument need. They need claims like the following to be coherent: "given all our evidence, theism is very unlikely to be true but, still, it is *epistemically* reasonable for us to believe that it is true."

I conclude that, unlike the moral Open Question Argument, the epistemic Open Question Argument is not convincing. Consequently, in the

epistemic realm, the door is open to embrace analytic reductionism and to enjoy its attendant theoretical advantages. The only disadvantage is to those of us who were hoping to be able to use the premise of moral-epistemic parity to support robust moral realism. Unfortunately for us robust moral realists, epistemic normativity appears not, after all, to be a partner-in-guilt of irreducible moral normativity.

4 What If the Proposed Analysis of Epistemic Reasonability Is Not Extensionally Adequate?

4.1 Richard Rowland's Counterexample

In a recent paper, Richard Rowland argues against the moral error theory on the grounds that it implies that there are no epistemic reasons (Rowland 2013). In doing so, he considers the position that I favour, "that a conceptual analysis of epistemic reasons in nonnormative terms is more plausible than a conceptual analysis of moral reasons in nonnormative terms" (Rowland 2013, 7). But he rejects this moral-epistemic disparity on the grounds that the particular analysis that he considers—one that reduces facts about epistemic reasons to facts about probability—is not extensionally adequate.

I don't believe that Rowland's arguments should make us worry that epistemic facts are not reducible after all. First, as a general matter, since epistemic reductionism is a doctrine in metaepistemology, we don't have to worry whether our working analysis of epistemic facts in terms of facts about likelihoods in perfectly extensionally adequate. Our central question is this: supposing (perhaps contrary to fact) that our theory is extensionally correct—in other words, that the corresponding biconditional is necessarily true—should we conclude (a) that, as in the unclehood and bachelorhood cases, the theory is analytic and what the right-hand side says is just another way of saying what the left-hand side says, or should we conclude (b) that, as in the moral case, the extensionally correct biconditional is a synthetic necessary truth rather than an analytic one?

In the previous section, I argued for (a). That argument suggests that epistemic reductionism is correct even if our supposition (that the particular biconditional under consideration is true) is mistaken. Showing via first-order epistemological considerations, as Rowland tries to do, that the theory is in fact not extensionally correct does not undermine this metaepistemological position.

These abstract points can be illustrated by examining Rowland's particular counterexample. Although Rowland's specific target is my (2009) paper, the analysis he criticises is a little different from the main one discussed in that paper.[11] According to the analysis Rowland focuses on,

> the epistemic predicate "is an epistemic reason to believe *p*" means the same as the naturalistic predicate "is a fact that raises the probability that *p*." (Rowland 2013, 8)

Rowland thinks that this equivalence fails, but not for open-question-type reasons. It fails, according to Rowland, for first-order epistemological reasons: it isn't even extensionally correct.

Consider the fact that I am in my office today. This raises the probability that I will die in my office today. That is to say, if I were at no point in my office today, the probability that I die in my office today would have been less than it actually is. "But," Rowland claims, "it does not seem that the fact that I am in my office today is a reason to believe that I will die in my office today; I am perfectly healthy, and probably safer in my office than anywhere else" (2013, 8).

Rowland's counterexample is interesting, but I don't think that it is successful in the end. Moreover, even if it were, it wouldn't matter for our central concern, which is not the precise extensional adequacy of any particular analysis but the prospects of going reductive once an extensionally adequate theory has been settled on.

About Rowland's particular counterexample, my considered judgment is that the fact that I am in my office today *is* a reason to believe that I will die in my office today. Intuitively, it *is* a consideration in favour of that belief. After all, and as Rowland agrees, that I am in my office today *makes it more likely* that I will die there today. How could it not then *support the belief* that I will die in my office today? Notice how weak the claim in

question is. Saying that the fact that I am in my office today is *a* reason to believe that I will die in my office today is just to say that it adds *some* support for the belief, or that it counts in its favour *at least somewhat*. It does not imply that one *should* believe that I will die in my office today. Of course one shouldn't believe that (given the other details of Rowlands' case: that I am perfectly healthy and that my office is very safe). But there can be reasons in favour of beliefs even when those beliefs are, all things considered, unreasonable to have.

That the fact in question does provide a reason in this case can be seen by comparing two irrational believers. The first knows that I am in my office and believes, irrationally and despite overwhelming evidence to the contrary, that I will die there today. In the second scenario, at no point today am I in my office. Our second believer knows this and yet believes, also irrationally and despite even more overwhelming evidence to the contrary, that I will die there today. Both believers are irrational, but the second one is more irrational. Why? The best explanation of this will involve the claim that the first believer's belief had at least one thing going for it that the second believer's belief did not: namely, that I was, in that first case, at least in my office, a fact that at least made my dying there more likely. The first believer is less irrational because he had at least this reason for his belief. Thus, that I am in my office today is indeed a reason to believe that I will die in my office today.

But even if you agree with Rowland's counterexample, that doesn't affect our larger question, the question of epistemic reductionism in general, abstracted away from the extensional adequacy of any particular reduction. To answer the question that concerns us, we can first *suppose* that in fact Rowland is wrong and that the analysis that he discusses is extensionally adequate. Then, within the scope of this supposition, we consider whether it is more plausible to go reductive or to go non-reductive. When we do, we find that there are no open-question-type pressures to go non-reductive. Supposing that all and only cases of a fact giving a reason to believe are cases of probability raising, the thought "I see that this fact raises the probability that *p*, but I don't think it gives me any reason to believe that *p*" isn't just false but is plausibly regarded as incoherent.

Another way to see why Rowland's putative counterexample should not worry the epistemic reductionist, even if it is a successful counterexample,

is to notice that once, via first-order epistemological theorising, we repair our analysis to avoid the counterexample, we can apply the open-question test again, and again it will support reductionism over non-reductionism. This is not the place to get lost in the salt mines of first-order epistemology, but here is a quick illustration. Perhaps we learn from Rowland's example that our analysis should be formulated as follows:

> for a fact to be a reason *to increase one's credence* in *p* just is for that fact to raise the probability that *p*, given one's evidence.

This theory avoids Rowland's counterexample. The fact that I am in my office today clearly is a reason to *increase one's credence* in the proposition that I will die in my office today (presumably from a very low credence to a slightly higher but still very low credence), even if it isn't a reason to *positively believe* that I will die in my office today. Now imagine someone saying, "I see that the fact that you are in your office today increases the probability that you will die in your office today, but, still, I don't think it gives me any reason to increase my credence that you will die in your office today." That has the same air of incoherence about it that our earlier examples did. This is further confirmation for epistemic reductionism. I conclude that Rowland has not given us good reason to doubt this doctrine.

4.2 Cuneo and Kyriacou's Counterexample

Terence Cuneo and Christos Kyriacou's recent paper, "Defending the Moral/Epistemic Parity," is a sustained attack of my (2009) attempt to establish moral-epistemic disparity (Cuneo and Kyriacou 2018). They advance a number of interesting criticisms, one of which begins as a kind of counterexample to my working analysis of epistemic reasonability, the theory that for a belief to be epistemically reasonable for some subject is for it to be probable given the subject's evidence.[12]

In Cuneo and Kyriacou's counterexample, you walk into a room and have a visual experience as of a table in front of you. On this basis, you form the belief that there is a table in front of you. Just before you step into the room, however, there is a small placard that enters your visual

field but that you do not read or even notice. The placard indicates that you are about to encounter a masterful *trumpe l'oeil* mural containing an image of a table. Cuneo and Kyriacou write,

> Given all your available evidence, it follows that the proposition *that there is a table in front of me* is not likely and, thus, according to Heathwood's proposal, it is not reasonable for you to believe that there is a table before you. But, by all appearances, it is reasonable for you to believe that there is a table before you. By no fault of your own or your eyes, you simply missed taking into account some available but not easily detectable information. (Cuneo and Kyriacou 2018)

I think that Cuneo and Kyriacou are too quick to assume that my proposal implies that it is not reasonable for you to believe that there is a table before you. For it is not clear that the information on the placard is a part of your evidence. Indeed, intuitively, I would say, what is written on the placard is not a part of your evidence. And note that some prominent accounts of evidence possession agree. According to Timothy Williamson's view, for example, a person's evidence consists of everything they know (Williamson 2000, 184–85). Since, in the case above, you don't know what the placard says, this is not a part of your evidence on Williamson's account. And according to Richard Feldman's account, a person's evidence at a time consists of everything that they are thinking about or are aware of at that time (R. Feldman 1988). The information on the placard is not a part of your evidence on this account either. My theory leaves it unsettled what it is for some proposition to be a part of a subject's evidence. This is by design, to avoid getting mired in first-order epistemological controversies. But I suspect that the correct view of evidence possession will agree with the views of Williamson and Feldman that, in the imagined case, the information on the sign is not a part of your evidence.[13]

In concluding that, on my proposal, it is not reasonable for you to believe that there is a table before you, Cuneo and Kyriacou cite your "available" evidence. I concede that there is a reading of "available evidence" on which the information on the placard is *available* to you. And in my previous paper I slid incautiously between the notions of evidence *in a subject's possession* and evidence *available to a subject*. I consider the

present reply to Cuneo and Kyriacou, then, as an opportunity to lay out my position more clearly and carefully. Though I hasten to add that my main argument should still go through despite my incautiousness. That argument asks us only to *suppose* that the theory in question is extensionally adequate, or immune from counterexample, and then to consider whether, given that, one should prefer a reductive or a non-reductive version of the theory.

5 Cuneo and Kyriacou on "Relevant Evidence"

A theme throughout Cuneo and Kyriacou's paper is that, for the purposes of my analysis, which makes use of the notion of a person's evidence, it will not do to appeal to the person's *total* evidence. That would be too indiscriminate, as the *trumpe l'oeil* case is meant to bring out. In Cuneo and Kyriacou's view, the best alternative is to appeal the subject's *relevant* evidence, where this amounts to the evidence that the subject *ought* to take into account (Cuneo and Kyriacou 2018). They worry that while this, an account stated in terms of relevant evidence, might make for an extensionally correct analysis, it will no longer be a descriptive analysis of the epistemic notion—and thus my case for moral-epistemic disparity would be undermined.

But now that we have clarified that the analysis is appealing to evidence that the subject *possesses*, and not evidence that is in some wider sense *available* to the subject, it is plausible for the view to appeal to the total evidence that the subject possesses. One worry of Cuneo and Kyriacou that goes away when we restrict the theory to possessed evidence is the worry that it is not possible for a subject to base a belief on their total evidence. For perhaps Feldman is correct and a subject's evidence—their total evidence—consists merely of everything that they are currently thinking. If so, Cuneo and Kyriacou's worry disappears, since there is presumably no difficulty in basing a belief on everything that one is currently thinking. Or perhaps Williamson's more capacious view is correct and a subject's evidence is everything that they know. Again, there seems to be no great worry here; a person can base a belief on everything

they know. And if there is a worry, that may simply be a reason to favour a more restrictive view of evidence possession, such as Feldman's, over a more capacious one. We, or, better, first-order epistemologists, can in fact construct their theory of what it is to possess evidence with the aim of making the theory of epistemic reasonability more plausible.[14] It is not our place to settle such matters. But once they are settled by first-order epistemological investigation, there is no reason to think that the epistemic Open Question Argument will be any more effective in undermining epistemic reductionism.

Let me make one final point, concerning the idea that the notion of relevant evidence should be used and understood normatively, as Cuneo and Kyriacou suggest. I have just explained why we don't need to take the route of appealing to relevant evidence. But even if we were to take this route and were convinced that the notion of relevant evidence should be understood normatively—as meaning *evidence that the subject ought to take into* account—this would not thereby undermine the case for epistemic reductionism. Because the normative supervenes on the descriptive, we know in advance that there is some (possibly complex and disjunctive) descriptive property on which the property of *being evidence that some subject ought to take into account* supervenes. If we were to take Cuneo and Kyriacou's suggested route, then, given the general failure of the Open Question Argument for epistemic notions, once we have identified this descriptive property, we would be entitled to reduce the normative notion of *evidence that a subject ought to take into account* to it.[15]

6 Cuneo and Kyriacou's Argument for Moral and Epistemic Parity as Regards the Prospects for Reduction

In addition to airing the above suspicions about the kind of reductive analysis of epistemic normativity that I favour, Cuneo and Kyriacou present a positive argument for the view that morality is on a par with epistemic rationality. They call attention to a challenging kind of dynamic for formulating different sorts of reductive analyses of normative notions. On the one hand, if the *analysans* is conceptually rather similar to the

analysandum, then the analysis is more likely to seem plausible but less likely to seem genuinely reductive (that is, the *analysans* will include, or at least might seem to include, normative concepts, just like *analysandum*). If, on the other hand, the *analysans* is conceptually very dissimilar from the *analysandum*, then the analysis is more likely to seem genuinely reductive but less likely to seem true. A successful reductive analysis must steer a middle ground between these two pitfalls.

Cuneo and Kyriacou think that, when looked at in light of this challenge, the moral and epistemic cases are closely analogous. We can formulate a spectrum of analyses in each case, and the prospects of success for the various positions on the spectra will seem similar in each case. This supports, Cuneo and Kyriacou believe, moral-epistemic parity as regards the prospects for reduction.

More specifically, Cuneo and Kyriacou think that my analysis of epistemic reasonableness and an imagined analysis of a moral notion (moral goodness) occupy similar positions on their respective spectra such that, in each case, it is both hard to say whether the analysis is true and, even if it is true, unclear whether the *analysans* contains a normative notion. Their imagined analysis of moral goodness is this:

> a thing is morally good just in case it would be prized by an ideal agent in idealised conditions. (Cuneo and Kyriacou 2018)

Because, they suggest, this theory is perhaps somewhat plausible yet not clearly true and also not so clearly genuinely reductive, they think the moral case is relevantly similar to the epistemic case. For in the epistemic case they think that my analysis of epistemic reasonableness in terms of probability is perhaps somewhat plausible yet not clearly true and also not so clearly genuinely reductive. They conclude that "we have as much reason to believe that we can furnish a reductive analysis of epistemic concepts as we have to believe that we can furnish a reductive analysis of moral concepts". (Cuneo and Kyriacou 2018)

I continue to think, however, that crucial disparities remain. One difference is that Cuneo and Kyriacou's imagined analysis of moral goodness, as an ideal-observer analysis, admits of two familiar interpretations, one of which is clearly not reductive and the other of which clearly is. The

covertly normative, non-reductive version is one in which the definition of "ideal" or "idealised" contains normative terms; the truly reductive version contains only psychological and other naturalistic notions in these definitions. That the *analysans* in the moral case admits of these two familiar interpretations, one clearly normative and the other clearly not, is one difference between the moral and epistemic cases. The other, more crucial difference becomes manifest once this first difference has been brought to light: if the analysis of moral goodness is interpreted as a genuinely reductive one, it will fall prey to the Open Question Argument.[16] This is in contrast to our epistemic analysis, which is of course intended to be understood as a genuinely reductive one. Even when this analysis is understood as intended, as genuinely reductive, an Open Question Argument against it fails. Thus, the crucial moral-epistemic disparity remains.

A truly reductive version of Cuneo and Kyriacou's imagined theory might hold, as Roderick Firth's well-known reductive version of the theory held, that the ideal agent in idealised conditions has these features: this person is omniscient with respect to non-ethical facts, omnipercipient, disinterested, dispassionate, consistent, and in other respects normal (Firth 1952, 333–45). It is a very difficult empirical question what things such a being would prize, but let's suppose that it turns out that such a being would prize (intrinsically, or for its own sake) even malicious pleasure. Even so, the following remains an open question: I see that a being who is omniscient with respect to non-ethical facts, omnipercipient, disinterested, dispassionate, consistent, and in other respects normal would prize malicious pleasure for its own sake, but is malicious pleasure really good in itself? A person who held that the answer to this question is "No" would not thereby indicate that she meant something different by "good in itself," would not thereby manifest conceptual confusion, and would not appear to be holding an incoherent position. Thus, a genuinely reductive version of Cuneo and Kyriacou's imagined theory falls prey to the Open Question Argument.

The other main version of the view, the version that is covertly normative, arguably does not fall prey to the Open Question Argument. One simple (if simplistic) way for the view to have an ultimately normative *analysans* would be for the characterisation of the ideal observer to include

omniscience about *ethical* facts together with a disposition to prize things in accord with one's ethical beliefs. The following is not an open question: I see that a being who is omniscient with respect to ethical facts and who prizes only what they judge to be good would prize malicious pleasure, but is malicious pleasure really good? A "No" answer to this question is plainly incoherent.

But my proposed version of epistemic reductionism is not covertly normative. Or at least that is and has been my position, and the argument of Cuneo and Kyriacou described in this section has given us no reason to think that my version of epistemic reductionism has a covertly normative *analysans*.[17] But Cuneo and Kyriacou have attempted to give such reasons elsewhere in their paper. To these I now turn.

7 Is Probability Normative?

Cuneo and Kyriacou argue that if my analysis of epistemic reasonability is to be plausible, the notion of probability at work in it needs to be covertly normative. They present a dilemma:

> Suppose, on the one hand, that we attempt to analyze the concept being reasonable by appeal to an *objective intrinsic* account of probability, as Heathwood suggests. If we do, then we will not be able to fashion an adequate analysis of the concept being reasonable. Suppose, on the other hand, we attempt to analyze the concept being reasonable by appeal to a *conditional epistemic* account of probability. If we do, then we might arrive at an adequate analysis of this concept. That analysis, however, would provide no reason to believe that probability facts are descriptive; to the contrary, we'll suggest, it would provide reason to believe that such facts are normative, in which case Heathwood's case for the disparity thesis would collapse. (Cuneo and Kyriacou 2018)

I agree that neither of these accounts of probability is suitable for the role needed by my reductive analysis of epistemic reasonableness. Objective intrinsic probability, as Cuneo and Kyriacou understand this expression, won't do because a subject may have no evidence pertaining to an event's objective intrinsic probability. To take a variant of their

example, suppose you have a biased coin in your hand—both sides are "tails"—but I have every reason to think that it is a normal, fair coin. The objective intrinsic probability of its landing on "tails" on each of the next ten tosses is 1 (or very nearly 1), but it would not be reasonable for me to believe that it will land on "tails" on each of the next ten tosses.

Conditional epistemic probability, as Cuneo and Kyriacou understand it, won't do either. For as they understand it, it is a normatively loaded notion, amounting either to a normatively loaded kind of ideal epistemic agent account or to an account that appeals, again, to relevant evidence, which they believe must be understood normatively.

I reject both of these understandings of the relevant sort of probability. The notion that is playing a role in my analysis, the notion of *a claim being likely, given some subject's evidence* is at once *subject-relative* and *objective*. It is subject-relative in that the likelihood is relative to the subject's evidence. And it is objective in that whether that body of evidence makes the claim likely is an objective fact, made true by objective relations between that evidence and the claim in question, and not, say, by the agent's wishes or whims. This kind of objective subject-relativity is familiar. It appears, for example, in the unremarkable notion of *a claim being entailed by some subject's beliefs*. Such a notion is obviously subject-relative: what is entailed is entailed by this particular subject's beliefs. It is also objective: whether some proposition is entailed by some specified set of propositions is an objective fact about the world.

8 Concluding Remarks

I believe that, though the Open Question Argument shows that reductionism in ethics is not plausible, it does not show the same about reductionism in epistemology. I have here tried to lay out more fully and clearly why I believe this, and to defend this position against a number of recent objections. Since the Open Question Argument is the main obstacle to reductive views about normative phenomena, the door is open to be epistemic reductionists. This, unfortunately, is bad news for robust moral realists like me who might have hoped to find support from the epistemic realm, in the form of a partner-in-guilt, for robust moral realism.

Notes

1. For brief explanations of these advantages, see Heathwood (2013).
2. Another interesting defense of moral-epistemic parity can be found in Greco (2015). I lack the space to discuss Greco's arguments here, but see Côte-Bouchard (2017).
3. Furthermore, they are not themselves natural properties. This addition is required to distinguish non-naturalism from non-reductive naturalism.
4. More exactly, they should hold that the property of *basic* intrinsic goodness is necessarily coextensive with the property of being a state of pleasure (see Harman 1967 and Feldman 2000). I ignore this complication in what follows.
5. A terminological note: in what follows, since it won't matter, I won't fuss over the differences between natural and descriptive properties, and I will also sometimes use the expression "non-moral property" for the same thing. I also won't fuss over the difference between reductionism and naturalism.
6. This phrase is due to Milo (1995, 182, 190–93). See Shafer-Landau (2003, 15–17) for the same sort of definition of moral realism.
7. Provided that pleasure is stance-independent, which is in fact a controversial view. See Bramble (2013) for a stance-independent account of pleasure and Heathwood (2007) for a stance-dependent theory.
8. The Open Question Argument is due to Moore (1903). For a helpful interpretation, see Feldman (2005).
9. Incidentally, I believe that synthetic reductionism fails for other reasons: moral terms don't function the way that synthetic reductionism requires them to function. See Barnett (2002) and Gampel (1996).
10. To emphasise again: it is not my claim that this biconditional in fact has no counterexamples. Nor am I claiming that it is in no need of clarification or refinement. For example, does "likely to be true" here mean merely a likelihood of greater than 0.5, or does it mean something greater than that, or perhaps something both greater than that and vague? For another example, might we want to include something to the effect that the belief must be *based on* the evidence? Or perhaps we are talking here about propositional rather than doxastic justification, so that appeals to the basing relation are not needed? Also, what is it for something to be a part of a person's evidence (this will we discuss briefly below (Sect. 4.2)). These are all important issues, but since my project is

the metaepistemological one, I don't want to—and don't much need to—get bogged down in the controversies of first-order epistemology.

11. The main one discussed in my (2009) paper is the same analysis that is on display in Sect. 3 above. In that paper, I do allude in passing to the closely related analysis that Rowland focuses on (Heathwood 2009, 89–90).
12. I don't respond to every one of Cuneo and Kyriacou's concerns in the main text. One of these concerns is that even if I am right that *epistemic reasonability* is reducible, this does not show that all epistemic notions are reducible. This is a fair point. My reductive analysis of reasonable belief is only a first step towards showing that the entire epistemic realm can be explained naturalistically. But epistemic reasonability (in other words, epistemic justification) is, it is fair to say, *the* central normative notion in epistemology. Cuneo and Kyriacou also object that my positon commits non-naturalists in metaethics to the view "that there are not one but two types of reason properties: one that is wholly descriptive and one that is not" (Cuneo and Kyriacou 2018). This is true. But I just think that it is a straightforward consequence of the fact that epistemic facts are naturalistically reducible while moral facts are not, a position that, I have been arguing, we have good reason to accept.
13. And if it really is some epistemologist's view that, in the example, you have, as a part of your evidence, the information that you are looking at a *trumpe l'oeil* mural, then I take it that the epistemologist should say that it is in fact not reasonable for you to believe that there is a table before you. After all, you had decisive evidence to the contrary.
14. Cf. Thomas Kelly's remarks about "theoriz[ing] in the opposite direction": "to the extent that one has independent intuitions about what an individual would be justified in believing in a given scenario, such intuitions will shape one's views about what evidence must be available to an individual so situated—and therefore, one's views about the more general theoretical issue about what evidence is, or what sorts of things can and cannot qualify as evidence" (Kelly 2014, sec. 1).
15. This argument assumes that the "ought" in *evidence that a subject ought to take into account* is an epistemic "ought." If this were some other kind of "ought," such as a moral "ought," the strategy described above would not succeed (at least given, what I believe, that moral notions are not reducible to descriptive notions). In general, if the moral and the epistemic are in this way entangled, in that whether a belief is reasonable depends on moral factors, my argument for moral-epistemic disparity

would fail. For a recent argument for this kind of moral-epistemic entanglement, see Case (forthcoming).

16. The idea that an ideal observer analysis will fall prey to the Open Question Argument if and only if it is understood reductively is recognised by advocates of the Open Question Argument in ethics, such as, for example, Michael Huemer: "I think that there are two ways of understanding this sort of theory [the ideal observer theory], one that makes it non-reductionist and immune to the Open Question Argument, and another that makes it reductionist but vulnerable to the Open Question Argument" (Huemer 2005, 68).
17. I should emphasise that Cuneo and Kyriacou don't mean the argument that I discuss in this section to stand on its own. It is part of a larger line of argument, which includes the part that I address in the next section.

References

Barnett, David. 2002. Against a Posteriori Moral Naturalism. *Philosophical Studies* 107 (3): 239–257.

Bramble, Ben. 2013. The Distinctive Feeling Theory of Pleasure. *Philosophical Studies* 162 (2): 201–217.

Case, Spencer. forthcoming. From Epistemic to Moral Realism. *Journal of Moral Philosophy*.

Côte-Bouchard, Charles. 2017. Epistemological Closed Questions: A Reply to Greco. *Manuscrito* 40 (4): 97–111.

Cuneo, Terence, and Christos Kyriacou. 2018. Defending the Moral/Epistemic Parity. In *Metaepistemology*, ed. Conor McHugh, Jonathan Way, and Daniel Whiting. Oxford: Oxford University Press.

Feldman, Fred. 2000. Basic Intrinsic Value. *Philosophical Studies* 99 (3): 319–346.

———. 2005. The Open Question Argument: What It Isn't; and What It Is. *Philosophical Issues* 15 (1): 22–43.

Feldman, Richard. 1988. Having Evidence. In *Philosophical Analysis*, ed. D.F. Austin, 83–104. Dordrecht: Kluwer Academic Publishers.

Firth, Roderick. 1952. Ethical Absolutism and the Ideal Observer. *Philosophy and Phenomenological Research* 12 (3): 317–145.

Gampel, Eric H. 1996. A Defense of the Autonomy of Ethics: Why Value is Not like Water. *Canadian Journal of Philosophy* 26 (2): 191–209.

Greco, Daniel. 2015. Epistemological Open Questions. *Australasian Journal of Philosophy* 93 (3): 509–523.
Harman, Gilbert. 1967. Toward a Theory of Intrinsic Value. *Journal of Philosophy* 64 (23): 792–804.
Heathwood, Chris. 2007. The Reduction of Sensory Pleasure to Desire. *Philosophical Studies* 133 (1): 23–44.
———. 2009. Moral and Epistemic Open-Question Arguments. *Philosophical Books* 50 (2): 83–98.
———. 2013. Reductionism in Ethics. In *International Encyclopedia of Ethics*. Oxford: Blackwell.
Huemer, Michael. 2005. *Ethical Intuitionism*. 2005 ed. Basingstoke: Palgrave Macmillan.
Hume, David. (1738) 1975. *A Treatise of Human Nature*. Edited by L.A. Selby-Bigge and P.H. Nidditch. 2nd ed. Oxford: Clarendon.
Kelly, Thomas. 2014. Evidence. Edited by Edward N. Zalta. *The Stanford Encyclopedia of Philosophy (Fall 2014 Edition)*. http://plato.stanford.edu/archives/fall2014/entries/evidence/.
Milo, Ronald. 1995. Contractarian Constructivism. *The Journal of Philosophy* 92 (4): 181–204.
Moore, G.E. 1903. *Principia Ethica*. Cambridge: Cambridge University Press.
Rowland, Richard. 2013. Moral Error Theory and the Argument from Epistemic Reasons. *Journal of Ethics and Social Philosophy* 7 (1): 1–24.
Shafer-Landau, Russ. 2003. *Moral Realism: A Defence*. Oxford: Clarendon.
Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press.



From *Moral* Fixed Points to *Epistemic* Fixed Points

Christos Kyriacou

1 Introduction

In a recent paper that is already gaining traction, Cuneo and Shafer-Landau (2014) argued that there are moral conceptual truths that are substantive in content, what they called "moral fixed points."[1] According to their essentialist theory of concepts (2014, 409–11), the moral fixed points are conceptual truths in virtue of the semantic relation of satisfaction necessarily obtaining among the essences of the involved constituent concepts (e.g. "Torturing kids for fun is *pro tanto* wrong").[2] That is, in the case of moral conceptual truths (i.e. "X is F") the essence of the moral concept F necessarily satisfies the essence of the subject concept X and applies to its substantive content.

As a result, it could not be the case that something is X but not F (i.e. torturing of kids for fun but not *pro tanto* wrong).[3] If some fail to acknowledge this much, they are somehow conceptually deficient with "wrong." Perhaps they are confused, or do not understand the proper meaning of "wrong" and what it implies. Thus, Cuneo and Shafer-Landau

C. Kyriacou (✉)
Department of Classics and Philosophy, University of Cyprus, Nicosia, Cyprus
e-mail: ckiriakou@gmail.com

© The Author(s) 2018
C. Kyriacou, R. McKenna (eds.), *Metaepistemology*, Palgrave Innovations in Philosophy, https://doi.org/10.1007/978-3-319-93369-6_4

(2014, 412–14, 438–40) have indicated that error-theorists (and other antirealists) that fail to accept such truths are likely to be conceptually deficient with regard to moral concepts.[4]

In this paper I argue that insofar as we have some reason to postulate *moral* fixed points on the basis of explanatory value, we have equal reason to postulate *epistemic* fixed points on the same basis. To this effect, I show that the two basic reasons Cuneo and Shafer-Landau (2014) offer in support of moral fixed points naturally carry over to epistemic fixed points. In particular, epistemic fixed points exhibit the four "marks" of conceptual truths that Cuneo and Shafer-Landau identify and can be utilised to explain and address important challenges to epistemic realism (from epistemic disagreement, epistemic supervenience, and remarkable coincidence). Moreover, postulating epistemic fixed points helps epistemic realism avoid challenges that epistemic antirealism runs into, such as that it entails epistemic self-defeat and global scepticism. I conclude that insofar as we have some reason to postulate moral fixed points on the basis of explanatory value, we have equal reason to postulate epistemic fixed points.

The paper develops as follows. In Sect. 2 I lay out the basics of the moral fixed points proposal, as I understand it, and in Sects. 3 and 4 show how the same basic reasons that count in favor of postulating moral fixed points, also count in favor of postulating epistemic fixed points. In Sect. 5 I sum up with a conclusion and some ruminations on its wider dialectical significance.

2 Cuneo's and Shafer-Landau's (2014) Moral Fixed Points: An Outline

Some of the examples of moral fixed points that Cuneo and Shafer-Landau (2014) offer are the following[5]:

- It is *pro tanto* wrong to engage in the recreational slaughter of a fellow person.
- It is *pro tanto* wrong to impose severe burdens on others simply because of their physical appearance.

- There is some moral reason to offer aid to those in distress, if such aid is very easily given and comes at very little expense.
- It is *pro tanto* wrong to humiliate others simply for pleasure.

Cuneo and Shafer-Landau offer a detailed argument for the existence of such "moral fixed points" that is constituted by three basic reasons. First, they suggest that postulating moral fixed points explains what they identify as four "marks" of conceptual truths: *apriority*, necessity, denial evokes bewilderment to competent speakers and framework status (2014, 407–08).[6] So, if there are any conceptual truths at all, there could be moral conceptual truths as well. Second, they suggest that postulating moral fixed points helps address three long-standing challenges to moral realism: moral disagreement, moral supervenience and the objection from remarkable coincidence (2014, 422–31). Third, they argue that moral fixed points are resistant to a battery of objections that can be run against it, such as from conceptual deficiency, the open question argument and the "unconnected heap of duties" objection (2014, 431–40).

On the basis of the aforementioned three reasons, they conclude that the notable *overall* explanatory value of the moral fixed points supports their existence. Of course, their argument is not meant to be conclusive, not even perhaps an inference to the best explanation of the various data they adduce.[7] Rather, they think that the three reasons they put forth jointly make a reasonable case for the moral fixed points proposal. In what follows, due to space restrictions, I show that the first two basic reasons in support of the moral fixed points naturally carry over to the epistemic case and, therefore, support epistemic fixed points as well. Unfortunately, discussion of the third reason, namely, how epistemic fixed points could be resistant to various objections (parallel to those run against the moral fixed points proposal) cannot be pursued here. It is, however, relatively easy for anyone to understand how this third reason could also be explored.

Let us set out to present the first reason constitutive of Cuneo's and Shafer-Landau's argument, namely, the four "marks" of conceptual truths that moral fixed points also exhibit. According to their proposal, conceptual truths (such as "actuality entails possibility") are epistemologically *a priori*, that is, justified without any direct appeal to experience.

This *a priori* justification is justification that is conferred in virtue of the essences of the constituent concepts. Understanding "actuality," "entails," and "possibility," is sufficient for competent speakers to understand that "actuality entails possibility"—at least to those initiated into the meaning of these quasi-technical concepts. Conceptual truths are also metaphysically necessary in the sense that they could not have been otherwise in some other possible world. Given the essences of "actuality," "entails" and "possibility," it could not have been the case that "actuality does not entail possibility."

In addition, psychologically speaking, the denial of conceptual truths tends to evoke bewilderment to competent speakers. If you try to explain to an otherwise competent speaker—initiated into these quasi-technical concepts—that "actuality does not entail possibility," it is likely that you would receive a stare of incredibility and bewilderment in return.[8] Finally, conceptual truths seem to have framework-status in the sense that they delimit what is permissible or appropriate concept application (e.g. we should not assert that heart transplant—which is already actual—is not possible). Thus, conceptual truths put in place standards in virtue of which someone can properly apply concepts, such as actuality and possibility.

These four marks of conceptual truths are, of course, not conclusive evidence for a proposition being a conceptual truth, but insofar as there are any conceptual truths, the marks are indicative evidence of a proposition being a conceptual truth. They are also rather uncontroversial marks of *prima facie* conceptual truths, presumably acceptable even by Quineans who are prone to deny the reality of any *ultima facie* conceptual truths and would proceed to debunk these "marks" as ultimately misguiding evidence.[9] Quinean scepticism about conceptual truths aside, however, moral fixed points seem to exhibit these four marks of conceptual truths and this hints that insofar as there are any conceptual truths, there could be moral conceptual truths as well. Let us elaborate the point.

According to the Cuneo and Shafer-Landau broadly Fregean, essentialist theory of concepts, the moral fixed points are *a priori* known conceptual truths. That is, in the case of moral conceptual truths (i.e. "X is F") the essence of the moral concept F necessarily satisfies the essence of the subject X and applies to its substantive content. It could not be the

case, metaphysically speaking, that something is X but not F (i.e. torturing of kids for fun but not *pro tanto* wrong). This much of the story explains the exemplification of the marks of *apriority* and necessity by alleged moral fixed points.

Thus, Cuneo and Shafer-Landau (2014, 412–14, 438–40) have indicated that error-theorists that fail to accept such truths are likely to be conceptually deficient with regard to moral concepts.[10] This explains the denial evokes bewilderment mark. Moral fixed points also exhibit framework-constitutive status in the sense that they delimit what is right and wrong, permissible and impermissible from the moral point of view (and concept application). If for example "torturing kids for fun is *pro tanto* wrong" is a moral fixed point, it is morally wrong to act accordingly and, therefore, it is impermissible. This explains the framework status mark. So, moral fixed points exhibit the identified four marks of conceptual truths and, therefore, have as much claim to be conceptual truths as any other.

We can now turn to the second constitutive reason of the Cuneo and Shafer-Landau argument, namely, that the proposal helps address important challenges to moral realism: from moral disagreement, remarkable coincidence and moral supervenience. Very briefly, according to the epistemological version of the challenge from moral disagreement, which is also the more acute version of the challenge, the extent of deep and persistent moral disagreement between epistemic peers *pro tanto* justifies the conclusion that there are no justified moral beliefs. If there could be justified moral beliefs, deep and persistent disagreement would at least gradually give way to convergence, but this is not observed to be happening. So, there can be no robust moral epistemology and moral realism, which is customarily committed to such a moral epistemology, is implausible.

Cuneo and Shafer-Landau (2014, 423) respond, first, that there is more agreement than the argument from disagreement allows and that, indeed, moral fixed points are almost universally accepted. Second, they point out that if there are moral fixed points, then this entails that those who fail to grasp them are not sufficiently competent in moral thinking and, therefore, are not epistemic peers of the realists (despite appearances). In other words, they are conceptually deficient with regard to moral concepts and what they imply. As they note (2014, 424), conceptual defi-

ciency might be "the cumulative effect of bias, prejudice, various kinds of irrationality, factual ignorance or limited imagination" (2014, 424). Hence, moral disagreement does not *pro tanto* justify moral scepticism and, hence, does not undermine the reality of moral knowledge (and moral realism).

Cuneo and Shafer-Landau (2014, 425–28) also argue that the argument from remarkable coincidence can be addressed by their proposal. Very briefly, the argument is usually propounded by genealogical debunkers (evolutionary, cultural and others) such as Harman (1975) and Street (2006) and is roughly the following. If we had a different genealogical history, there could have been very different moral beliefs considered as truths because there is a multitude of possible moral systems. Given this possibility, it would be remarkable coincidence if our moral beliefs came to reliably track and represent moral facts of the correct moral system (even if there are such facts and system).

According to Cuneo and Shafer-Landau, the argument rests on a false premise that the moral fixed points proposal exposes. The false premise is the premise that there are no "contentful conceptual constraints" (2014, 420) on what can count as a moral system (for beings like us, in worlds like us). On the contrary, the moral fixed points proposal suggests that there are contentful conceptual constraints on what can count as a moral system, which is, in effect, to deny the argument from remarkable coincidence the required premise to take off in the first place. Thus, the moral fixed points proposal addresses the remarkable coincidence objection.

The third challenge is the one from moral supervenience—what they call the "Humean Challenge"—advocated by many anti-realists in the past and recently restated in a crisp way by McPherson (2012). Very briefly again, various anti-realists have argued that non-naturalists owe us an explanation of how irreducible moral properties supervene on non-moral, natural properties. Non-naturalists usually either tacitly assumed, or explicitly appealed, to brute connections between the moral and the natural, which the antirealists thought that in reality explain nothing and leaves us with a metaphysical mystery to be dispelled.[11]

McPherson expresses aptly this metaphysical concern about non-naturalism with what he calls the "Modest Humean Thesis": "Commitment to brute necessary connections between discontinuous properties count

significantly against a view" (2012, 217). McPherson takes "discontinuous properties," to mean that any two properties, A and B, that are neither reducible to one another, identical to one another, nor metaphysically continuous with each other are discontinuous. Given that non-naturalism is indeed committed to such brute necessary connections between discontinuous moral and natural properties, moral supervenience is typically considered by anti-realists to count seriously against the plausibility of any form of non-naturalism.

Although Cuneo and Shafer-Landau make clear that they are "deeply suspicious" of the Modest Humean constraint, they suggest that moral fixed points can address it and set the moral supervenience challenge to rest (2014, 429).[12] As they explain, moral fixed points suggest that there are necessary connections between the essences of moral and natural concepts and if there are such connections then the supervenience challenge for non-naturalism has a relatively easy answer. There are necessary moral fixed points that are part and parcel of the fabric of reality and these fixed points by their very nature ground necessary connections between moral and natural concepts and their essences.[13] So there are no brute, inexplicable connections between the moral and the natural because the necessary connection is explicable by means of conceptual necessity. That is, given moral fixed points, necessary connections between moral and natural essences exist, the supervenience requirement is satisfied and the mystery is dispelled.

All in all, the moral fixed points proposal indicates that the long-standing challenges from moral disagreement, moral supervenience and remarkable coincidence fail to make contact with non-naturalist realism and are therefore, if not dialectically harmless, at least not insurmountable for the non-naturalist. Given the prominence of these three antirealist arguments against non-naturalism, this seems a significant explanatory payoff of the proposal.

With the sketch of the two basic reasons that Cuneo and Shafer-Landau develop in support of the moral fixed points at hand, let us now show how we could—by parity of reasoning—support epistemic fixed points as well.

3 From Moral to Epistemic Fixed Points: A Parity Argument

The contention of the paper is that the argument for the moral fixed points proposal could, in principle, be extended to support epistemic fixed points as well—and this is, perhaps, not surprising given the well-attested normative analogues between the moral and epistemic domains (e.g. categorical reasons, objectivity, motivation etc.).[14] That is, by parity of reasoning, we could postulate epistemic fixed points that exhibit the identified four "marks" of conceptual truths, such as *apriority*, necessity, denial evokes bewilderment and framework status, and we can use the proposal to address respective challenges to epistemic realism, namely, from epistemic disagreement, epistemic supervenience and remarkable coincidence. But first a list of candidate epistemic fixed points:

- Knowledge entails truth.
- Knowledge entails epistemic justification (of some kind).
- Justified true belief does not entail knowledge.
- It is *pro tanto* epistemically rational to believe what is best epistemically justified, given relevant evidence at one's possession.
- It is *pro tanto* epistemically irrational to believe what is epistemically unjustified, given relevant evidence at one's possession.
- It is *pro tanto* epistemically irrational to believe what is based on wishful thinking, self-deception, doxastic akrasia, or other sorts of epistemic irrationality.
- Epistemically rational belief entails justification (of some kind).
- Epistemically rational belief entails *pro tanto* aiming at the truth/avoiding falsity.
- Epistemic justification (of some kind) is, *ceteris paribus*, truth-conducive.
- Understanding-why is *pro tanto* epistemically valuable.
- Knowledge is *pro tanto* more epistemically valuable than true belief.
- It is *pro tanto* epistemically irrational to misinform others for fun (in full awareness of the fact).

- It is *pro tanto* epistemically irrational to rely on unreliable sources of information (in full awareness of the fact).

As with the moral fixed points, a case can be made that epistemic fixed points exhibit the four marks of conceptual truths: they are *a priori* justified, metaphysically necessary, psychologically their denial tends to evoke bewilderment in competent speakers and they have framework-constitutive status for epistemic discourse and practice.

Take for example the candidate fixed point that "knowledge entails truth," which is almost universally accepted by epistemologists as a necessary truth.[15] It is, arguably, *a priori* justified in virtue of the essences of the constituent concepts ("knowledge," "entails," "truth"). It is necessary in the sense that it could not have been otherwise (for rational beings like us) in a different possible world (like ours), its denial evokes bewilderment in competent speakers—at least to those initiated to the quasi-technical philosophical concepts involved—and it is framework-constitutive of "knows" concept application (and appropriate knowledge thought and talk).

The framework-constitutive character of the epistemic fixed points is an especially important mark of *prima facie* conceptual truths because it indicates that they cannot be rationally denied. If they cannot be rationally denied and it is *pro tanto* epistemically rational to believe what is best epistemically justified, given relevant evidence at one's possession—which is yet another candidate epistemic fixed point in our list—then we have further support for epistemic fixed points. To illustrate, take the factivity condition of knowledge, that is, "knowledge entails truth." It could be an epistemic fixed point because, as we have seen, it exhibits the four marks of conceptual truths, and it seems irrational to deny it (and it is unclear what a rationally binding, conceptual truth it is, if it is not a fact of sorts).[16] This can be argued for by means of a *reductio ad absurdum* that indicates that "knowledge entails truth" is a prima facie necessary and indispensable epistemic fixed point.[17]

Suppose we assume the popular knowledge norm of assertion: we should, *ceteris paribus*, assert that *p* only if we know that *p*.[18] For the sake of clarity, I mark the linguistically implicit operation of the knowledge norm of assertion with "(I know that)." We may understand the linguisti-

cally implicit operation of the knowledge norm of assertion as a conveyed implicature that takes the asserted proposition as a complement. For example, the assertion (I know that) "Mary is from Bristol" says that "Mary is from Bristol" but conveys via implicature (and means) that "I know that Mary is from Bristol."

Suppose now we assert that (I know that) "knowledge does not entail truth." It seems that we are committed to implicit epistemic self-defeat. For if the assertion is true, then intuitively we will know that is true that knowledge does not entail truth and therefore, by the same lights it will not be true that we know that knowledge does not entail truth (because, intuitively, knowledge requires truth). Hence, we land in epistemic self-defeat. Therefore, we cannot rationally assert (to know) that "knowledge does not entail truth" because intuitively the knowledge assertion itself entails truth and the upshot is epistemic self-defeat.[19]

But for argument's sake, let us suppose for a moment that "knowledge does not entail truth" and forget about epistemic self-defeat. Two other unpalatable implications follow. First, the supposition would allow that we can *really* know falsehoods, which is very counterintuitive. I cannot be coherently understood to (really) know that Hillary Clinton is the new president of the US, unless I am insincere, using the concept loosely, making a pun or have an implausibly revisionary, perhaps oxymoronic, understanding of "know."

Second, it seems that nothing can be really known if the factivity condition is not satisfied and, therefore, it seems that we are saddled with global scepticism about knowledge if the factivity condition is denied. This is the case because we do not *really* know of what is non-true, that is, false. And if nothing can be known, global scepticism about knowledge follows, which seems again the wrong result because, on the one hand, it is epistemically self-defeating and, on the other hand, at least some necessary-Cartesian propositions are intuitively known. On the former, it is epistemically self-defeating to assert (that I know) that there is no knowledge. For, if there is no knowledge we have good reason to deny that we know that there is no knowledge and thereby good reason to retract the assertion of the claim. On the latter, I can provisionally doubt all knowledge, but—on pain of epistemic self-defeat—I cannot doubt

that at least some thinking is necessary for such doubting. So, I can know that thinking is necessary for doubting.[20]

The above *reductio* suggests that the assertion "knowledge does not entail truth" is necessarily false and, therefore, by deduction its negation "knowledge entails truth" is necessarily true. Moreover, that it is a necessary truth that "knowledge entails truth" can be supported by the fact that we cannot even universally *conditionalise* knowledge (à la Evers and Streumer 2016) without presupposing *some* knowledge. Let us explain. It seems that we can assert without incoherence that (I know that) "*If anything is knowledge*, knowledge entails truth." This is the case because if true the assertion is an instance of knowledge, satisfies the antecedent and therefore entails the consequent, namely, that knowledge entails truth. Yet this presupposes that there is a robust knowledge relation (that entails truth) and that we have at least some knowledge, namely, of the conditional itself.

What is more, it seems that we cannot coherently assert the denial of "*If anything is knowledge*, knowledge entails truth." We cannot rationally assert (that I know) its denial without falling into another *reductio*. For, if we assert (I know that) "It is not the case that, *if anything is knowledge*, knowledge entails truth," this would entail that (I know that) at least some instances of knowledge do not entail truth (which as we have seen is very counterintuitive). But if we assume for the sake of the argument that some instances of knowledge do not entail truth, then we do not really know that "It is not the case that, *if anything is knowledge*, knowledge entails truth" because it could itself be one of the cases in which knowledge does not entail truth. And even if it is one of those cases, according to the factivity denier we will "know" something without being true, which seems very counterintuitive.

These difficulties can be easily resolved if we assume that it is an epistemic fixed point that "knowledge entails truth." Given that we easily resolve complex difficulties with such a simple concession, it seems that we have good reason to accept that there is at least one epistemic fixed point (and by parity of *reductio* reasoning, there could of course be more). So, it seems plausible that there are at least some epistemic fixed points and we can be minimally reliable in knowing them, although no doubt we are far from infallible.

The above illustration indicates that epistemic fixed points are, arguably, framework-constitutive principles that are necessary and indispensable for epistemic discourse and reasoning.[21] If anyone plays the epistemic game, then she has to abide by these principles, if she is to play the *real* game at all. This is the case because the epistemic fixed points are principles that need to be presupposed as necessary if any proper epistemic reasoning and inquiry is to take place. They are *indispensable* principles of epistemic reasoning.

The *reductio* argument raises a number of concerns and I can afford to discuss only two here. An obvious first objection to the *reductio* argument is that it is of limited scope. That is, even if the *reductio* argument for epistemic fixed points works, it does not apply to the case of moral fixed points. This is the case because denying moral fixed points is not epistemically self-defeating in the same way that denying the factivity condition is. Perhaps we cannot rationally assert the negation of some epistemic fixed points and this indicates that they are indispensable for theoretical reasoning. But this does not seem the case with moral fixed points because *prima facie* I can coherently, without any obvious epistemic self-defeat, deny that "It is *pro tanto* wrong to torture kids for fun." If this is right, the extension of the mark of framework-constitutive status from moral to epistemic fixed points backfires because it shows that it does not apply to the moral case in the same way.

Very briefly, the rejoinder could be that it is due to conceptual deficiency that it is not acknowledged that such a denial breaks the standards constitutive of morality, namely, the moral fixed points. It might be true that there is no equivalent epistemic self-defeat but this is because of the specifically *epistemic* nature of epistemic fixed points. Surely, we cannot rationally deny the standards constitutive and indispensable for any theoretical-epistemic reasoning (moral, scientific or other).[22] But equally we cannot rationally deny the moral standards indispensable for practical-moral reasoning, although there is not the same kind of epistemic self-defeat. The difference in nature of the two domains explains away the intuition about the difference of the moral/epistemic fixed points in regard to rational deniability, without defeating the parity in terms of framework status between moral and epistemic fixed points.

A second objection is that the presented *reductio*-style, indispensability argument does not seem to extend and apply to all candidate epistemic fixed points in the suggested list. Perhaps it applies to some fixed points such as the factivity condition but not to others, such as that "It is *pro tanto* epistemically irrational to misinform others for fun (in full awareness of the fact)." Besides, some might advocate some form of epistemic egoism, where we should promote personal truth-acquisition and not interpersonal truth-acquisition. Perhaps we have reason to promote our own epistemic good, but have no reason to promote the epistemic good of others. If so, we can rationally deny, without incoherence, that "It is *pro tanto* epistemically irrational to misinform others for fun (in full awareness of the fact)."

The objection raises questions we cannot delve into here, but I would concede that the *reductio* argument seems to apply more clearly to some candidate epistemic fixed points than others. This concession would suffice for current purposes, which is to uphold the case for at least *some* epistemic fixed points vis-à-vis the Cuneo/Shafer-Landau case for some moral fixed points. Perhaps such concerns could be assuaged, but the scope of the paper does not allow for such an extended digression here.

Let us now move on to the second reason in support of epistemic fixed points, namely, what explanatory payoffs they might yield. Such a payoff would add more support to the epistemic fixed points proposal.

4 Explanatory Payoffs: Disagreement, Remarkable Coincidence, Supervenience

As we have seen, Cuneo and Shafer-Landau use moral fixed points to address three long-standing challenges to moral realism, moral disagreement, moral supervenience and remarkable coincidence (2014, 422–31). The same style of explanatory argument could be used in support of epistemic fixed points (vis-à-vis the epistemic version of these challenges).

First, it could be used to deflect an argument from epistemic disagreement against epistemic knowledge (and realism).[23] Like in the moral case, we could point out that just because there is deep and persistent epis-

temic disagreement (e.g. about climate change, the causes of the recent financial crisis, or even about epistemic normativity) it does not follow that there are no epistemic facts and epistemic knowledge. Deep and persistent disagreement might be even widespread, but this need not defeat epistemic knowledge (and epistemic realism). As in the case of moral disagreement, at first we should note the near universal acceptance among experts of the necessary status of at least some epistemic fixed points, such as the factivity condition. *A fortiori*, any residual disagreement might be due to conceptual deficiency in regard to epistemic concepts and it might be "the cumulative effect" of cognition-inhibiting factors such as confusion, bias, factual error, vice, unsound arguments et cetera. In conclusion, deep and persistent disagreement need not undermine epistemic knowledge (and realism) and epistemic fixed points help us explain how this might be the case.

Second, we can use epistemic fixed points to address remarkable coincidence with respect to epistemic facts and knowledge. The epistemic version of the remarkable coincidence argument runs as follows. If we had a different genealogical history, there could have been very different epistemic beliefs considered as truths because there is a multitude of possible epistemic systems. Given this possibility, it would be remarkable coincidence if our epistemic beliefs came to reliably track and represent epistemic facts of the correct epistemic system (even if there are such facts and systems).

The response is to deny that there is theoretical space for such wide-ranging coincidence because it seems that it could not have been the case that very different epistemic beliefs and systems could have been rationally considered as truths and the correct epistemic system (in worlds like ours, for beings like us).[24] There are "contentful conceptual constraints" on what can count as an epistemic system and via conceptual means alone we can have access to such conceptual constraints-truths (that are framework-constitutive of epistemic rationality). As the *reductio* argument with the factivity condition has indicated, there are limits to what can be an epistemic system because there are contentful conceptual constraints that such a system must satisfy.

A fortiori, epistemic fixed points can help us avoid related problems with the remarkable coincidence argument that epistemic anti-realism

faces, such as the so-called "self-debunking problem" for evolutionary debunking (and other genealogical debunking) and the global skepticism that inexorably ensues from such self-debunking.[25] The self-debunking problem is the problem that the standard evolutionary debunking of normative beliefs as unjustified seems to debunk epistemic beliefs, which are also arguably normative, and this leaves genealogical debunking arguments against normative beliefs epistemically self-defeating.[26] It leaves genealogical arguments epistemically self-defeating because if there are no justified beliefs *at all*, then genealogical arguments are themselves unjustified and we have no good reason to endorse them. And of course if there are no justified beliefs at all, then global scepticism about knowledge follows.

But if there are epistemic fixed points in virtue of which genealogical arguments can be run, then there is no self-debunking problem for genealogical arguments. There are epistemic facts-standards in virtue of which debunking arguments can be launched (without committal to epistemic self-defeat) and these are the epistemic fixed points. Of course, postulating epistemic fixed points undercuts the aspirations of anti-realists to wholesale debunk normative beliefs as unjustified, but this is a welcome implication for the normative realist. We can still, however, use genealogical reasoning to debunk some beliefs as unjustified (perhaps, as some have suggested, thermoreceptive beliefs and positive illusions beliefs)[27] but not just any beliefs (such as epistemic or moral fixed points beliefs) and this also helps with a noted *demarcation problem* for evolutionary debunking arguments.

This is the problem of asking in virtue of what regulative metaepistemic norm evolutionary considerations render a belief justified or debunk it as unjustified.[28] The problem is serious because it seems that, in the absence of a regulative metaepistemic norm, any appeal to evolutionary considerations in order to justify or debunk beliefs would be ad hoc. That is, it would be ad hoc because it could always be asked why we do not appeal to the same evolutionary considerations in order to justify instead of debunk some beliefs (and vice versa). Insofar as evolutionary arguments are sometimes used to debunk, and other times to justify beliefs, the adhocness concern seems potent and reasonable.

The epistemic fixed points proposal indicates that debunking arguments cannot indiscriminately apply to normative epistemic beliefs because there are at least some indispensable epistemic beliefs-truths (i.e. epistemic fixed points) for any theoretical reasoning, genealogical reasoning included. Moreover, no global scepticism about knowledge ensues from the self-debunking problem of genealogical debunking because access to basic epistemic truths, such as the fixed point of the factivity condition, is established via conceptual means alone. So, there is at least some knowledge (such as of the factivity condition) and global scepticism about knowledge is false (or at least implausible).

Third, epistemic fixed points can also help us with an account of epistemic supervenience. Although epistemic supervenience has not been as much explored as an argument against epistemic realism (as its moral counterpart has), a prospective supervenience argument against non-naturalist epistemic realism could be run.[29] The argument would proceed in the usual way. It would challenge the non-naturalist to explain how epistemic facts and properties supervene on natural facts and properties, given that epistemic and natural properties seem to be, in McPherson's words, "metaphysically discontinuous."

On first sight, it would seem that the only available answer to the non-naturalist is to bite the bullet and accept that the supervenience relation is brute and *sui generis* and no further explanation is to be asked, which is only to play into the hands of the anti-realists because it leave us with a metaphysical mystery to be dispelled. In essence, the challenge boils down to a dilemma: to explain the metaphysical connection between the epistemic and the natural in a way that does not posit metaphysically mysterious brute connections or explain why brute connections are not particularly mysterious (or at any rate mysterious, but to be countenanced).

As in the moral case, we could posit necessary connections between the epistemic and the natural set up by the essences of the involved epistemic concepts, but these need not be particularly brute and inexplicable. Necessary connections might even exist between two normative-epistemic concepts as well as between a normative and a natural concept. In the case of the factivity condition, a necessary connection exists between two normative concepts, knowledge and truth. That is, the essence of truth necessarily applies to the essence of knowledge and therefore truth

supervenes on knowledge. Knowledge is sufficient for truth and necessitates it and, therefore, for any proposition *p*, if *p* is known by some S, then *p* is true. The factivity condition is a case where two *prima facie* non-natural, normative concepts are involved in an epistemic supervenience relation.

But there could also be epistemic supervenience between epistemic and natural concepts-essences. Such a case is the candidate fixed point in our list "It is *pro tanto* epistemically irrational to believe what is produced by wishful thinking." Epistemic (ir)rationality is a *prima facie* non-natural, epistemic concept but wishful thinking seems to be a natural, psychological concept that individuates a belief-forming process. This is, roughly, the process of forming beliefs on the basis of our mere preferences and wishes (and not on the basis of given evidence). The respective fixed point indicates that any belief produced by the process of wishful thinking is sufficient for epistemic irrationality and necessitates it because the essence of "epistemic irrationality" necessarily applies to the essence of "beliefs produced by wishful thinking." In other words, "epistemic irrationality" supervenes on "beliefs produced by wishful thinking" in virtue of conceptual necessity.

Let us now sum up with some ruminations on the wider dialectical significance of the *prima facie* normative parity between moral and epistemic fixed points.

5 Conclusion

I have argued that insofar as we have some reason to postulate moral fixed points, we have equal reason to postulate epistemic fixed points. This conclusion indicates that moral and epistemic fixed points might be "on a par" and, therefore, the respective normative "fixed points" proposals should stand or fall together. This implies that anyone who wishes to deny moral fixed points should also—by parity of reasoning—deny epistemic fixed points and it is surely much harder to deny *both* proposals than to deny only moral fixed points.

For one thing, as our discussion of the factivity condition has shown, anyone inclined to deny the factivity condition verges on rational incoherence. The result is of dialectical importance because the *prima facie*

parity between moral and epistemic fixed points further strengthens the prospect for the *overall* parity between the moral and the epistemic, which is an important premise in companions-in-guilt arguments in support of moral (and, inversely, epistemic) realism. Hence, anyone who wishes to undermine the moral-epistemic parity in general, should also venture to deny the moral-epistemic fixed points parity in particular, which is not to presume that it is either unassailable or easily undermined.[30]

Notes

1. See Ingram (2015, forthcoming), Killoren (2016), Evers and Streumer (2016), Copp (2018) for critical reactions and Kyriacou (n.d, forthcoming-b, 2017b) for a sympathiser to the moral fixed points proposal.
2. Of note, is that Cuneo and Shafer-Landau (2014, 410–11) distinguish between *conceptual* truths and *analytic* truths. Conceptual truths are propositions that hold in virtue of the essences of the constituent concepts and need not have been hitherto discovered and linguistically expressed. Analytic truths are sentences that hold in virtue of the meaning of constituent words and have been hitherto discovered and linguistically expressed.
3. The view of Cuneo and Shafer-Landau (2014) has affinities with the moderate rationalist theory of *a priori* justification that Bonjour (1997) has defended. Compare Bonjour: "The sentence in question is necessarily true because it expresses a necessary relation between certain properties, and it is of course in virtue of its meaning that it does this" (1997, 102). Another view of a *priori moral* justification can be found in Huemer (2005, chap. 5) and Swinburne (2015). For a proponent of both *a priori* justification and *a priori* moral justification, see Peacocke (2003, chaps. 6 and 7). For a general defense of *a priori* intuition, see Bealer (1998) and for *a priori* moral intuition, Audi (2015).
4. See Ingram (2015) for an argument that the conceptual deficiency charge backfires and defeats the moral fixed points proposal and Kyriacou (2017b) for a reply on behalf of the moral fixed points proposal. Ingram (forthcoming) offers a rejoinder to Kyriacou (2017b). For Kyriacou's re-reply see n.d.
5. Cuneo and Shafer-Landau (2014, 404) make clear that moral fixed points are truths that apply to rational beings like us, in worlds like us. They are not supposed to be absolute truths in any unrestricted sense.

6. This is not to suggest that we could not in principle explain these marks in a way more friendly to the antirealist. See Evers and Streumer (2016, 4 fn. 9) for how such an explanation could go.
7. Cuneo and Shafer-Landau (2014, 403) distinguish between two non-naturalist stories that could metaphysically invest the moral fixed points proposal: *minimal* and *robust* moral non-naturalism. Minimal moral non-naturalism asserts that "there are nonnatural moral truths, but there are no nonnatural moral properties or facts. All moral properties and facts are natural." Robust moral non-naturalism is ontologically more permissive and asserts that "there are both nonnatural moral truths and nonnatural moral properties and facts." They suggest that the moral fixed points proposal is compatible with both positions and that, although they "want to remain officially neutral on which of these views is true," they favour a view of the robust style (at least something "in the spirit …if perhaps not the letter" of such a view, as they put it). As nothing of substance hinges on the distinction in this paper, I remain neutral in regard to which of the two is the more promising.
8. Moral (and other) fixed points need not be very obvious at first sight and might bear different degrees of self-evidence (cf. Cuneo and Shafer-Landau 2014, 413–14). Compare, for instance, the quite evident "knowledge requires truth" with Gettier's less evident lesson, namely, "justified true belief is insufficient for knowledge."
9. See Juhl and Loomis (2009) for a round introduction to the Carnap-Quine debate on analyticity as well as the debate that ensued after Quine's seminal work. Traditional conceptual analysis seems also to be making a comeback in recent debates, see Jackson (1998), Juhl and Loomis (2009) and McGinn (2012) for some defences.
10. Kyriacou (2017b, forthcoming-b, n.d.) distinguishes between conceptual and meta-conceptual deficiency and submits that error theorists are likely to be only meta-conceptually deficient, if deficient at all. For simplicity's sake, I will ignore the distinction for the purposes of this paper.
11. For some discussion of the supervenience challenge to moral non-naturalism, see Ridge (2014, sec. 6).
12. For one thing, as Cuneo and Shafer-Landau (2014, 429 fn. 66) note, Wielenberg (2014, chap. 3) has stressed that the Modest Humean Thesis appears self-defeating because it posits itself a brute connection between the discontinuous properties "entailing a brute necessary connection between discontinuous properties" and "being unreasonable to believe."

13. While formulations of the moral supervenience challenge are often phrased in terms of explaining the connection between moral and natural *properties*, the moral fixed points proposal explains moral supervenience in terms of moral and natural *concepts-essences*. As far as I can see, nothing substantial hinges on this difference because the supervenience challenge calls for an explanation of the metaphysical "bridging" between the moral and the natural domain, be it in terms of properties or essences.
14. See Kim (1988), Chrisman (2007), Cuneo (2007), Kyriacou (2012, 2016a, b), Rowland (2013) and Cuneo and Kyriacou (forthcoming) for the moral-epistemic parity.
15. Compare Stanley: "The factivity of knowledge is not just a truth about knowledge, but a necessary truth about knowledge. That is, not only is it the case that x's knowing that *p* entails that *p*, but it is necessarily true that if someone knows that *p*, then *p*." (2005, 112) Hazlett (2010) has argued that the *ordinary* concept of knowledge is non-factive, although he concedes that epistemologists are probably right that the concept of knowledge is factive. For a rejoinder that defends that even the ordinary concept of knowledge is factive, see Hannon (2013). At any rate, the factivity condition is here intended to apply to the concept of knowledge.
16. See Schaffer (2009, 375) for the intuitive grounding of truth in fact.
17. For an appeal to an indispensability argument from practical deliberation for the existence of moral facts, see Enoch (2011, chap. 3). In essence, we present a parallel indispensability argument from theoretical deliberation-reasoning for the existence of epistemic facts. Interestingly, if knowledge is also the norm for practical reasoning, as some think (e.g. Hawthorne 2004; Stanley 2005; Williamson 2000), then the indispensability argument from theoretical deliberation for the existence of epistemic facts is by extension an indispensability argument from practical deliberation.
18. See Hawthorne (2004), Stanley (2005) and Williamson (2000) for the knowledge norm of assertion.
19. We should distinguish between *logical* self-defeat that implies contradiction and rebutting (e.g. the self-referential semantic paradoxes) and *epistemic* self-defeat that implies undercutting defeat (e.g. Descartes' cogito). Both sorts of self-defeat provide us with good epistemic reason to reject propositions or arguments. See Fumerton (1995, 43–53) for some discussion of the distinction between epistemic and logical self-defeat.

20. See Rowland (2013) for a similar argument in a different dialectical context.
21. For the idea that there might be epistemic facts that are indispensable for epistemic reasoning, see Cuneo (2007, 229).
22. Or, maybe, we can and we should? Besides, why not be epistemic fictionalists and proceed "as if" there are such indispensable fixed points while accept that these are only epistemically useful fiction? Epistemic fictionalism raises delicate issues that we cannot pursue here but one concern for the fictionalist proposal is that normative, epistemic authority does not seem to be merely fictional. But if there is epistemic authority there must also be epistemic facts in virtue of which the authority exists and binds us. Otherwise, epistemology turns out to be relative to one's fictionalist commitments (and these might differ widely) and is reduced to *schm*epistemology (see Kyriacou 2016a for some discussion). Unfortunately, we have to cut off the topic here. For some critical discussion of epistemic fictionalism see Cuneo (2007, chap. 4).
23. The epistemology of disagreement literature has recently paid attention to the possible skeptical implications of the "equal weight view" about peer disagreement. See Matheson and Carey (2013) for some discussion. For some more work on epistemology of disagreement see the essays in Feldman and Warfield (2010). Also, see Ahlstrom-Vij (forthcoming, sec. 5) for an appeal to disagreement about epistemic normativity in support of epistemic anti-realism.
24. For the moral analogue of this argument against evolutionary genealogical debunking, see Wielenberg (2014, chap. 4).
25. For the self-debunking problem for genealogical debunkers see Kyriacou (2016a, 2017a, forthcoming-a), Pust (2001) and Vavova (2014). For arguments from epistemic self-defeat run against epistemic anti-realism, see Boghossian (2006), Cuneo (2007), Lynch (2009) and Rowland (2013).
26. The obvious response that epistemic beliefs are non-normative is not easy to come by. See Heathwood (2009) for an argument that epistemic reasonability beliefs are non-normative and Cuneo and Kyriacou (forthcoming) and Kyriacou (forthcoming-a, sec. 4) for a response. For a rejoinder, see Heathwood (2018 in Chap. 3).
27. See De Cruz et al. (2011) for a possible debunking of thermoreceptive beliefs and McKay and Dennett (2009, 532) for a possible debunking of positive illusion beliefs.

28. For versions of this problem see the discussion in Bradie (1990, 35–36), Kahane (2011), Shafer-Landau (2012, 35), Woods (2018, sec. 4). Shafer-Landau (2012), in particular, makes explicit that evolutionary debunking arguments about a philosophical domain quickly over-generalise to domains that seem beyond serious epistemological doubt and, therefore, we need to disambiguate the metaepistemic norm in virtue of which debunking arguments run and confer unjustifiedness.
29. For some general discussion of epistemic supervenience, see Kyriacou (2016b, sec. 3) and Turri (2010).
30. I would like to thank David Enoch for helpful comments.

References

Ahlstrom-Vij, Kristoffer. forthcoming. The Costs of Epistemic Realism. In *Metaepistemology*, ed. Conor McHugh, Jonathan Way, and Daniel Whiting. Oxford: Oxford University Press.

Audi, Robert. 2015. Intuition and Its Place in Ethics. *Journal of the American Philosophical Association* 1 (1): 57–77.

Bealer, George. 1998. Intuition and the Autonomy of Philosophy. In *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, ed. Michael DePaul and William Ramsey, 201–240. Lanham, MD: Rowman & Littlefield.

Boghossian, Paul. 2006. *Fear of Knowledge: Against Relativism and Constructivism*. Oxford: Oxford University Press.

BonJour, Laurence. 1997. *In Defense of Pure Reason: A Rationalist Account of a Priori Justification*. Cambridge: Cambridge University Press.

Bradie, Michael. 1990. Should Epistemologists Take Darwin Seriously? In *Evolution, Cognition and Realism*, ed. Nicholas Rescher, 33–38. Lanham, MD: University Press of America.

Chrisman, Matthew. 2007. From Epistemic Contextualism to Epistemic Expressivism. *Philosophical Studies* 135 (2): 225–254.

Copp, David. 2018. Are there Substantive Moral Conceptual Truths? In *Moral Skepticism*, ed. Diego Machuca. New York: Routledge.

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

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

Cuneo, Terence, and Christos Kyriacou. forthcoming. Defending the Moral/Epistemic Parity. In *Metaepistemology*, ed. Conor McHugh, Jonathan Way, and Daniel Whiting. Oxford: Oxford University Press.

Cuneo, Terence, and Russ Shafer-Landau. 2014. The Moral Fixed Points: New Directions for Moral Nonnaturalism. *Philosophical Studies* 171 (3): 399–443.

Enoch, David. 2011. *Taking Morality Seriously: A Defense of Robust Realism*. Oxford: Oxford University Press.

Evers, Daan, and Bart Streumer. 2016. Are the Moral Fixed Points Conceptual Truths? *Journal of Ethics and Social Philosophy* 10 (1): 1–9.

Feldman, Richard, and Ted A. Warfield. 2010. *Disagreement*. Oxford: Oxford University Press.

Fumerton, Richard. 1995. *Metaepistemology and Skepticism*. Lanham, MD: Rowman & Littlefield.

Hannon, Michael. 2013. "Knows" Entails Truth. *Journal of Philosophical Research* 38: 349–366.

Harman, Gilbert. 1975. Moral Relativism Defended. *Philosophical Review* 84 (1): 3–22.

Hawthorne, John. 2004. *Knowledge and Lotteries*. Oxford: Oxford University Press.

Hazlett, Allan. 2010. The Myth of Factive Verbs. *Philosophy and Phenomenological Research* 80 (3): 497–522.

Heathwood, Chris. 2009. Moral and Epistemic Open-Question Arguments. *Philosophical Books* 50 (2): 83–98.

———. 2018. Epistemic Reductionism and the Moral-Epistemic Disparity. In *Metaepistemology: Realism and Anti-realism*, ed. Christos Kyriacou and Robin McKenna. Basingstoke: Palgrave Macmillan.

Huemer, Michael. 2005. *Ethical Intuitionism*. Basingstoke: Palgrave Macmillan.

Ingram, Stephen. 2015. The Moral Fixed Points: Reply to Cuneo and Shafer-Landau. *Journal of Ethics and Social Philosophy* 9 (1): 1–5.

———. forthcoming. Are Moral Error Theorists Intellectually Vicious? *Journal of Ethics and Social Philosophy*.

Jackson, Frank. 1998. *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Oxford University Press.

Juhl, Cory, and Eric Loomis. 2009. *Analyticity*. Abingdon: Routledge.

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

Killoren, David. 2016. Why Care About Moral Fixed Points? *Analytic Philosophy* 57 (2): 165–173.

Kim, Jaegwon. 1988. What is "Naturalized Epistemology?". *Philosophical Perspectives* 2: 381–405.

Kyriacou, Christos. 2012. Habits-Expressivism About Epistemic Justification. *Philosophical Papers* 41 (2): 209–237.

———. 2016a. Are Evolutionary Debunking Arguments Self-Debunking? *Philosophia* 44 (4): 1351–1366.

———. 2016b. Metaepistemology. Edited by James Fieser and Bradley Dowden. *Internet Encyclopedia of Philosophy*. http://www.iep.utm.edu/meta-epi/.

———. 2017a. Expressivism, Question Substitution and Evolutionary Debunking. *Philosophical Psychology* 30 (8): 1019–1042.

———. 2017b. Moral Fixed Points and Conceptual Deficiency: Reply to Ingram. *Journal of Ethics and Social Philosophy* 11 (3): 1–9.

———. forthcoming-a. Evolutionary Debunking: The Milvian Bridge Destabilized. *Synthese*: 1–19.

———. forthcoming-b. Moral Fixed Points, Rationality and the "Why Be Moral" Question. *Erkenntnis*.

———. n.d. Moral Fixed Points, Error Theory and Meta-Conceptual Deficiency.

Lynch, Michael. 2009. Truth, Value and Epistemic Expressivism. *Philosophy and Phenomenological Research* 79 (1): 76–97.

Matheson, Jonathan, and Brandon Carey. 2013. How Skeptical is the Equal Weight View? In *Disagreement and Skepticism*, ed. Diego Machuca, 131–149. Abingdon: Routledge.

McGinn, Colin. 2012. *Truth by Analysis: Games, Names, and Philosophy*. New York: Oxford University Press.

McKay, Ryan T., and Daniel C. Dennett. 2009. The Evolution of Misbelief. *Behavioral and Brain Sciences* 32 (6): 493–561.

McPherson, Tristram. 2012. Ethical Non-naturalism and the Metaphysics of Supervenience. In *Oxford Studies in Metaethics Vol. 7*, ed. R. Shafer-Landau, 205–234. Oxford: Oxford University Press.

Peacocke, Christopher. 2003. *The Realm of Reason*. Oxford: Oxford University Press.

Pust, Joel. 2001. Against Explanationist Skepticism Regarding Philosophical Intuitions. *Philosophical Studies* 106 (3): 227–258.

Ridge, Michael. 2014. Moral Nonnaturalism. Edited by Edward N. Zalta. *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/entries/moral-non-naturalism/.

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

Schaffer, Jonathan. 2009. On What Grounds What. In *Metametaphysics: New Essays on the Foundations of Ontology*, ed. David Manley, David Chalmers, and Ryan Wasserman, 347–383. Oxford: Oxford University Press.

Shafer-Landau, Russ. 2012. Evolutionary Debunking, Moral Realism, and Moral Knowledge. *Journal of Ethics and Social Philosophy* 7 (1): 1–37.

Stanley, Jason. 2005. *Knowledge and Practical Interests*. Oxford: Oxford University Press.

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

Swinburne, Richard. 2015. Necessary Moral Principles. *Journal of the American Philosophical Association* 1 (4): 617–634.

Turri, John. 2010. Epistemic Supervenience. In *Blackwell Companion to Epistemology (2nd Edition)*, ed. Jonathan Dancy, Ernest Sosa, and Matthias Steup, 340–343. Oxford: Blackwell.

Vavova, Katia. 2014. Debunking Evolutionary Debunking. In *Oxford Studies in Metaethics*, ed. R. Shafer-Landau, vol. 9, 76–101. Oxford: Oxford University Press.

Wielenberg, Erik J. 2014. *Robust Ethics: The Metaphysics and Epistemology of Godless Normative Realism*. Oxford: Oxford University Press.

Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press.

Woods, Jack. 2018. Mathematics, Morality, and Self-Effacement. *Noûs* 52 (1): 47–68.



Normative Reasons for Mentalism

Eva Schmidt

The aim of this paper is to connect the traditional epistemological issue of justification with what one might call the "new reasons paradigm" coming from the philosophy of action and metaethics. More specifically, I will show that Conee and Feldman's mentalism, a version of internalism about justification, can profitably be spelled out in terms of subjective normative reasons. On the way to achieving this aim, I will argue that it is important to ask not just the oft-discussed ontological question about epistemic reasons—what kind of entities are they?—but also: Reasons *in which sense* are fundamental to justification?

My entry point is the debate between externalism and internalism, with a focus on mentalism/evidentialism (Sect. 1). I motivate an insertion of reasons into mentalism and argue that this leads to the question of which kind of reasons are fundamental to justification (Sect. 2). Next, I argue that mentalists should spell out their view by appeal to normative rather than deliberative or explanatory reasons (Sect. 3). I then discuss

E. Schmidt (✉)
Institute of Philosophy, University of Zurich, Zurich, Switzerland
e-mail: eva.schmidt@philos.uzh.ch

© The Author(s) 2018
C. Kyriacou, R. McKenna (eds.), *Metaepistemology*, Palgrave Innovations in Philosophy, https://doi.org/10.1007/978-3-319-93369-6_5

how exactly mentalism with normative reasons should be elaborated (Sect. 4) and conclude by connecting the results to a plausible version of epistemological disjunctivism (Sect. 5).

1 Mentalism/Evidentialism About Justification

To say that a belief is justified is to ascribe to it a certain—*the* central—positive epistemic normative status. Take

The vitamin C case
I believe
(1) that vitamin C is water-soluble.
Moreover, I remember
(2) that I learned that vitamin C is water-soluble in high school
and have the background belief
(3) that vitamin C powder (the kind available in drugstores) can be mixed with water.

In virtue of my memory (2) and my belief (3), I have *ex ante* justification for my belief (1)—it is suitable for me to have this belief, given these mental states. If I also base my belief (1) on my memory and my background belief, it is *ex post* justified (Goldman 1979, 21). Assume that (2) and (3) are the only mental states I have with any bearing on this issue. Then, my belief (1) is all-things-considered justified.

But under which conditions is a subject's belief justified? Internalists affirm, and externalists deny, that a belief's justification depends only on factors that are given from the believer's perspective. One motivation for internalism comes from the so-called new evil demon problem (Cohen and Lehrer 1983). Imagine your non-factive mental duplicate who—unlike you—is constantly deceived by an evil demon, so that her perceptual experiences, memories, and beliefs are mostly incorrect. Given that you and your duplicate reason equally well from the same experiences, memories et cetera, it seems that your beliefs are both equally justified. But if this is correct, then a believer's situation beyond her own mental states and processes cannot make a difference to her justification.

Correspondingly, *mentalism*, as one version of internalism, maintains that the justificatory status of a subject's beliefs strongly supervenes on her non-factive mental states and other mental conditions (Conee and Feldman 2001, 234).[1] Mentalism is combined with evidentialism by Conee and Feldman (2004)—the subject's mental states, as her evidence, determine the degree to which her beliefs are justified. Evidentialism has been phrased as follows:

> (ES) The epistemic justification of anyone's doxastic attitude toward any proposition at any time strongly supervenes on the evidence that the person has at the time (Conee and Feldman 2004, 101).

> (EJ) Doxastic attitude D toward proposition *p* is epistemically justified for S at *t* if and only if having D toward *p* at *t* fits the evidence that S has at *t* (Conee and Feldman 2004, 83).[2]

(ES) makes a strong supervenience claim: The dependence of justification on evidence is such that even across possible worlds, subjects who have the same evidence (mental states) have the same justification for their beliefs. This claim ensures that the new evil demon problem doesn't affect evidentialism, but note that it is quite weak: It allows that non-mental phenomena that don't vary between possible worlds fix justificatory status. Conee and Feldman claim that evidence/mental states determine justification, but worries have been raised about whether supervenience can be used to express such a dependence claim (McLaughlin and Bennett 2018).

(EJ), on the other hand, is a biconditional: A belief is justified if and only if it *fits* with, or is supported by the subject's evidence/mental states. "Fit" can spelled out in different ways, for instance, as saying that the probability of the truth of the belief is raised by the evidence.[3]

Next, (ES) and (EJ) are concerned with *ex ante* and all-things-considered, or overall justification. They don't say whether the subject also bases her belief on her evidence—whether it is *ex post* justified. And her evidence includes *all* of her mental states. But what exactly is the subject's *evidence*? Seeing as we're looking at a combination of mentalism and evidentialism, a subject's evidence consists of the information that she has and can operate on immediately in forming beliefs or, put

somewhat differently, her beliefs, memories, experiences, and other mental states (Conee and Feldman 2008, 88).

Let me now show how to connect this to the reasons paradigm.

2 Reasons for Mentalism

It can be argued that the subject's evidence is identical with her epistemic reasons. First off, it is widely held that "evidence" and "epistemic reason" are more or less synonymous (Kelly 2014). In most, if not all cases of belief, "epistemic reasons" talk and "evidence" talk can be used interchangeably. Conee and Feldman seem to agree (cf. Conee and Feldman 2004, 1, 2008, 88), as does McCain (2014, 10), another evidentialist. This identity can be further supported by noticing that evidence just is what bears on the justification or reasonableness of a belief (Goldman 2011, 255)[4]; but so are epistemic reasons (Gibbons 2010, 344). Finally, what distinguishes epistemic reasons from practical reasons is that only the former bear positively on the truth of the beliefs they support. Evidence, as something that indicates that a belief is true, has just the kind of connection to truth that is specific to epistemic reasons.[5] So we should identify the two.

If this is correct, mentalists/evidentialists should accept *mentalism with reasons*:

> **R-Mentalism**: A subject's belief that *p* is epistemically justified if and only if (and because) the subject has epistemic reasons, overall, for believing that *p*.

Let me clarify. R-Mentalism is restricted to the case of belief; it leaves out the temporal dimension; and it is concerned with all-things-considered *ex ante* justification. It incorporates the idea from (EJ) that one has justification for a belief just in case one's belief fits one's all-things-considered evidence/reasons, but without mentioning "fit" explicitly. That a subject has reasons, overall, *for* believing that *p*, just is to say that so believing fits with her reasons.

Next, in reaction to the aforementioned worries about the supervenience claim from (ES), R-Mentalism incorporates a "because"-claim

instead. R-Mentalism is thereby stronger than mentalism/evidentialism—a subject's beliefs are justified in virtue of the reasons that she has, where this may be understood as a grounding claim (cf. Schmidt forthcoming). R-Mentalism is a kind of mentalism as long as the grounding facts, that is, facts about reasons possession, are mental facts, which can be tested by confronting it with the new evil demon problem.[6]

Even if you reject the identification of evidence with reasons,[7] note that it is *useful* to see what would follow from endorsing R-mentalism. To begin with, there hasn't been much in-depth exploration of how the reasons paradigm relates to traditional epistemological positions. Moreover, even if R-Mentalism and mentalism/evidentialism are genuinely distinct views, they are still close relatives. By getting clear on how the former is best developed, we can probably draw conclusions about the latter as well.

One such conclusion concerns the nature of evidence. As Kelly (2014) points out, "evidence" is a multi-faceted term that can, for example, be understood as that which justifies belief, that which indicates the truth of a proposition, or as something that rational subjects heed. Moreover, Conee and Feldman (2008) sometimes speak of mental states and sometimes of propositions or contents as evidence; in their (2011, 303), they even once identify evidence with facts. Introducing R-Mentalism allows us to tap the resources of the debate over the nature of reasons so as to gain some clarity here.

One might object that, as a matter of fact, epistemologists have discussed the question of what evidence and epistemic reasons consists in—facts propositions, or mental states.[8] So maybe to gain clarity on the nature of evidence/reasons, it is sufficient to turn to this literature. There is no need either to delve into the complicated debate over practical reasons or to accept R-Mentalism.

I disagree. Determining the nature of the mentalist's evidence does not yet settle a further important question concerning mentalism, which requires the reasons idiom: Should the reasons which give rise to epistemic justification be understood as reasons in an explanatory, a deliberative, or a normative sense?

Let me briefly introduce these kinds of reasons for action.[9] *Normative reasons* are considerations that *count in favour* of an agent's φ-ing or that

justify her φ-ing, even if the agent is not aware of them. Take the following example.

> *The hiker and the snake*
>
> A hiker in the desert comes upon a snake, which will bite her unless she remains immobile. The hiker has learned from a reliable source, and thus believes
>
> (4) that the snake will bite unless she turns and runs.
>
> She turns and runs (cf. Parfit 2011).

That the snake will bite the hiker unless she remains immobile is a normative reason for the hiker to remain immobile. *Explanatory reasons*, by contrast, are reasons that explain why the agent φ-s, the beliefs or desires that cause as well as make the agent's φ-ing rationally intelligible (Davidson 1980). For instance, the hiker turns and runs because she believes (4) and desires not to be bitten, which explains her action and makes it rationally intelligible. Finally, *deliberative reasons* are the reasons *in the light of which* an agent φ-s. These are the considerations that she weighs or takes into account as she makes up her mind to φ, and which persuade her to φ. In the example, it is considerations such as (4) in the light of which she turns and runs.

Two important observations concerning explanatory and deliberative reasons: First, there is a close match between them, though the former come into view from a third-person perspective and the latter from a first-person perspective. It is exactly the considerations that are available to the agent as the contents of her rationalising mental states in whose light she acts. *Vice versa*, her mental states rationalise her action exactly because they have a content in whose light the action seems the thing to do.

Second, explanatory and deliberative reasons, so understood, provide *ex post* explanations, but also an *ex ante* understanding of the agent. Her beliefs and desires determine which courses of action available to her would be rationally intelligible. Correspondingly, if we know the considerations available to the agent, we can tell which action will antecedently seem to her to be the one to perform in light of these considerations.

The three varieties of reasons reflect different, yet mutually compatible ways in which we employ "reasons" talk, and can naturally be adapted to

the case of belief. First, normative epistemic reasons are considerations that count in favour of adopting a certain belief. Second, explanatory epistemic reasons are used in explanations that cite the mental states that cause someone to form a belief and that rationalise her believing so. Third, deliberative reasons make a belief intelligible by providing the considerations that persuade the believer to adopt her belief.

The importance of these distinctions derives from the fact that even the question, "What is the nature of evidence?," doesn't touch upon the issue of *whether justification is normative bedrock*, according to mentalism. That is, merely figuring out whether evidence consists in mental states or propositions, say, doesn't help with the question whether justification is the first and fundamental normative epistemic standing, or whether there are other, more basic, normative epistemic phenomena, to which justification can be traced back. However, if we endorse R-Mentalism and ask, "*Are the reasons appealed to by mentalism normative reasons, explanatory reasons, or deliberative reasons for belief?*," we ask whether the reasons that determine the justification of a subject's beliefs are themselves already normative, or whether they are non-normative, merely explanatory or deliberative reasons. Is mentalism a view on which epistemic normativity begins with justification (to which reasons as a purely descriptive factor give rise) or one on which it begins with epistemic normative reasons?

Two remarks before addressing these questions. First, philosophers of action will balk at the suggestion that epistemic justification goes back to non-normative reasons. On the accepted picture, actions are justified in virtue of normative reasons *by definition*, and the same has to hold for belief. However, traditional mentalism can be understood as the claim that certain *non*-normative phenomena determine justification. This understanding is available even if we combine mentalism with reasons rather than evidence, at least given that there is a purely descriptive notion of epistemic reasons. (I will defend this possibility in the next section.) In this context, note further that Conee and Feldman deny that "epistemic concepts are to be analyzed in … deontological terms" (2001, 239). Rejecting a connection between justification and normative reasons would fit well with their negative attitude concerning epistemic obligations.

Second, these options for mentalism bear on the issue of meta-epistemological anti-realism (Kyriacou 2016). If spelled out as appealing

to merely explanatory or deliberative reasons, R-Mentalism makes room for an elimination of normative reasons altogether, and thus for a very strong kind of anti-realism. For recall that explanatory and deliberative reasons are supposed to be reasons in a merely descriptive sense. Even if she appeals to normative reasons, the mentalist is under pressure to accept a kind of mind-dependence of epistemic normativity: Mentalism involves the claim that the facts about the reasons that the subject possesses are mental facts. This suggests that normative reasons are dependent, for their existence, on the subject's (but not thereby on an external appraiser's!) mental states. As will become clear below, this can be either because a normative reason consists in the fact that the subject has certain mental states, or because a consideration is a normative epistemic reason only once it is the content of one of the subject's mental states. On some views, this mind-dependence is sufficient to make R-Mentalism a kind of anti-realism because there are then no facts about normative reasons independent of our mental states (Street 2010).

Now, let's look at combinations of R-Mentalism with deliberative, explanatory, and normative reasons.

3 Which Kind of Reasons for R-Mentalism?

In this section, I discuss varieties of mentalism that result from plugging in different kinds of reasons into R-Mentalism, while holding on to the central mentalist assumption—brought out by the new evil demon problem—that changes beyond the believer's mind make no difference to her justification.

Here's how to identify R-mentalism's reasons with *deliberative* reasons—the contents of her mental states. In the vitamin C case, it is suitable for me to believe (1), that vitamin C is water-soluble, in the light of the remembered consideration (2), that I learned this in school, and the consideration that vitamin C powder can be mixed with water, the content of my background belief (3). These are my deliberative reasons, which I take to bear positively on whether to believe (1). Plausibly, my belief (1) is *ex ante* justified because of how the situation presents itself to me.

Here is the proposal:

> **D-Mentalism**: A subject's belief that *p* is epistemically justified if and only if (and because) the subject has *deliberative* epistemic reasons, overall, for believing that *p*.

In other words, where (and because) the subject's available deliberative reasons present her with a situation which, all things considered, bears positively on the belief that *p*, the latter is justified. As is proper for a mentalist view, the same holds for any possible subject who has the same deliberative reasons.

A lot rides on how we spell out the idea that (the subject takes it that) the considerations bear on whether to adopt the belief. To illustrate this, say that I don't remember (2) and that, rather than believing (3), I believe

> (5) that vitamin *B* powder can be mixed with water.

Imagine that I am a bit confused today and that I take consideration (5) to have a positive bearing on whether to believe (1), so that I would form this belief in the light of (5). But certainly, despite this, my belief (5) does nothing to justify my belief (1).

To avoid this problem, D-Mentalism needs to make deliberative reasons more demanding than they initially appear. When a subject takes a consideration to bear on whether to adopt a belief, she treats the consideration as a normative reason—it is an *apparent* normative reason for her (Parfit 2011, 35). Accordingly, among the considerations that would persuade the subject to adopt a belief, only those *which would be normative reasons to form the belief were they true*, count as deliberative reasons. This would be to supply the mentalist's notion of a deliberative reason with an account of *substantial rationality*, which subjects exhibit who are responsive to what their reasons are reasons for. (Substantive rationality allows for factual mistakes, but not for mistakes about what one's reasons favour.)

Unfortunately, on this proposal, D-Mentalism loses its special interest of providing a completely descriptive foundation for epistemic justification. While deliberative reasons are not themselves normative—merely

apparent normative reasons only *appear* to have normative force—they do conceptually presuppose the notion of a normative reason. For any "bad" case, in which there is a merely apparent reason for the subject to believe that *p* because she is mistaken about an empirical fact, there is a possible "good" case in which she is in touch with the facts and her mental states are correct. Her deliberative reason will then also be a normative reason to believe that *p*, and will contribute to the explanation of the belief's justification. But then normative reasons are fundamental to justification, sometimes because they are directly involved, sometimes because our conception of apparent reasons presupposes them.[10]

Here's an additional problem: The same content can be believed, desired, imagined, et cetera (cf. Turri 2009). So, it may be that I merely *imagine* that vitamin C powder can be mixed with water. Obviously, if this is the only consideration available to me that has a bearing on my belief (1), the latter is not justified even though the consideration is an apparent reason for believing (1). In response, we can restrict deliberative reasons to the contents of presentational attitudes with an assertoric character—only these contribute to justification (Schroeder 2008). But this can't be right either, for my non-assertoric imagining that vitamin C powder can be mixed with water (or my feeling of queasiness, which has no consideration as its content) *does* justify a belief about it, namely, my belief that I am imagining that vitamin C powder can be mixed with water (or that I feel queasy) (Conee and Feldman 2008, 87). Here, it is plausibly the mental state as a whole (including its content or phenomenal character) which is a reason for the subject to believe that she is in the state, and in virtue of which her belief is justified. But then D-Mentalism can't be correct—it is not only her available considerations that contribute the subject's justification, but also her mental states.

Rather than looking for further defenses of D-Mentalism, I take this as sufficient motivation to turn to *mentalism with explanatory reasons*. For there is a close parallelism between deliberative and explanatory reasons anyway (cf. Sect. 2).[11] Moreover, E-Mentalism is concerned with mental states as individuated by their *contents* or their phenomenal character, for otherwise we cannot see how a subject's mental states, taken together, can rationalise her belief with its particular content. So, E-Mentalism doesn't exclude mental contents from being constitutive of epistemic reasons and

justification, it merely insists that the former don't exhaust a subject's reasons.

Here is the view:

> **E-Mentalism**: A subject's belief that *p* is epistemically justified if and only if (and because) the subject has *explanatory* epistemic reasons, overall, for believing that *p*.

To illustrate: My *memory* (2) and my *background belief* (3) are mental states that would make my believing (1) rationally intelligible. So, my belief (1) is *ex ante* justified. Moreover, it is justified *because* of my mental states, and *mutatis mutandis* for my mental duplicates—so, this is a version of mentalism.

E-Mentalism is very close to Conee and Feldman's intended view: It ascribes a central justificatory role to mental states—they themselves are the reasons. It can allow for them to determine a belief's justificatory status even if they are not presentational attitudes, or don't have a (consideration as their) content. For instance, my imagining that vitamin C powder can be mixed with water would rationalise my belief that this is what I imagine (but *not* my belief (3) that vitamin C powder can be mixed with water).

As above, more needs to be said about what is involved in *rationalising* a belief. For it can look like my belief (5), that vitamin *B* powder can be mixed with water, together with my confusion, *does* make my belief (1) rationally intelligible. We can fully understand why I would make the (unwarranted) transition to (1). But, the E-Mentalist can reply, that doesn't make the transition rational. To make this response stick, however, she needs to provide us with an account of what makes belief rational.

An appeal to substantial rationality would once again force her to involve normative reasons in her attempt to capture the basis of justification. The contrast of E-Mentalism with mentalism with normative reasons would be lost. A better option, which fits quite nicely with the idea of rationalising explanations, is to appeal instead to *structural rationality*.[12]

Structural rationality is the kind of rationality that a person exhibits when her mental states cohere with each other. For instance, the structur-

ally rational believer doesn't believe that *p and* believe that not-*p*, nor does she believe that she has insufficient evidence to believe that *p* while believing that *p* (e.g. Broome 1999). Some philosophers argue that structural rationality is itself normative, but many oppose this idea (e.g. Broome 2007). E-Mentalists could join the latter camp. The result would be a view akin to coherentism: A belief is justified for a subject if and only if (and because) an explanation of the belief is available which appeals to its coherence with the subject's pre-existing mental states. For instance, my memory (2) and belief (3) would rationally explain my belief (1) in that an adoption of the latter would cohere with (2) and (3). My belief (1) is therefore justified. If we make our coherence requirements rich enough, by contrast, my belief (1) wouldn't be rationalised by my belief (5), that vitamin B powder can be mixed with water. For, plausibly, it is structurally irrational to adopt a belief like (1) out of thin air, without having any other attitudes from which one might infer to it.

I believe that this is the best available version of R-Mentalism that builds epistemic justification from purely descriptive phenomena. It may seem surprising to first deny that rationality is normative, and then to claim that it gives rise to the normative epistemic standing of justification. However, this is unproblematic. For one, this is just one way of capturing the idea that the normative supervenes on the descriptive. For another, the E-Mentalist might interpret "justification" as an *evaluative* status of a belief. Explanatory reasons are things that make things reasonable—they are the attitudes that cohere with the belief in question and thus show how it is rational. This, in turn, is fundamental to the positive evaluation of the belief as justified.

Unfortunately, E-Mentalism faces some major difficulties. It disconnects a belief's justification from its truth—it doesn't show how justification is truth-conducive, even though what is particular about *epistemic* (rather that moral or pragmatic) justification is that it is a normative standing which flows from belief's connection with truth. That her mental states cohere with each other and that she is thus structurally rational is a feature of the subject that doesn't touch upon her connection to the world, upon whether she is sensitive to the way things really are. Relatedly, this renders it obscure what distinguishes epistemic from practical structural rationality: If we cannot appeal to justification's connection with

truth, how are we to tell which coherence requirements are needed for epistemic rationality and thus justification, and which ones concern practical rationality? For instance, E-Mentalism risks the implausible result that *akrasia* undermines a subject's epistemic justification.

There are two sub-difficulties which articulate this general worry in different ways. First, an explanatory reason's truth makes no difference to how good it is as a reason. Whether my belief (3) or my memory (2) are true, or whether I believe falsely or misremember, does not impede belief (3) and memory (2)'s roles as reasons that explain why, assuming I'm structurally rational, I would believe (1). But intuitively, true reasons are better than false ones; that my reasons are true typically means I'm in touch with the world. This relates to the truth-connection: When my reasons are true, believing according to them gives me pretty good chances that beliefs justified by these reasons will also be true. Not so when my reasons are false. Something appears to be defective about them. E-Mentalism cannot capture this intuition.

Second, E-Mentalism is vulnerable to opponents who claim that internalists are concerned with rationality, but not with the central epistemic standing, justification. For instance, Littlejohn (2012) and Sylvan (2014) suggest that internalism can at best provide an account of "*excusable*" belief. That the subject was (structurally) rational is to say that she exhibited no faulty dispositions in reaching her belief, so if her belief is mistaken, this is excusable. However, this falls short of the belief having the genuinely positive epistemic standing of justification. Similarly, Plantinga (1990) argues that internalist "justification" is not the central epistemic ingredient to knowledge. Applying this to E-Mentalism, adding structural rationality to true belief doesn't result in knowledge.

The strength of such criticisms, when applied to E-Mentalism, derives from the distance between structural rationality and truth. Her structural rationality doesn't bring the believer close enough to epistemic success/truth. It seems as if this way of spelling out justification indeed misses the point.

To sum up, both D-Mentalism and E-Mentalism face severe problems. On the one hand, there is the difficulty of correctly delineating the subject's epistemic reasons, for example excluding justification by merely imagining something, while allowing for justification of beliefs about

one's own mental states. On the other hand, there is the problem of the notion of rationality needed to ensure that not any mental state that would persuade a subject counts as a reason. If we appeal to substantial rationality, we include normative reasons in the foundation of justification and thus fail to provide a non-normative alternative to mentalism with normative reasons; but if we appeal to structural rationality, our notion of justification is (too) disconnected from truth. So, while there is conceptual space for mentalism that incorporates non-normative reasons, no such version is after all defensible. Mentalism with normative reasons is the only game in town. I turn to varieties of this view next.

4 Varieties of Mentalism with Normative Reasons

Normative reasons for belief are facts that count in favour of the subject's adopting certain beliefs. Since we are looking for a mentalist account, the relevant facts can't go beyond the subject's mind. This suggests that normative reasons consist in facts about the subject's own mental states, as formulated by *mentalism with mental state facts as normative reasons*:

> **M-Mentalism**: A subject's belief that *p* is epistemically justified if and only if (and because) the fact that she has mental states m, n, … is a normative reason, overall, for her to believe that *p*.

In the vitamin C case, the fact that I remember (2) and believe (3) *counts in favour* of my believing (1). Therefore, my belief (1) is justified. By contrast, that I believe (5), that vitamin B powder can be mixed with water, does not count in favour of my believing (1); the belief is not thereby justified. Importantly, the fact about the subject's mental states includes *all* her mental states; that she has all of them counts in favour or her believing that *p*, and in virtue of this, her belief is justified. Thereby, M-Mentalism is clearly a mentalist account.

The fact that the subject has certain mental states *counts in favour* of her adopting a certain belief. It is appropriate for the subject to form certain beliefs, given that she has certain other mental states. This could

naturally be spelled out by appeal to coherence requirements, if—contrary to what I assumed in the discussion of E-Mentalism—we think that structural rationality is normative: It gives rise to normative reasons to adopt certain beliefs.

A worry for this account is that, by restricting normative reasons to mental state facts, it implausibly limits what, as a normative reason, determines a belief's justification. It conflicts with our practice of citing worldly facts rather than the fact that we have certain attitudes (regarding these facts) as normative reasons to believe. What justifies my belief (1)? The fact (2) *that I learned it in school* and the fact (3) *that vitamin C powder can be mixed with water*. What I'm going by in adopting my belief is the facts that my mental states represent, not the fact that I have these mental states. M-Mentalism makes it look as though only structural rationality, or coherence between mental states, provides normative reasons and thus justification for belief. But on our ordinary conception, the world beyond our mind also favours certain beliefs. Plausibly, this is so because worldly facts can make it probable that a belief is true. For instance, given that I learned this in school, it is extremely likely that vitamin C is indeed water-soluble. So, we should allow that worldly facts that reliably indicate that a belief is true are epistemic normative reasons *and* help determine a belief's justification. M-Mentalism, since it excludes this highly plausible possibility, should be abandoned.[13]

Here's a proposal that takes the facts that our mental states represent as the normative reasons that determine which beliefs are justified for a subject, *mentalism with objective normative reasons*:

> **O-Mentalism**: A subject's belief that *p* is epistemically justified if and only if (and because) the facts represented by her mental states, overall, count in favour of her believing that *p*.

In the vitamin C case, I have justification for my belief (1) if and only if my memory (2) and my belief (3) represent normative reasons, overall, for me to adopt this belief. They do: The fact that vitamin C powder can be mixed with water reliably indicates that vitamin C is water-soluble, as does the fact that I learned this in school. So, these facts count in favour of so believing, and my belief (1) is justified in virtue of this.

Unfortunately, O-"Mentalism" isn't a kind of mentalism: It's not (just) in virtue of the subject's mental states that her beliefs are justified. Duplicates don't necessarily have the same justification. Consider:

> *The vitamin C duplicate case*
> Unlike me, my duplicate lives in a world in which vitamin C powder cannot be mixed with water. Neither of us remembers (2), but we both believe (1) and (3).

My belief (1) is justified, for my belief (3) represents the fact that vitamin C powder can be mixed with water, which is a normative reason to believe that vitamin C powder is water-soluble. In my duplicate's world, however, it is not the case that vitamin C powder can be mixed with water. So, her belief (3) doesn't represent a fact that counts in favour of her adopting (1); her belief (1) is not justified. The justificatory status of our beliefs (1) is not the same even though we are duplicates. For O-"Mentalism," then, a subject's belief isn't justified entirely in virtue of her mental states, so this proposal cannot be used to spell out mentalism.

But there is a version of mentalism in the vicinity. We merely have to allow that not just the *facts*, but also the *merely apparent* facts represented by a subject's presentational attitudes may be normative reasons. For this, we can appeal to Schroeder's (2008) *subjective* normative reasons, considerations—but not necessarily facts—from the subject's perspective, that speak in favour of her φ-ing. These are merely apparent normative reasons (cf. Sect. 3), relating to substantial normativity, except that on Schroeder's picture, they involve *genuine normativity*. Given that the world appears to her in a way that favours φ-ing, it is rational for her to φ. We expect her to φ and think that, in the light of the information available to her, she ought to φ. To illustrate, if the hiker justifiably believes (4), that the snake will bite her unless she turns and runs, she has a subjective normative reason to turn and run. Given what she believes, she would be crazy to stand still (or at least irrational and hard to make sense of). Nonetheless, there is an *objective* normative reason for her not to move: The *fact* that the snake will only bite if she turns and runs objectively counts in favour of her remaining immobile.

Making room for these two kinds of normative reasons allows us to make full sense of the hiker's normative situation, of a sense in which she ought to turn and run and of a sense in which she had better not move.

We can then introduce *mentalism with subjective normative reasons*:

> **S-Mentalism**: A subject's belief that *p* is epistemically justified if and only if (and because) the apparent facts represented by her mental states, overall, count in favour of her believing that *p*.

Certain facts appear to obtain to the subject, namely those that her presentational attitudes, such as belief and perceptual experience, represent. Given these apparent facts, which constitute the situation that the subject takes herself to be faced with, there are beliefs that are substantially rational for her to form, and these beliefs are ones that she ought to have or that she has subjective normative reasons to adopt. These are the beliefs for which she has justification.[14]

The vitamin C duplicate case causes no trouble for S-Mentalism. With respect to our subjective normative reasons, my duplicate and I are in the same boat: Given that we both labour under the assumption that vitamin C powder can be mixed with water, it is rational for us to believe (1). The apparent fact that belief (3) presents us with is a subjective normative reason for both of us to believe (1). In virtue of this, both of us have equal justification to believe (1).

But here is a more problematic case.

> *The berries case*
>
> My duplicate and I both see
>
> (6) that there are such-and-such berries on that bush.
>
> We both have a disposition to form the belief
>
> (7) that there are edible berries on that bush
>
> immediately on the basis of experiencing (6), for the berries just "look" edible to us. In my world, berries of this kind are edible, but in my duplicate's world, they are poisonous. The fact (6), represented by my experience, reliably indicates that there are edible berries on the bush, but not so the fact (6) represented by my duplicate's experience.

The trouble for S-Mentalism is that I have justification for belief (7), as for me the perceived fact (6) makes it probable that (7) is true and thus is a normative reason to believe (7). But my duplicate lacks such justification because in her world what she perceives, that there are such-and-such berries on that bush, doesn't reliably indicate (7) and so is no reason. So, my duplicate and I differ with respect to the justification of our respective beliefs (7).

To reply, in order to have justification, the subject needs to *possess* the relevant normative reasons. It is not sufficient that the subject's mental states represent apparent facts that favour so believing. To possess reasons, she also has to be able to be guided by the subjective normative reason to an appropriate belief (cf. Sylvan 2015). So, if my duplicate and I are both clueless about whether berries of this kind are edible, neither of us is sensitive to the fact that there are berries in the bush *as* a reason to believe (7). Neither of us then possesses (6) as a reason to believe (7), and neither of our beliefs is justified. This fits with everyday judgments about reasons and justification. Imagine an elderly lady looking at someone pulling the trigger on a paint ball gun that is pointed at her—she has a presentational attitude representing a normative reason to believe that she will be marked with paint. If she is clueless about such things, however, we would not say that she possesses a reason, or has justification for this belief.

But maybe in the berries case, I possess (6) as a reason, but my duplicate doesn't? After all, only I am guided by my experience to the correct belief. In response, if I have the capacity to be guided by subjective normative reasons, that is, by *apparent* facts, so does my duplicate. For there are no differences between duplicates with respect to their apparent reasons, and so both my duplicate and I either have or lack the competence to move from appearances as of edible berries to the belief that these berries are edible. This is compatible with the fact that in my duplicate's case, appearances were deceiving.

This leads to an improved version of S-Mentalism,

> **S-Mentalism***: A subject's belief that *p* is epistemically justified if and only if (and because) the subjective normative reasons she possesses, overall, count in favour of her believing that *p*.

All in all, of the three proposals discussed in this section, O-"Mentalism" fails as a version of mentalism since it does not guarantee that facts about reasons possession are exclusively mental facts. M-Mentalism is implausible because it limits the normative reasons relevant to justification to mental state facts. This leaves S-Mentalism* as the best version of R-Mentalism.[15]

5 Results

The idea with which I started is that the close ties between evidence and epistemic reasons, together with the openness of mentalism with respect to the kind of reasons we insert, allows us to spell out the view in interestingly different ways, even in such a way that the relevant reasons are non-normative. This should give reasons theorists who try to transfer their views from the practical realm to epistemology pause: We cannot automatically get from normative reasons as the foundation of practical justification to normative reasons as the basis of epistemic justification. More work has to be done to show that the practical picture can be transferred, one-to-one, to epistemology.

Here, I have done so by arguing that the only way to keep the reasons involved in R-mentalism non-normative disrupts the connection between justification and truth: If we take reasons as rationalisers, and rationality as mere structural rationality, we end up with reasons that are not truth-conducive. But if we understand reasons as apparent reasons, we put normative reasons at the foundations of justification after all.

It may seem that my discussion of different versions of mentalism has been unfair. On the one hand, won't mentalists be able to resist my argument by sticking with evidence instead of reasons? On the other hand, doesn't my proposal, in which S-Mentalism* is the central pillar of mentalism, also face the problem that it severs the truth-connection? I will close the paper by addressing these two worries.

As to the first worry, I concede that the mentalist/evidentialist can dig in her heels and insist that evidence is the basis of justification, while denying that this commits her to reasons in any of the senses discussed

here. Still, my discussion, together with the ambiguity of "evidence," puts the burden on my opponent to clarify her notion of evidence, and its relation to epistemic reasons, in order to ensure that it does not so commit her.

As to the second worry, traditional versions of internalism like E-Mentalism explicate justification by appeal to the limited information that is available to the subject independently of her external situation. They can neither account for the truth-connection nor give epistemic priority to the good case in which the subject is in touch with the world. S-Mentalism*, by contrast, takes as its base case one in which the subject is presented with facts that count in favour of adopting a belief (they actually indicate its truth). From such objective normative reasons, it derives what the subject has subjective normative reason to believe—and is hence justified to believe—in indistinguishable cases where these facts are merely apparent. In this sense, then, S-Mentalism* gives priority to the good case. It thereby makes room for a truth-connection: Justification connects to the truth in that it is due to one's subjective normative reasons—apparent facts that derive their normative strength from their objective "twins" (i.e., the corresponding objective normative reasons that *do* make the belief's truth probable).

An additional advantage of S-Mentalism* that is thrown into relief by contrast with E-Mentalism is that it can be combined with an externalist view along the lines of O-"Mentalism," resulting in a kind of epistemological disjunctivism (cf. Hornsby 2008; Schroeder 2015). Objective and subjective normative reasons are not mutually exclusive. Rather, we can use these two types of reasons to highlight different aspects of how the subject's situation bears on what she had better believe—on the one hand, the subject's *actual* situation, which favours certain beliefs, so that her justification is best when she is in touch with these objective normative reasons. On the other hand, the subject's situation, *as she takes it to be*, subjectively favours her believing certain things. Plausibly, reasons of both kinds bear on the justification of her beliefs.

Such an epistemological disjunctivism is then clearly a realist view of normative reasons. Both the way it spells out subjective normative reasons and the combination of S-Mentalism* with O-"Mentalism" rely on the existence of mind-independent objective normative reasons.

If mentalism is understood along these lines, the contention that internalism fails to address the central epistemic normative standing—justification—is misguided. The subject's subjective normative reasons, thanks to their close connection with their objective "twins", allow for her to have beliefs that are not just excusable, but genuinely epistemically proper. Nonetheless, the contrast with objective normative reasons brings out a respect in which the justification provided by S-Mentalism falls short of epistemic success: That I have subjective normative reasons does not yet guarantee that I am in touch with the world.[16]

Notes

1. I focus on *non-factive* mental states and duplicates throughout.
2. To focus the discussion, I will restrict myself to a thesis about the justification of belief, rather than doxastic attitudes generally. Since it plays no role for my argument, I will spend no time on the claim involved in (ES) and (EJ) that only *concurrent* evidence makes a difference to justification.
3. Note that Conee and Feldman (2008) instead appeal to explanatory relations.
4. Conee and Feldman (2011) are not too happy with this suggestion, which Goldman uses to argue that their view is circular—but in response, they appeal to encyclopedia entries which, again, equate evidence with grounds or reasons.
5. I have defended the view that there is a broader gap between *practical* reasons and evidence. See Schmidt (2017).
6. One get-out strategy for mentalists/evidentialists is to reject the move from supervenience to grounding. Because of limited space, I cannot further motivate the grounding claim, so I here merely acknowledge that my proposal is indeed stronger than the original mentalist/evidentialist claim.
7. Cf. Littlejohn (2013) and Sylvan and Sosa (2018).
8. See Hawthorne and Magidor (2018), Littlejohn (2012), McCain (2014), Neta (2008), Turri (2009), Williamson (2000), among others.
9. Cf. Alvarez (2017), Hieronymi (2011), Mantel (2018), Olson and Svensson (2005), among many others.

10. Contrary to my suggestion above that deliberative reasons are naturally incorporated into an anti-realist view, D-Mentalism would then be a realist position: It would presuppose the existence of mind-independent normative reasons after all.
11. With one significant difference, though: More entities come into view as reasons if we turn to explanatory reasons, for these aren't restricted to presentational attitudes that have considerations as their contents.
12. There is a third option, which I cannot address here: Appealing to a *purely subjective* notion of rationality (Foley 1987).
13. Plausibly, to *possess* reasons, one has to have certain mental states no matter what, but nonetheless the non-mental facts may be the reasons. Cf. Schmidt (forthcoming).
14. Just like D-Mentalism, S-Mentalism faces the problem that not any kind of mental state—e.g. an imagining—that represents a subjective normative reason to adopt a belief makes a difference to its justification. It can be solved by restricting the mental states in S-Mentalism to presentational attitudes. Regarding beliefs *about one's own mental states*, the correct account is indeed M-Mentalism. The fact that the subject has a mental state is a normative reason for her to believe that she does.
15. Again, it should be combined with M-Mentalism to account for the justification of our beliefs about our own mental states.
16. I thank Chrisoula Andreou, Vuko Andrić, Kevin Baum, Christoph Fehige, Susanne Mantel, Jean Moritz Müller, Stephan Padel, Katia Samoilova, Kurt Sylvan, Christian Wendelborn, Ulla Wessels, and Hong Yu Wong for helpful suggestions on this paper.

References

Alvarez, Maria. 2017. Reasons for Action: Justification, Motivation, Explanation. Edited by Edward N. Zalta. *The Stanford Encyclopedia of Philosophy (Winter 2017 Edition)*. https://plato.stanford.edu/archives/win2017/entries/reasons-just-vs-expl/.

Broome, John. 1999. Normative Requirements. *Ratio* 12 (4): 398–419.

———. 2007. Is Rationality Normative? *Disputatio* 2 (23): 161–178.

Cohen, Stewart, and Keith Lehrer. 1983. Justification, Truth, and Coherence. *Synthese* 55 (2): 191–207.

Conee, Earl, and Richard Feldman. 2001. Internalism Defended. In *Epistemology: Internalism and Externalism*, ed. Hilary Kornblith, 231–260. Hoboken, NJ: Wiley-Blackwell.

———. 2004. *Evidentialism: Essays in Epistemology*. Oxford: Oxford University Press.

———. 2008. Evidence. In *Epistemology: New Essays*, ed. Quentin Smith, 83–104. Oxford: Oxford University Press.

———. 2011. Replies. In *Evidentialism and Its Discontents*, ed. Trent Dougherty, 283–323. Oxford: Oxford University Press.

Davidson, Donald. 1980. Actions, Reasons, and Causes. In *Essays on Actions and Events*, ed. Donald Davidson, 1–19. Oxford: Clarendon.

Foley, Richard. 1987. *The Theory of Epistemic Rationality*. Cambridge, MA: Harvard University Press.

Gibbons, John. 2010. Things That Make Things Reasonable. *Philosophy and Phenomenological Research* 81 (2): 335–361.

Goldman, Alvin. 1979. What is Justified Belief? In *Justification and Knowledge*, ed. George Pappas, 1–25. Dordrecht: Reidel.

———. 2011. Toward a Synthesis of Reliabilism and Evidentialism? In *Evidentialism and Its Discontents*, ed. Trent Dougherty, 254–280. Oxford: Oxford University Press.

Hawthorne, John, and Ofra Magidor. 2018. Reflections on the Ideology of Reasons. In *The Oxford Handbook of Reasons and Normativity*, ed. Daniel Star, 113–139. Oxford: Oxford University Press.

Hieronymi, Pamela. 2011. Reasons for Action. *Proceedings of the Aristotelian Society* 111 (3pt3): 407–427.

Hornsby, Jennifer. 2008. A Disjunctivist Conception of Acting for Reasons. In *Disjunctivism: Perception, Action, Knowledge*, ed. Adrian Haddock and Fiona Macpherson, 244–261. Oxford: Oxford University Press.

Kelly, Thomas. 2014. Evidence. Edited by Edward N. Zalta. *The Stanford Encyclopedia of Philosophy (Fall 2014 Edition)*. http://plato.stanford.edu/archives/fall2014/entries/evidence/.

Kyriacou, Christos. 2016. Metaepistemology. Edited by James Fieser and Bradley Dowden. *Internet Encyclopedia of Philosophy*. http://www.iep.utm.edu/meta-epi/.

Littlejohn, Clayton. 2012. *Justification and the Truth-Connection*. Cambridge: Cambridge University Press.

———. 2013. The Russellian Retreat. *Proceedings of the Aristotelian Society* 113 (3pt3): 293–320.

Mantel, Susanne. 2018. *Determined by Reasons: A Competence Account of Acting for a Normative Reason*. Abingdon: Routledge.
McCain, Kevin. 2014. *Evidentialism and Epistemic Justification*. Abingdon: Routledge.
McLaughlin, Brian, and Karen Bennett. 2018. Supervenience. Edited by Edward N. Zalta. *The Stanford Encyclopedia of Philosophy (Spring 2018 Edition)*. https://plato.stanford.edu/archives/spr2018/entries/supervenience/.
Neta, Ram. 2008. What Evidence Do You Have? *British Journal for the Philosophy of Science* 59 (1): 89–119.
Olson, Jonas, and Frans Svensson. 2005. Regimenting Reasons. *Theoria* 61 (3): 203–214.
Parfit, Derek. 2011. *On What Matters*. Oxford: Oxford University Press.
Plantinga, Alvin. 1990. Justification in the 20th Century. *Philosophy and Phenomenological Research* 50 (Suppl.): 45–71.
Schmidt, Eva. 2017. New Trouble for 'Reasons as Evidence': Means That Don't Justify the Ends. *Ethics* 127 (3): 708–718.
———. forthcoming. Possessing Epistemic Reasons: The Role of Rational Capacities. *Philosophical Studies*. https://doi.org/10.1007/s11098-017-1025-z/.
Schroeder, Mark. 2008. Having Reasons. *Philosophical Studies* 139 (1): 57–71.
———. 2015. Knowledge is Belief for Sufficient (Objective and Subjective) Reason. In *Oxford Studies in Epistemology Vol. 5*, ed. John Hawthorne and Tamar Szabó Gendler, 226–252. Oxford: Oxford University Press.
Street, Sharon. 2010. What is Constructivism in Ethics and Metaethics? *Philosophy Compass* 5 (5): 363–384.
Sylvan, Kurt. 2014. On the Normativity of Epistemic Rationality. Thesis, Rutgers University.
———. 2015. What Apparent Reasons Appear to Be. *Philosophical Studies* 172 (3): 587–606.
Sylvan, Kurt, and Ernest Sosa. 2018. The Place of Reasons in Epistemology. In *The Oxford Handbook of Reasons and Normativity*, ed. Daniel Star, 555–574. Oxford: Oxford University Press.
Turri, John. 2009. The Ontology of Epistemic Reasons. *Noûs* 43 (3): 490–512.
Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press.



Epistemic Consequentialism: Haters Gonna Hate

Nathaniel Sharadin

1 Introduction

Something's *final value* is value it has that doesn't depend on its standing in any relation to anything else.[1] Familiarly, a *hedonistic* theory of *practical* value holds that:

1. Pleasure has positive final practical value.
2. Pain has negative final practical value.
3. Nothing else has (positive or negative) final practical value.[2]

Hedonism is naturally paired with the following practical first-order normative theory:

> **Practical Outcome-Consequentialism** (POC): An agent A's action φ at time t is practically correct iff A's φ-ing at t is included in the practically best outcome available to A at t.

N. Sharadin (✉)
Department of Philosophy, The College of New Jersey, Ewing, NJ, USA
e-mail: natesharadin@gmail.com

© The Author(s) 2018
C. Kyriacou, R. McKenna (eds.), *Metaepistemology*, Palgrave Innovations in Philosophy, https://doi.org/10.1007/978-3-319-93369-6_6

Together, hedonism and POC tell us that the practical correctness of an action is a function of the pleasure and pain present in outcomes available to an agent at a time.[3]

Less familiarly, a *veritistic* theory of *epistemic* value holds that:

1. True belief has positive final epistemic value.
2. False belief has negative final epistemic value.
3. Nothing else has (positive or negative) final epistemic value.[4]

Veritism is naturally paired with the following *epistemic* first-order normative theory:

> **Epistemic Outcome-Consequentialism** (EOC): An agent A's belief that *p* at *t* is epistemically correct iff A's belief that *p* at *t* is included in the epistemically best outcome available to A at *t*.

Together, veritism and EOC tell us that the epistemic correctness of believing some proposition is a function of the truth and falsity present in the outcomes available to an agent at a time. From here on I'll talk about beliefs being epistemically *rational*, *reasonable*, or *justified* rather than merely *correct*.[5]

Recently, the package of veritism and EOC has come under attack on a number of fronts.[6] Here, I'm interested in one particular criticism: that this package allows for illicit "trade-offs" between propositions. The criticism is simple to state but, as we'll see, it's difficult to pin down. The simple way to state the criticism is via an example. I'll borrow one from Selim Berker, who has recently pressed the case against the pairing of veritism and EOC:

> **Scientist**: Suppose I am a scientist seeking to get a grant from a religious organisation. Suppose, also, that I am an atheist: I have thought long and hard about whether God exists and have eventually come to the conclusion that He does not. However, I realise that my only chance of receiving funding from the organisation is to believe in the existence of God: they only give grants to believers, and I know I am such a bad liar that I won't be able to convince the organisation's review board that I believe God exists unless I genuinely do. Finally, I know that, were I to receive the grant, I would use

> it to further my research, which would allow me to form a large number of new true beliefs and to revise a large number of previously held false beliefs about a variety of matters of great intellectual significance. (Berker 2013a, 363–64)[7]

According to Berker, "most of us" would agree that forming a belief in God's existence in these circumstances would not be epistemically reasonable, rational, or justified (2013a, 364). The problem, according to Berker, is that if we accept both veritism and EOC then we appear to be barred from saying this agreeable thing. This is because, were I to form a belief in God's existence in these circumstances, I'd be making what according to the pairing of EOC and veritism looks like an excellent *trade-off*:

> I would be trading off the pursuit of one thing of final epistemic value (true belief and the avoidance of false belief with regard to the proposition <God exists>) in order to greatly facilitate my pursuit of a variety of other things of final epistemic value (true belief and the avoidance of false belief with regard to the propositions that make up the subject matter of my scientific research). (2013a, 364)

Such trade-offs between *propositions* licensed by EOC and veritism are reminiscent of trade-offs between *persons* licensed by some versions of POC and hedonism: just as the latter view seems to sometimes require us to sacrifice one person's pleasure for the "greater practical good," the former seems to sometimes require us to sacrifice one proposition's truth for the "greater epistemic good" (Berker 2013a, 373, 375). However, says Berker, in the epistemic case such trade-offs are always illicit, and this is because

> [W]hereas in the ethical case there is some room for arguing that it is not a normative mistake to ignore the separateness of persons, […] in the epistemic case there is no wiggle room: *the epistemic separateness of propositions is nonnegotiable*. (2013a, 365)

But: Is it? I am not so sure. In what follows I have two principle aims. The first is to figure out what, exactly, the "epistemic separateness of

propositions" amounts to and, along the way, to figure out why, exactly, philosophers such as Berker think it's non-negotiable. Since the *epistemic* separateness of *propositions* is supposed to be an analogue to the *moral* separateness of *persons*, I'll explore arguments for the nature and negotiability of the former in parallel with arguments over the nature and negotiability of the latter. Armed with a clearer understanding of the separateness of propositions, I'll suggest that *h*ardcore *a*nti-*t*rade-off-*ers* (henceforth *haters*, for short) of Berker's sort face a dilemma. On the one hand, haters can model their understanding of the epistemic separateness of propositions on the moral separateness of persons. But, I'll argue, we have been given no reason to think the separateness of propositions, so understood, is absolutely non-negotiable. Indeed, as I'll explain, the moral analogue militates in *favour* of negotiability. On the other hand, haters can decline to use the moral analogy. But then we are left with no meaningful content to the idea that propositions are epistemically separate. In that case, we clearly have no reason to accept that propositions are in any interesting (non-metaphorical, non-question-begging) way epistemically separate and so, *a fortiori*, no reason to think their epistemic separateness is absolutely non-negotiable. My second principle aim is to diagnose why haters might have been led to think that the epistemic separateness of propositions was absolutely non-negotiable, and so that trade-offs are always banned. I'll argue that they've been led to this view via a methodological mistake. Let me make two brief preliminary remarks before getting on with things.

First: as I noted above *all* haters, that is, all those who think that epistemic trade-offs are always forbidden, are my target. But Berker in particular will be my foil: I'll take him as representative of the class of views that finds something seriously disturbing about the idea that it can be epistemically permissible to trade off one proposition's truth against another's. I focus on Berker for two reasons. First, Berker is clearest, among haters, about what the structure of the argument against trade-offs is supposed to be (this will be important in pursuing my second principle aim in the latter half of the paper). Second, Berker is, as far as I can tell, alone among haters in explicitly appealing to the analogues between the separateness of propositions, on the one hand, and the separateness of persons, on the other. Since I'm going to spend the first half

of the paper arguing that this analogy isn't helpful for haters, it might seem I'm loading the die by focusing on Berker in this way (since only he appeals explicitly to it). But as a matter of fact I think it's a *good analogy*, in the sense that it gives voice to the relevant sentiment that is behind haters' thinking. If Berker hadn't offered the analogy on the part of haters, I'd happily do so myself. Unfortunately, as I'll argue in the first half of the paper, the analogy doesn't help—and might well undermine—haters' position. But that doesn't mean it's not worth exploring in some detail.

The second preliminary remark is this: if I'm successful in achieving my aims, I'll have defended the pairing of EOC and veritism against the charge that it allows for objectionable trade-offs. I'm not the first to offer such a defence. But previous defences have all, in various ways, attempted to show that haters' understanding of veritism or of EOC are somehow flawed. The idea behind these previous defences, then, has been that EOC and veritism do not have the implausible result haters ascribe to it: contrary to what haters say, EOC and veritism do not "ignore the separateness of propositions" or allow for the trade-offs to which haters object.[8] By contrast, my defence of EOC and veritism begins by accepting haters' charge that EOC and veritism sometimes ignores the separateness of propositions and allows for trade-offs. I'm interested in what, precisely, it means to say this and why, precisely, we should care if it's true. What I'm going to argue is that if we make good sense out of the claim that propositions are epistemically separate, then it turns out we shouldn't think such separateness is non-negotiable; if instead we make a hash out of the idea of propositions' separateness, then their separateness *really* isn't non-negotiable. Let's get on with it.

2 The Moral Separateness of Persons and the Epistemic Separateness of Beliefs

The claim that propositions are "separate" is a metaphor: to be meaningful, it must be unpacked. We are told that propositions are "*epistemically* separate," but this turns out to mean that propositions are the sort of thing such that trading one's truth off against another's is always illicit.

That's obviously not much help in the present context. Happily, haters do provide some guidance: they point us to the separateness of different things in a different normative domain they think is analogous to the separateness enjoyed by propositions, namely the *moral* separateness of *persons*. So, let's take them at their word and explore the analogy in the hopes of refining our understanding of the separateness of propositions.[9] What does it mean to say that persons are morally separate?

The claim that persons are morally separate is no less metaphorical. But proponents of the idea in the moral domain have said something somewhat more concrete about what they mean by this idea. Tellingly, moral philosophers regularly appeal to the separateness of persons in the course of arguing against versions of the pairing noted above between hedonism and POC. Since epistemologists such as Berker appeal to the analogous idea in the course of arguing against the analogous pairing of veritism and EOC, it seems we might be on the trail of our quarry. To get us started, here are a few representative samples of the idea on the moral side of things:

> [Utilitarianism's] view of social cooperation is the consequence of extending to the society the principle of choice for one man, and then, to make this extension work, conflating all persons into one […] Utilitarianism does not take seriously the distinction between persons. (Rawls 1971, 27)

> There are only individual people, different individual people, with their own individual lives. Using one of these people for the benefit of others, uses him and benefits the others. Nothing more. What happens is that something is done to him for the sake of others. Talk of an overall social good covers this up. (Intentionally?) To use a person in this way does not sufficiently respect and take account of the fact that he is a separate person, that his is the only life he has. (Nozick 1974, 32–33)

> [Consequentialism] treats the desires, needs, satisfactions, and dissatisfactions of distinct persons as if they were the desires, etc., of a mass person. (Nagel 1970, 134)

> [S]uppose that mankind is a super-person, whose greatest satisfaction is the objective of moral action…. But this is absurd. Individuals have wants, not

> mankind; individuals seek satisfaction, not mankind. A person's satisfaction is not part of any greater satisfaction. (Gauthier 1963, 126; cited in Norcross 2009)

> The basic problem with standard versions of core consequentialism is that they fail to recognize adequately the *normative separateness of persons*. Psychological autonomous beings (as well, perhaps, as other beings with moral standing) are not merely means for the promotion of value. They must be respected and honored, and this means that at least sometimes certain things may not be done to them, even though this promotes value overall. (Vallentyne 2006, 29; cited in Norcross 2009)

Importantly, despite some of the misleading language (especially in the Nagel, Nozick, and Gauthier quotes), the point here is not a metaphysical one about the fact that persons enjoy distinct existences (cf. Norcross 2009). And, correspondingly, the fundamental mistake those who ignore the separateness of persons make is not a metaphysical one. For even supposing there were, metaphysically speaking, such a thing as a "mass-" or "super-person" comprising what we ordinarily think of as individual persons, those concerned with the moral separateness of persons would think this was neither here nor there when it came to how, morally speaking, we should treat the component persons that make up such a super-person. The mistake, in other words, is not in thinking that super-persons *exist*, but is instead in thinking that super-persons are, when it comes to what we are morally obliged to do, the appropriate *foci* of our normative thought.

Vallentyne's quote is perhaps the clearest on this point, though each of these authors makes it clear elsewhere that it is the *normative*—not metaphysical—separateness of persons that concerns them. As Vallentyne characterises it, the normative separateness of persons consists in the fact that they must be "respected and honored" and not merely treated as "means for the promotion of value." And in his later work, Rawls says that it is distinctive of persons that they are "self-authenticating sources of valid claims," requiring our respect (1995, 32, 72, 100). Below I'll discuss this in some more detail. For now, we can simply note that each of these authors, in different ways, understands the *normative separateness*

of persons as placing restrictions on the appropriateness of certain trade-offs between what is good for one person and what is good for another. Since it's something like this latter claim that haters want to make regarding the *epistemic separateness of propositions* (i.e., that there are restrictions on the appropriateness of trade-offs between them), it looks like we've found our analogue.

Let's dig in to the analogy a bit further. First notice that, in the moral domain, it is *persons* that are separate in the sense that there are restrictions on trading off what's good for one person against what's good for another. This suggests that the correct epistemic analogy is in terms of the separateness of *beliefs* not *propositions*. To explain: recall our two theories of value. Hedonism claims *persons' pleasure* has positive final practical value.[10] Veritism claims *true belief* has positive final epistemic value. Restrictions on trade-offs between *persons* on the basis of their separateness are restrictions on trading off one person's thing hedonism says is valuable against another person's thing hedonism says is valuable. So, if we mean to make the analogous claim in epistemology, we should not say that the issue is over whether there is a ban on trade-offs between *propositions*. Veritism doesn't say anything about the value of *propositions*. If there are restrictions on trade-offs, they are restrictions on trading off what's good for a particular belief against what's good for a separate belief. Our question about the epistemic separateness of *propositions* is, properly, a question about the epistemic separateness of *beliefs*.[11]

With this in mind, the idea is that, just as *persons* are separate in the sense that there are restrictions on their good being be traded off against *another person's good*, so *beliefs* are separate in the sense that there are restrictions on their *truth* being traded off against *another belief's truth*. This seems to be exactly what haters have in mind. Recall Berker's example of the atheistic scientist. What was supposed to be objectionable, according to Berker, was that in those circumstances he was considering trading off the negative value of his false belief in God against a range of other true beliefs concerning the subject matter of his scientific research.

All this should be uncontroversial. It's simply making explicit the analogue between the moral separateness of persons and the epistemic

separateness of beliefs (propositions) alluded to by haters such as Berker. The interesting thing, as I'll now explain, is what happens when we take this analogy seriously.

3 Interpersonal and Interdoxastic Aggregation of Value

If persons are normatively separate in the way just described, what follows with respect to the permissibility of *trade-offs* between persons, that is, interpersonal trade-offs? An insistence on the absolute non-negotiability of the normative separateness of persons would entail it was *never* morally permissible to sacrifice (any amount of) one person's good in order to bring about (any amount of) any number of persons' goods. In other words, insistence on the absolute non-negotiability of the normative separateness of persons entails an absolute ban on interpersonal trade-offs. But an absolute ban on the interpersonal trade-offs leads to absurd first-order results. It entails that it is *never* morally permissible to sacrifice one person's life in order to save five lives, or five million, or even five billion; a strict ban on interpersonal trade-offs entails it's never morally permissible to give one person a *short mild headache* in order to save five billion lives. Assuming, as Alistair Norcross puts it, that we don't want to follow those who insist on such a ban to the "funny farm," we should reject an *absolute* ban on interpersonal trade-offs and accept, instead, that it is at least *sometimes* morally permissible to trade-off one person's good against another's (2009, 80).[12]

Why accept the moral permissibility of interpersonal trade-offs? One reason is, as I've just said, is that rejecting it leads to absurd results—we end up at the funny farm. Another reason, which matters more in the present context, is that *even the most ardent proponents of the moral separateness of persons accept interpersonal trade-offs*. This is worth pausing over for a moment. Rawls is often credited with his celebration of the separateness of persons. But even Rawls doesn't end up endorsing a view according to which interpersonal trade-offs are completely off the table. Even he thinks interpersonal trade-offs are *negotiable*.

Perhaps the most obvious illustration of the negotiability of such trade-offs in Rawls's theory is his commitment to a principle of distributive justice requiring that inequalities in wealth and income be justified by their positive contribution to the position of the least well-off in a society. This is his famous "Difference Principle," which says that differences in inequality in income and wealth arising from the arrangement of the basic structure of society are justified only if those differences make the least well-off in society better off than they otherwise would be under some alternative arrangement (Rawls 1971, especially the discussion beginning on p. 76). Implicit in the Difference Principle, then, is the permissibility of some interpersonal trade-offs. This is because the Difference Principle forbids (mere) Pareto improvements in a distribution where some people, namely the best-off in a society, could be made better off without worsening—but also without *improving*—the position of the worst-off in that society. Hence the Difference Principle says it is licit—indeed, that we are *required*—to trade-off the good of some persons (those comprising the best-off) against the good of some other persons (those who comprising the worse-off).

I myself don't have a theoretical stake in the moral separateness of persons. But if those who care most for and have thought hardest about the moral separateness of persons think that separateness is compatible with *at least some* interpersonal trade-offs, I'm inclined to believe them. So, here I won't be arguing any further for this idea. Let emphasise how minimal the assumption I'm making here is. Again, I'm not arguing, here, in favour of any particular conception of how, precisely, we ought to trade-off value between persons. The Difference Principle's particular claim about how to *value* trade-offs is controversial. That's not surprising, since it amounts to the claim that we should value the difference between the good of the best-off after a (mere) Pareto improvement and their good prior to that improvement at *zero*. But we don't need to accept the Difference Principle's idiosyncratic claim about how to value trade-offs between persons to think it's obvious that, sometimes at least, trade-offs between persons are morally permissible and, most importantly, are not ruled out by the "separateness of persons." And I'm certainly not arguing that *unrestricted interpersonal trade-offs*, that is, the view according to which there are no theoretical restrictions on how we can trade-off value

between persons, is compatible with the separateness of persons. That would be a quite controversial assumption.

What about the corresponding claims in the epistemic case? What does the epistemic separateness of beliefs, understood analogously to the moral separateness of persons, give us reason to think about the permissibility (or not) of trade-offs? Here we come to the crux of things. For notice that, as we've just seen, a commitment to the moral separateness of persons is broadly compatible with interpersonal trade-offs. But then, by analogous reasoning, it seems we should accept trade-offs between beliefs, that is, *interdoxastic trade-offs*.

Here is why. The epistemic separateness of beliefs is best understood as a normative matter just as the moral separateness of persons is best understood as a normative (not metaphysical) matter. In the moral case, separateness entails a set of restrictions on how persons can be treated: restrictions, in other words, on trading-off one's good against another's. But, again, *those restrictions do not constitute absolute bans* on interpersonal trade-offs. That way lies the funny farm, and in any case the proponents of the moral separateness of persons don't themselves sign up for such absolute bans on trade-offs. So, in the epistemic case, the separateness of beliefs should be understood, I'm now busy suggesting, as entailing a set of restrictions on how beliefs can be treated: restrictions, in other words, on trading-off one's truth against another's. But the idea that there are *some* restrictions on interdoxastic trade-offs does not amount to the idea that there is an *absolute ban* on them. Moreover, insofar as we take seriously the analogue from the moral case, we've good reason to think even the most ardent proponents of the epistemic separateness of beliefs will, like their counterparts in moral theory, allow for (restricted) trade-offs.

At the very least, and here's the important point, we've been given no reason by haters to think the situation is substantively different in the epistemic case than it is in the moral. Worse, it appears that the best way of cashing out the otherwise obscure metaphor regarding the separateness of beliefs, namely by analogy with the moral separateness of persons, gives us positive reason for thinking trade-offs will sometimes be permissible. In fact, since haters *themselves* encourage the analogy between the two domains, they thereby encourage the thought that what goes for trade-offs between morally separate persons (they're permissible, with

restrictions) goes for trade-offs between epistemically separate beliefs (they're permissible, with restrictions). Haters therefore face a dilemma.

On the one hand, they can have a plausible conception of the epistemic separateness of beliefs. The most plausible conception—suggested by haters themselves, and modelled on the moral separateness of persons—gives us no independent reason for thinking that there is an *absolute ban* on trade-offs between beliefs. If we understand the epistemic separateness of beliefs as the epistemic analogue of the moral separateness of persons, we should accept that trade-offs will sometimes be epistemically permissible.

On the other hand, haters can reject the conception of the epistemic separateness of beliefs modelled on the moral separateness of persons. I haven't argued, here, that this rejection would be unworkable or even unreasonable. If haters reject the conception of the epistemic separateness of beliefs modelled on the moral separateness of persons, they cannot then remain silent on what the epistemic separateness of beliefs is supposed to amount to. Such metaphors are meaningless unless unpacked. And, importantly, it cannot turn out that what it is for beliefs to be epistemically separate is for it to be such that their truth cannot ever be appropriately traded-off against another belief's truth. That would beg the question against the proponent of the package of EOC and veritism. I'm not optimistic about the prospects for haters giving a meaningful alternative conception of beliefs' epistemic separateness that doesn't simply beg the question in this way; but whether it can be done remains to be seen: if they take this route, the burden is now on haters to tell us what they mean.

This might seem unfair to haters. Berker, at least, is explicit about the fact that he thinks there are structural disanalogies between the moral and the epistemic domains (2013a, b, 2015). And it's reasonable to assume Berker thinks it is these disanalogies that explain why the moral separateness of persons is negotiable and therefore allows for trade-offs whereas the epistemic separateness of beliefs is not, and so does not. After all, it's Berker himself who points to the disanalogy in setting up his claim initially. As he says, *unlike in the moral case*, "in the epistemic case there is no wiggle room" (2013a, 365). But what I am presently urging is that we are owed an explanation for this lack of wiggle room. And the appeal to

the metaphor of the separateness of beliefs (or propositions) is no help at all. At best, it confuses matters.

4 Cases and Cases

Still, you might have worries. Perhaps I'm placing far too much emphasis on a catchy turn of phrase. (In my defence, I'm trying here not just to bury the phrase but to make sure it's dead when I do.) This can seem unfair to the hater. What she cares about isn't defending her use of a bit of language (the "separateness of propositions"), but instead her view about *trade-offs* (they're always forbidden). And after all, the worry continues, Berker has an explanation for the lack of wiggle room in the epistemic case that doesn't appeal to anything about separateness: according to him, epistemic rationality is *non-promotive*, in the sense that it's not concerned with promoting some valuable end such as true belief. So, what's important is not the epistemic separateness of beliefs (propositions), whatever that turns out to mean. What's important (a hater might say) is why we should think that epistemic rationality is non-promotive. So (says our hater), forget about arguments over separateness; instead, focus on arguments over whether epistemic rationality is promotive. In this section, I address this issue.

So: Why do haters think that epistemic rationality is non-promotive, that is, not concerned with promoting some independently valuable end such as true belief? The answer is precisely the one Berker gives: because of our judgements about cases. Recall, again, haters' intuitions, which we're invited to share, about cases such as Berker's:

> **Scientist**: Suppose I am a scientist seeking to get a grant from a religious organisation. Suppose, also, that I am an atheist: I have thought long and hard about whether God exists and have eventually come to the conclusion that He does not. However, I realise that my only chance of receiving funding from the organisation is to believe in the existence of God: they only give grants to believers, and I know I am such a bad liar that I won't be able to convince the organisation's review board that I believe God exists unless I genuinely do. Finally, I know that, were I to receive the grant, I would use

> it to further my research, which would allow me to form a large number of new true beliefs and to revise a large number of previously held false beliefs about a variety of matters of great intellectual significance.

According to Berker, "most of us" would not think that a belief in God was reasonable, or rational, or justified, in these circumstances. As we know, haters' point isn't about how to *value* trade-offs: their point is not that in this case, all things considered, it would be unreasonable to believe in God in order to further one's research (because, say, the research furthered isn't sufficiently important). Instead, their point is that such trade-offs are *always forbidden*. It's important to be clear on this point, so I'll pause over it for a moment.

It can help to put things in terms of reasons. Haters' point is not that the epistemic reasons in favour of believing in God (because it will further one's research) are *outweighed* by the epistemic reasons against doing so (because, say, it's something one thinks is false). Instead, haters' view is that the epistemic reasons in favour of believing in God in this case *do not exist*. We know this is haters' view because their central thought isn't one about how to *value* trade-offs—about how much (important) truth we'd have to put up against how much (important) falsehood for things to come out equals-pequals. If that were haters' central thought, then it'd be utterly mysterious why they think such trade-offs are always forbidden. After all, it'd be possible to imagine the sort of case where the scales tip in favour of the trade-off. But, and again, haters think such trade-offs are always forbidden. In the language of the misleading metaphor from the previous section: the epistemic separateness of beliefs is *non-negotiable*.

But as I'll now explain, this way of arguing for the non-promotive nature of epistemic rationality, and so in favour of an absolute ban on trade-offs, via our judgements about cases, suffers a serious methodological mistake that vitiates its probative value. The mistake is one Mark Schroeder identifies in a different context, and it'll be helpful to work our way into the mistake as made by haters by first looking at the instance identified by Schroeder:[13]

> In a first version of the case, you see Tom Grabit come out of the library, pull a book from beneath his shirt, cackle gleefully, and scurry off.

> Intuitively, you have a reason to believe that Tom just stole a book from the library. But in a revised version of the case, Tom has an identical twin, Tim, from whom you cannot visually distinguish him. If you're aware of this, then it turns out that you don't have any reason to believe that Tom stole a book after all. Right? Of course that's right. This, after all, is a classic case of an undercutting defeater, and as everyone knows, undercutting defeaters make it the case that things that would otherwise have been reasons for you, instead are not.
>
> This is a natural thought about the case, but your intuitions mislead you. That you still do have some reason to believe that Tom stole a book can be observed by comparison with yet a third version of the case. In the third version, Tom and Tim have a third identical sibling, Tam. In this case, you have even less reason to believe that Tom stole a book than in the second, and so in the second it can't have gone away entirely. By similar reasoning, you still have a reason to believe that Tom stole a book even in the third case, because there is a fourth case in which there are four identical siblings and your reason to believe that Tom stole a book is still worse. (2007, 93)

Schroeder's point explicitly concerns undercutting defeat. But in the present context, what's important is the more general lesson we can draw from Schroeder's examples, namely that we can be led into intuitive error about the (non-)existence of (epistemic) reasons by failing to draw the right comparisons between cases—either by using the *wrong* comparison case or by failing to use *any comparison case whatsoever*.[14] That, in short, is the mistake haters make. Let me explain.

If the lesson of Schroeder's Grabit example is that our intuitive verdicts regarding the existence of epistemic reasons can be misled by failing to use the correct comparison case, then failing to use *any comparison case whatsoever* in coming to an intuitive verdict regarding the existence of epistemic reasons—as we're asked to do in, for example, Berker's atheistic scientist case, might equally well mislead. And that, I think, is exactly what's happened. To see this, consider the following case, which I'll call:

> **Inept Scientist**: Suppose I am a scientist seeking to get a grant from a religious organisation. Suppose, also, that I am an atheist: I have thought long and hard about whether God exists and have eventually come to the conclusion that He does not. However, I realise that my only chance of

> receiving funding from the organisation is to believe in the existence of God: they only give grants to believers, and I know I am such a bad liar that I won't be able to convince the organisation's review board that I believe God exists unless I genuinely do. Finally, I know that, were I to receive the grant, I would use it to further my research, which would allow me to form a small number of new true beliefs and to revise a small number of previously held false beliefs about a variety of matters of minimal intellectual significance.

Note that Inept Scientist is exactly like Berker's Scientist, except that, whereas in Scientist the grant from the religious organisation allowed me to form a *large* number of new true beliefs and revise a *large* number of previously held false beliefs about a variety of matters of *great* intellectual significance, in Inept Scientist, I form a *small* number of new true beliefs and revise a *small* number of false beliefs about matters of *minimal* intellectual significance. I hereby invite you to think (presumably in agreement with haters) that forming a belief in God's existence would not be epistemically reasonable in these (Inept Scientist) circumstances. But that all-out verdict, that it would be epistemically unreasonable to form a belief in God's existence in these circumstances, isn't the verdict I'm interested in exploring.

Instead, consider the essentially *comparative* judgement about Inept Scientist as opposed to Scientist. We should agree that it is *less* reasonable to believe in God's existence in Inept Scientist than it is in Scientist. If you're not convinced, consider a *third* version of the case that is exactly like Inept Scientist except that one doesn't form *any* new true beliefs (and so only revises a small number of previously held false beliefs about matters of minimal intellectual significance). Or consider a *fourth* version of the case that is just like the third version but where still *fewer* false beliefs are revised. In each of these successive cases, the correct intuitive verdict seems to be that one is progressively *less justified* or *less reasonable* in forming a belief in God's existence in the circumstances one finds oneself in.

Let's put this again in the language of reasons. The intuitive verdict about Inept Scientist is that you have *less* reason to believe in God's existence in those circumstances than you do in Scientist. And similarly for the third and fourth versions of the case: these are cases where you have

progressively less and less reason to believe in God's existence. But if this is so, then in Scientist you must have *some* reason to believe in God's existence. This is because in Inept Scientist we know that your reason to believe hasn't gone away entirely; and similarly for the third and fourth versions of the case: you still have reason to believe in God's existence in Inept Scientist because in the third (and fourth) version of the case you have *less* (and still less). And so examples such as Scientist do not put the lie to a promotive conception of epistemic rationality, since when we consider such examples in comparison to other cases, we see that there *is* some epistemic reason in such cases, which is what the promotive conception of epistemic rationality predicts.

In sum, then: haters' view that epistemic trade-offs are always forbidden is not meant to be tied to verdicts about all-things-considered rationality in particular cases. Their view is that there is an *absolute* ban on trade-offs because epistemic rationality is non-promotive. So, it's not supposed to be essential to the force of the examples meant to support this view that we agree with haters about the all-things-considered verdict in any given case. What's supposed to be essential is the intuitive verdict that in the relevant cases there is no epistemic reason *at all* in favour of the belief in virtue of (potential) trade-offs. Let's suppose we grant, for the sake of argument, that we share this intuitive verdict in one-off cases. This does not show what it is supposed to show, that we share haters' view that epistemic rationality is non-promotive. This is because when we *contrast* our verdict about such cases with suitably modified versions of the cases, the intuitive verdict there is that there is *still less* reason in these modified cases—even if the *overall* verdict (i.e., the judgement about whether the trade-off is here permissible, whether it would be reasonable or justified) is the same. But if there is still less reason in the modified cases, then there must be *some* reason in the unmodified cases, contrary to our initial intuition. That is Schroeder's lesson. So, in the cases (modified and unmodified alike) there must be some reason, in virtue of the fact that doing so would promote true belief (and promote the avoidance of false belief), to believe. And so epistemic rationality is (at least in part) promotive, contrary to haters' claim.

I'll close by considering one final rejoinder. Perhaps this is still unfair to haters—and especially to Berker—since it might seem to rely on the

details of particular examples, and in other work Berker explicitly eschews such reliance as the right way to understand his arguments:

> I do not mean my argument to rest too heavily on these particular examples. The important point is not the examples themselves, but rather the structure of the examples. As I see it, my […] argument against epistemic consequentialism involves not a series of counterexamples, but rather a *recipe* for generating counterexamples. […] this recipe makes my style of argument particularly versatile: even if fault is found with the specific examples I have offered in this essay, my recipe allows me to find other examples of a structurally similar sort that suffice to make the same point. (2013b, 378; emphasis in original)

But what I've just argued is that there is a *structural flaw* in Berker's recipe, and so it can't be used to cook up the results he wants. He's failed to recognise that some ingredients in it, to taste right in the cake, are essentially comparative: the sugar must be weighed *against* the flour. Less culinarily: our intuitive verdict about any single case that believing *p* is unreasonable despite the fact that so believing would best promote overall epistemic value doesn't show that epistemic rationality is non-promotive. This is because, as we just saw, when we compare our verdict about such a case to one where believing *p* would *less well* promote the epistemic values, our judgement is that it would be *even more unreasonable* to believe *p* in that case. And that means, again as we just saw, that it must be that there is *some* reason to believe *p* in the first case—and so epistemic rationality must have *some* promotive component (for that is what explains the existence of the reason). At best, then, what our intuitive verdicts about these singleton, non-comparative, cases suggest is that there are *some* restrictions on epistemic trade-offs; our intuitive verdicts don't deliver the result that such trade-offs are absolutely forbidden.

5 Summary

Haters object to the pairing of veritism and EOC on the grounds that it ignores the epistemic separateness of beliefs (propositions) and thereby allows for objectionable epistemic trade-offs. I've argued here the

metaphor of separateness of beliefs is at best inert, at worst misleading. It's misleading if unpacked by analogy with the moral separateness of persons. This is because, in that domain, even the most ardent proponents of the separateness of persons don't think that separateness yields a ban on trade-offs. It's inert if it remains a metaphor: it can't do any argumentative work.

If the argumentative work is done elsewhere, it appears to be done by appeal to our intuitive verdicts about cases such as Berker's Scientist. But those verdicts are not probative, for they fail to recognise that our intuitive verdicts about the existence of epistemic reasons in particular cases can mislead when those verdicts aren't tested against the right (or indeed *any*) comparison cases. Once we introduce the relevant comparison cases, we see that our intuitive verdicts tell in favour of a promotive conception of epistemic rationality, that is, one where trade-offs are at least in principle permissible.

Let me close with a remark that's more confession than argument, but which is short. It makes pre-theoretical sense to me to think the activity of trading off persons' goods requires special justification. After all, we're dealing with a *person* and her welfare; of *course* special justificatory caution is required in circumstances where we're deliberately considering *harming* a *person* (even when it's to the benefit of others). At least, I have a sense for why I might feel this way, pre-theoretically, even if I end up in a position where I say otherwise. Why might this be? For one thing, and quite obviously, *I* am a person, and *I* do not particularly want to suffer harm even for others' benefit. For another, many of those who I care about most are *also* persons, and I do not much want that *they* be cut up to save the many. And even if I end up where the consequentialist in the practical domain wants me to be, it's important that I recognise the need to be *talked into it*. But, and maybe this is just my cold and unsympathising epistemic nature, I'm at a loss to make any good pre-theoretical sense out of the idea that trading off one belief's truth against another's requires, in a similar way, any sort of special justification. After all (I'm inclined to say) we're (just) dealing with a *belief*. Sure, the thought continues, I'm considering *believing falsely* in order to believe truly about some other matters. But it is just belief, just some mental state of mine; *I* am not the belief, nor are any who I care about most themselves beliefs. So, unlike in

the practical case, I don't feel the need to be *talked into* the position the epistemic consequentialist wants me to be in: I can happily start from there, and need to be convinced otherwise.

Notes

1. The issues here—terminological and otherwise—are fraught. For more, see Bradley (2006), Zimmerman (2001) and Korsgaard (1983). For present purposes, my use is simply stipulative.
2. See, for perhaps the most famous instance, Mill (1979). For contemporary versions of roughly this sort of hedonism, see Feldman (2004), Bradley (2009) and Heathwood (2006).
3. "Correctness" is a placeholder for other normative terms, e.g., "obligatory," "right," etc. These two ideas *don't* tell us plenty, such as how to "weigh up" pleasures and pains, what it means for an outcome to be available to an agent at a time, whether our obligations are to be understood as *actualist* or *possibilist* (c.f. Jackson and Pargetter 1986), and so on. We can safely ignore these complications.
4. See Berker (2013a, b, 2015); Goldman (1999, 2015). This is what Pritchard (2010) calls "T-monism," Zagzebski (2004) calls "epistemic value monism" and Ahlstrom-Vij dubs "veritistic value monism" (2013). The "monisms" here highlight clause 3 in our definition.
5. Here, I follow the opponents of EOC, especially Berker (2013a). More on this in a moment. As with the pairing of hedonism and POC, there is plenty we can't read off the combination of veritism and EOC. But, again, we can safely ignore these complications. For doubts about the parallel between the two pairings, see Horowitz (forthcoming).
6. On theoretical grounds, as in Littlejohn (2015, forthcoming) and Berker (2013a, b, 2015), and on empirical grounds, as in Andow (2017).
7. Berker himself borrows this example from Fumerton (2001) who in turn adapts it from one due to Firth (1998).
8. For a few cases, see Singer (2018), Ahlstrom-Vij and Dunn (2014), Goldman (2015), Dunn and Ahlstrom-Vij (2017); and for a kind of defence against empirical results reported in Andow (2017), see Roberts et al. (2017).
9. Rather than going this route, you might try to offer a *neutral* characterisation of the epistemic separateness of propositions in terms of their

alethic or probabilistic separateness (i.e., either that the propositions don't stand in entailment relations or that they are probabilistically independent). But this will not do, for obvious reasons: even if two propositions stand in the relevant entailment or probabilistic relations, haters will presumably think trading off one's truth against another is forbidden.

10. Hedonism doesn't *quite* say this: it says that pleasure (*whoever's*) is finally practically valuable. For present purposes, we can ignore this complication and assume that the only relevant bearers of pleasure are persons. This is because, when it comes to the epistemic analogue of hedonism, i.e., veritism, the only relevant bearers of *truth* are *beliefs*.
11. You might think that in the present context we distinguish between token beliefs on the basis of their propositional content. Thus the "separateness of propositions" and the "separateness of beliefs" comes to one. But that will not do. For veritism says that *your* true belief that *P* is as much a bearer of epistemic value as *my* true belief that *P*.
12. In fact Norcross argues for trade-offs without bound. But here I need only the weaker assumption.
13. In unpublished work, Daniel Wodak marshals this same diagnosis of a symmetrical methodological mistake in a different domain, viz. arguments regarding the reasons we have on the basis of normative (and in particular aesthetic) testimony. I am in complete agreement with what he says there, which is, roughly, that arguments for the conclusion that we lack reasons for attitudes on the basis of normative testimony that go by way of our intuitive verdicts about the wrong (or no) comparison cases are not probative.
14. Schroeder himself uses the case to draw a lesson about out intuitions regarding reasons quite generally, including our intuitions about *practical* reasons. See Schroeder (2007, 93 and following).

References

Ahlstrom-Vij, Kristoffer. 2013. In Defense of Veritistic Value Monism. *Pacific Philosophical Quarterly* 94 (1): 19–40.

Ahlstrom-Vij, Kristoffer, and Jeffrey Dunn. 2014. A Defence of Epistemic Consequentialism. *Philosophical Quarterly* 64 (257): 541–551.

Andow, James. 2017. Do Non-philosophers Think Epistemic Consequentialism is Counterintuitive? *Synthese* 194 (7): 2631–2643.

Berker, Selim. 2013a. Epistemic Teleology and the Separateness of Propositions. *Philosophical Review* 122 (3): 337–393.

———. 2013b. The Rejection of Epistemic Consequentialism. *Philosophical Issues* 23 (1): 363–387.

———. 2015. Reply to Goldman: Cutting Up the One to Save the Five in Epistemology. *Episteme* 12 (2): 145–153.

Bradley, Ben. 2006. Two Concepts of Intrinsic Value. *Ethical Theory and Moral Practice* 9 (2): 111–130.

———. 2009. *Well-Being and Death*. Oxford: Oxford University Press.

Dunn, Jeffrey, and Kristoffer Ahlstrom-Vij. 2017. Is Reliabilism a Form of Consequentialism? *American Philosophical Quarterly* 54 (2): 183–194.

Feldman, Fred. 2004. *Pleasure and the Good Life: Concerning the Nature, Varieties, and Plausibility of Hedonism*. Oxford: Clarendon Press.

Firth, Roderick. 1998. Epistemic Utility. In *In Defense of Radical Empiricism: Essays and Lectures*, ed. John Troyer, 317–334. Lanham, MD: Rowman & Littlefield.

Fumerton, Richard. 2001. Epistemic Justification and Normativity. In *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*, ed. Matthias Steup, 49–60. Oxford: Oxford University Press.

Gauthier, David. 1963. *Practical Reasoning the Structure and Foundations of Prudential and Moral Arguments and Their Exemplification in Discourse*. Oxford: Clarendon Press.

Goldman, Alvin. 1999. *Knowledge in a Social World*. Oxford: Oxford University Press.

———. 2015. Reliabilism, Veritism, and Epistemic Consequentialism. *Episteme* 12 (2): 131–143.

Heathwood, Chris. 2006. Desire Satisfactionism and Hedonism. *Philosophical Studies* 128 (3): 539–563.

Horowitz, Sophie. forthcoming. Epistemic Value and the Jamesian Goals. In *Epistemic Consequentialism*, ed. Kristoffer Ahlstrom-Vij and Jeffrey Dunn. Oxford: Oxford University Press.

Jackson, Frank, and Robert Pargetter. 1986. Oughts, Options, and Actualism. *Philosophical Review* 95 (2): 233–255.

Korsgaard, Christine. 1983. Two Distinctions in Goodness. *Philosophical Review* 92 (2): 169–195.

Littlejohn, Clayton. 2015. Who Cares What You Accurately Believe? *Philosophical Perspectives* 29 (1): 217–248.

———. forthcoming. The Right in the Good: A Defense of Teleological Non-consequentialism in Epistemology. In *Epistemic Consequentialism*, ed. Kristoffer Ahlstrom-Vij and Jeffrey Dunn. Oxford: Oxford University Press.

Mill, J.S. 1979. *Utilitarianism*. Indianapolis: Hackett.

Nagel, Thomas. 1970. *The Possibility of Altruism*. Oxford: Clarendon Press.

Norcross, Alastair. 2009. Two Dogmas of Deontology: Aggregation, Rights, and the Separateness of Persons. *Social Philosophy and Policy* 26 (1): 76–95.

Nozick, Robert. 1974. *Anarchy, State and Utopia*. New York: Basic Books.

Pritchard, Duncan. 2010. Knowledge and Understanding. In *The Nature and Value of Knowledge*, ed. Adrian Haddock, Alan Millar, and Duncan Pritchard, 3–88. Oxford: Oxford University Press.

Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.

———. 1995. *Political Liberalism*. New York: Columbia University Press.

Roberts, Pendaran, James Andow, and Kelly Schmidtke. 2017. Lay Intuitions About Epistemic Normativity. *Synthese*: 1–21.

Schroeder, Mark. 2007. *Slaves of the Passions*. Oxford University Press.

Singer, Daniel J. 2018. How to Be an Epistemic Consequentialist. *Philosophical Quarterly* 68 (272): 580–602.

Vallentyne, Peter. 2006. Against Maximizing Act-Consequentialism. In *Contemporary Debates in Moral Theories*, ed. James Dreier, 21–37. Oxford: Blackwell.

Zagzebski, Linda. 2004. Epistemic Value Monism. In *Ernest Sosa and His Critics*, ed. Greco John, 190–198. Oxford: Blackwell.

Zimmerman, Michael J. 2001. *The Nature of Intrinsic Value*. Lanham, MD: Rowman & Littlefield.

Part II

Epistemic Anti-realism



Knowledge, Reasons, and Errors About Error Theory

Charles Côté-Bouchard and Clayton Littlejohn

1 Introduction

Readers are probably familiar with this objection to the moral error theory. According to the moral error theorist, our moral claims are systematically in error. If our moral claims are systematically in error, we should expect the same would be true of our epistemic claims. The moral error theory, in other words, entails the epistemic error theory. This, however, is a step too far. The epistemic error theory cannot be true. Thus, the moral error theory cannot be true, either.

This line of objection treats the epistemic error theory as the *absurdum* of a *reductio*.[1] Although we don't accept the epistemic error theory, we think that this objection to the moral error theory rests on a mistake. The objection paints the moral error theorist as someone who denies the

C. Côté-Bouchard (✉)
Department of Philosophy, Rutgers University, New Brunswick, NJ, USA
e-mail: charles.cote79@gmail.com

C. Littlejohn
Department of Philosophy, King's College London, London, UK
e-mail: clayton.littlejohn@kcl.ac.uk

© The Author(s) 2018
C. Kyriacou, R. McKenna (eds.), *Metaepistemology*, Palgrave Innovations in Philosophy, https://doi.org/10.1007/978-3-319-93369-6_7

existence of a certain kind of reason and who is therefore committed to denying the existence of epistemic facts. We think that denying the existence of the relevant kind of reasons does not commit us to the epistemic error theory. Epistemic realism doesn't require the relevant kind of reason.

In Sect. 2, we shall present the companions in guilt objection to the moral error theory. In Sect. 3, we shall flesh out the kind of normativity and reasons that concern error theorists and realists. In Sect. 4, we shall argue that we should not think of epistemic norms as robustly normative. In Sect. 5, we shall explain why this does not undermine the error theorists' claim that morality is robustly normative.

2 Error Theories and Companions in Guilt

Error theories about some domain of thought and discourse D standardly take the following form:

> Conceptual Claim: It is conceptually necessary that D-claims represent facts with some feature F (the truth of D-claims turns on whether there are F-facts; they would be systematically in error if there are no F-facts, but not otherwise).
>
> Ontological Claim: There are no F-facts; nothing has feature F.
>
> Conclusion: D-claims are systematically in error.

In metaethics, error theorists argue that *moral* claims are systematically mistaken because they necessarily represent facts that would have to have a feature that nothing in our universe has (Garner 1990; Joyce 2001, 2006; Mackie 1977; Olson 2014; Streumer 2017). According to most error theorists following Mackie (1977), that feature has to do with the kind of *normativity*, *authority*, *prescriptivity*, or *to-be-doneness* that moral facts would have to have (see especially Garner 1990; Joyce 2001; Olson 2011, 2014). As Garner explains:

> It is hard to deny that the recognition of a moral fact (say the wrongness of an action), if such a thing were possible, would give us a justifying reason for refraining from performing that action. […] If we recognize a moral fact, then we recognize it as a moral fact, which is to say that we recognize it as directing us to be or act one way rather than another—and we recog-

> nize this whether or not we are moved to obey the directions. It is the peculiar combination of objectivity and prescriptivity […] that makes moral facts and properties queer […]. (Garner 1990, 143)

The problematic commitment of our moral discourse and thought, then, is a commitment to the existence of a certain kind of normativity and normative reason. It's difficult to state precisely what the problematic reasons are like and we shall try to get clear on this in the next section.

In response to this kind of moral error theory, Cuneo offers a "companions in guilt" argument based on the apparent parity between ethics and *epistemology*. Epistemic facts, he argues, are no less normative and reason-implying than moral facts. The error theorists' conceptual claim about moral thought and discourse is also true of epistemic thought and discourse: they essentially represent facts with the same relevant kind of normativity. But if, as error theorists maintain, there are no facts with that kind of normativity, then there are no epistemic facts either. The moral error theory therefore entails the epistemic error theory. This is what Cuneo calls the Parity Principle.

The problem, the objection goes, is that the epistemic error theory cannot be true. This is because it faces a fatal dilemma:

> [E]pistemic error theory is either (i) self-defeating insofar as it presupposes the very sort of entities it says do not exist or (ii) implies wholesale epistemological scepticism, according to which we have no epistemic reason to believe anything, including the error theory itself. An implication of this second horn is that epistemic error theory would be polemically toothless if true, as no one could make a rational mistake in rejecting the position. Since neither of these options are palatable […] we have powerful reason to reject the epistemic error theory. (Cuneo 2016, 78)

Therefore, since it entails the epistemic error theory, we should reject the moral error theory. This objection can be summarised as follows:

The Companions in Guilt Argument

1. If the moral error theory is true, then the epistemic error theory is true (Parity Principle).
2. The epistemic error theory is false.
3. Therefore, the moral error theory is false.

Error theorists have two main options in response to this objection. The first is to concede that moral and epistemic error theories are "companions," but deny their "guilt." It is, in other words, to attack the second premise and defend the epistemic error theory against Cuneo's dilemma. The second is to deny companionship and reject the Parity Principle. Whether or not Cuneo is correct about the "guilt" of the epistemic error theory, such an error theory is not entailed by the moral error theory in the first place. In what follows, we offer a version of the second strategy.

3 Normativity Thick and Thin

What error theorists reject is a particular kind of problematic normativity, one that involves a certain kind of categorical reason. Getting clear on the distinction between the unproblematic and the contentious is the business of this section.

Cuneo represents the disagreement between the moral error theorists and the moral realists as a disagreement about "robust" or "categorical" reasons. Some might treat these notions as interchangeable, but since there are different ways of drawing the distinction between the categorical and non-categorical, we think it will pay dividends to proceed carefully.

To start, let's look at the commitments of moral and epistemic realism, as Cuneo conceives of them. He thinks that the moral realist should accept some platitudes about morality. Some have to do with the content of morality (for example, that it cannot be wrong to watch Foot's hedgehogs in the moonlight). The platitude that interests us is what Cuneo calls the *Authority* platitude:

> [M]oral […] facts are or imply excellent reasons for us to behave in certain ways, regardless of our desires, goals, or the social roles we may occupy. In terminology that has become popular, these facts are or imply 'robust reasons.' (Cuneo 2016, 74)

Cuneo thinks that this feature of moral facts is what the moral realist and error theorist disagree about. He sees them as agreeing that the existence of moral facts turns on whether there are features of the world that

fit these descriptions and then disagreeing about whether there are such things.

As Cuneo sees things, the error theorist's problem with moral realism is that the moral realist is committed to the idea that there could be features of the world that are authoritative in the way suggested above. To show that there can be such features, he turns to the epistemic domain. He thinks that the epistemic realist has similar commitments. The epistemic realist, he thinks, accepts an epistemic authority platitude as with our conception of moral facts; it's partly constitutive of our conception of epistemic facts and claims that they have robust normative authority. If, as the epistemic realist believes, there truly are such features of the world, the moral error theorist's view must be mistaken.

We think that the connection between epistemic realism and the thesis that epistemic facts are robustly normative isn't as clear as Cuneo believes. Just to put our cards on the table, we think our epistemic speech can correctly pick out epistemic facts even if there are no robustly normative facts and reasons. The authoritativeness that interests Cuneo and the error theorists is connected to the existence of certain kinds of categorical reasons. They account for the inescapability that the moral realist associates with moral facts and the reasons that they provide, so in assessing the Parity Principle we should remember that the issue that concerns us is whether we should believe that the epistemic realist is committed to the existence of such reasons. The commitment isn't trivial, recall, because the issue is whether the existence of such reasons is entailed by epistemic realism.

One way of distinguishing the categorical from the non-categorical reasons is in terms of a connection to an individual's contingent desires or ends. The categorical reasons apply to everyone irrespective of their particular contingent ends or interests. If it's true that Agnes has a reason to work on her dribbling or practice her scales, it's probably only true because she wants to become a better midfielder or improve as a cellist. Moral reasons aren't supposed to be like this. If she loses her love of football, it might no longer be true that she has this reason. If she loses her love for her fellow persons it's still supposed to be true, the moral realist says, that she has good reason to promote their interests, keep her promises, and so on.

This isn't a terribly interesting understanding of categoricity. The reasons of etiquette and the law are like this, but they aren't supposed to have the kind of authority that the moral realist thinks moral reasons have. So, we need a more robust notion of robustness and categoricity to capture what's at issue. Brink (1992) distinguishes this first superficial kind of categoricity from more interesting notions. Let's focus on his understanding of authority. We might think that what makes moral reasons interesting is not that they apply to everyone irrespective of their contingent ends or desires but that, in some sense, it is irrational for an agent not to be moved by them even if heeding them wouldn't suit the agent's contingent ends or desires. This is how Brink understands the authoritativeness of moral reasons.

In this respect, it might seem that the realists' epistemic and moral reasons would be similar. The epistemic realists agree, presumably, that a thinker's beliefs might be irrational for reasons that have nothing to do with the thinker's contingent ends or inclinations. The moral realists agree, presumably, that a thinker's actions might be irrational for reasons that have nothing to do with the thinker's contingent ends or inclinations.

The similarities here might be purely superficial. Amongst those who seem to think that morality is a source of normative pressure that's different to, say, etiquette and accept something like the moral authority platitude are philosophers who caution us against using the term "irrational" too cavalierly. We shouldn't just use it so that we have something that we can say about the wicked to "hit him where it will hurt," as Foot (1978, 152) puts it. There is a special kind of unreasonableness that's supposed to be manifest when a wicked person does some unspeakable thing for no good reason. If we let such things give us the relevant sense of "irrational," it's hard to believe that anyone has taken seriously the idea that *this* is the kind of irrationality a person can be charged with if, say, they don't proportion their beliefs to the evidence. Let's consider a slightly different case, the case of instrumental rationality.

The standards of instrumental rationality are different to those of, for example, etiquette because while both are *categorical* in our first superficial sense, the standards of etiquette lack authority. But the standards of instrumental rationality are also supposed to differ from those of morality.

This difference isn't helpfully brought out by talking about standards that we'd be irrational to violate because that's true of the standards of instrumental rationality. When moral realists tell us that they think moral reasons are robustly normative, we think they want to distinguish the requirements of morality from a larger set of requirements of practical rationality. Consider the enkratic requirement and the requirements of instrumental reason:

1. If an agent judges that she ought to ϕ but doesn't intend to ϕ, she's irrational;
2. If agent intends to ϕ and believes that it's impossible for her to ϕ unless she ϕ_2s, she's irrational if she doesn't intend to ϕ_2.

These requirements display a problematic kind of generality. They seem to apply to all individuals whatever specific intentions and beliefs they have. If they hold with sufficient generality, they tell us that people shouldn't be akratic and shouldn't be instrumentally irrational but such requirements are non-substantive in an important sense. However silly or trivial our ends might be, we can violate them. In this way, they seem to differ in an important respect from the violation of moral standards.

What's missing, we think, is something about a standard that cannot be recovered from the observations about who or when the standard applies. It has to capture the thing that makes morality and its requirements important in the way that formal requirements of rationality of the kind just mentioned might not be. One suggestion might be that there are values associated with the moral reasons that call for certain responses and make us criticisable if we're not invested or committed to serving these values well. Such values might be lacking in the case of the rational requirements just mentioned. Let's label a difference here between *thin* and *robust* normative standards and reasons to conform to them.[2] We think that epistemic norms, like the requirements of rationality just mentioned, are thinly normative, not robustly so. And we think that this is sufficient for epistemic realism even though moral realism requires robust normativity.

When it comes to these thinly normative standards we think that it's important to appreciate that their form, shape, or content isn't a good guide to understanding the normative force or pressure that thinkers are under.

We think that pairs of agents, say, might be subject to different kinds of criticism for failing to meet some single standard of practical rationality because, say, there are differences in the substantive reasons that apply to them. Compare the case where Agnes violates the instrumental principle while playing pushpin to a case where she violates this trying to find a way to pay for her daughter's surgery. If we proceed in this way, we can make progress by asking whether epistemic realism requires reasons more robust than those associated with instrumental rationality or if our epistemic discourse commits us only to reasons that are thinly normative in a way moral reasons aren't supposed to be.

4 The Thin Epistemic Normativity Strategy

4.1 Thin Epistemic Normativity

We think moral error theorists should use the distinction between thin and robust normativity to reject Cuneo's parity premise. Instead of defending the epistemic error theory, they should reject the idea that epistemic facts are robustly normative in the first place. Instead, they should maintain that epistemic thoughts and claims are only thinly normative. Although they essentially imply norms or standards, they do not imply robust reasons. To put the point another way, although the authority platitude is true of morality, it is not true of epistemology. It is not an essential part of our conception of epistemic facts that they are robustly normative.

We can think of two reasons to suggest that epistemic norms are thin in a way that moral norms are not. The first has to do with the possibility of knowledge without accountability. The second has to do with some surprising features of the epistemic "ought."

Animals are not subject to moral standards. Your dog's actions cannot be morally right or wrong. More generally, since animals lack the rational capacities needed for agency and attributions of responsibility, they are exempt from robustly normative standards. They cannot have robust reasons that apply to them and the "oughts" that require such reasons do not apply to them.

At the same time, non-human animals can have knowledge. Their beliefs can therefore be epistemically successful or defective even if we wouldn't say that such attitudes could be justified or that these animals have met their intellectual responsibilities. That they can meet or fail to meet certain epistemic standards would be puzzling if epistemic standards were robustly normative. Since animals are exempt from robustly normative standards, they could not have knowledge if knowledge were like moral rightness in requiring conformity to robust reasons. So, another reason why thin epistemic normativity is more plausible than thin moral normativity is that unlike moral standards, epistemic standards can apply to animals even though they cannot be held accountable.[3]

There is a point, often credited to Harman (1986), which is also important to consider. He thinks that we aren't subject to criticism for failing to fill our heads with as many of the obvious consequences of uninteresting claims as we can. If you would see, for example, with just a moment's reflection that something is logical consequence of the fact that the state beverage of Delaware is milk, it's unlikely that we'd think that this is something you ought to believe. Notice that this isn't because your epistemic position with respect to this claim might be weak. Far from it. It might be better than many of the things that you may believe. It might be better than many of the things that you ought to believe. Still, if you don't believe it, it's unlikely that you can be subjected to criticism for this.

This observation seems to indicate a strange difference between the moral "ought" and the epistemic "ought." In any plausible moral outlook, there will be some things that the agent ought to do that have nothing to do with furthering this agent's non-moral ends but it isn't clear that there is any purely epistemic relation between a thinker and a proposition in virtue of which the thinker ought to believe this thing. In this regard, the connection between belief and epistemic standards seems more akin to the connection between arrangements of cutlery and the standards of etiquette than the connection between actions and moral standards. Think about the kinds of things that it seems a thinker ought to believe. If someone who is generally reliable confides in you that they've been abused, it's probably the case that you ought to believe them. And now compare the strength of epistemic position here to the strength of epistemic position in the case of the logical

consequences of the boring belief about Delaware's beverage. We doubt that the reason that one belief is in the ought to be believed pile and the other is not is down to a difference in strength of position. The difference seems to be one of importance.

How should importance be understood? We doubt very much that the importance can be understood epistemically. The beliefs are, for all that's been said, equally likely, equally true, and equally good candidates for knowledge. The difference, we submit, is that the most compelling cases of the ought to be believed are cases where two things happen. First, the relevant belief is epistemically kosher. Second, there are non-epistemic reasons to have this belief. They might be non-epistemic reasons to enquire, to settle a question, to trust your informants. We think that the force of the criticism that we level against a thinker for failing to believe what she ought to believe comes from non-epistemic reasons—reasons to form beliefs that meet epistemic standards.

In general, it's hard to see that epistemic standards generate reasons that require us to get into the belief-game. It's interesting to note, however, that epistemic standards generate directives for us even if we're not interested in settling some question correctly. If all the evidence indicates that the state beverage of Delaware is gin, you shouldn't believe that it is milk. There appears, then, to be a weird asymmetry here that we don't find in the moral case. There doesn't seem to be any value associated with this standard that calls for us to get in the belief-game. At best, the standards flag potential moves in that game as incorrect, unwarranted, et cetera. When we criticise people for failing to form beliefs that meet a standard, it seems to be that that's because of some non-epistemic reason for them to try to form a belief on some issue.[4]

This suggests that although the standard that we use to evaluate a belief might be epistemic and although it might apply to all thinkers at all times, its application doesn't tell us whether a thinker should care about whether they have the sanctioned or right attitudes. We think that when a thinker should care about having such attitudes, it's not purely down to meeting an epistemic standard; rather, it's partially down to non-epistemic factors.

In this regard, epistemic reasons are quite different to moral reasons. Moral reasons are the kinds of things that if you're not committed to serving them to the best of your abilities you are criticisable for this reason.

Perhaps this is because, as Raz (2014) has suggested, practical reasons are connected to a good that calls for a response of some kind and epistemic reasons are not inherently connected to any such good. Once we see this, we can see that epistemic reasons must be thinly normative in a perfectly good sense—a thinker's indifference with respect to meeting certain standards isn't itself grounds for criticism.

So, unlike moral facts, epistemic facts are only thinly normative. Unlike moral norms, and just like etiquette, fashion, and the law, epistemic norms are such that there is not necessarily a robust reason to conform to them. Consequently, for the same reason that it does not entail a legal error theory or a fashion error theory, the moral error theory does not entail an epistemic error theory. Moral claims are systematically mistaken, Garner, Joyce, and Olson tell us, because they essentially aim at representing robustly normative facts, which are non-existent. Etiquette, fashion, and the law, in contrast, do not embody the same mistake. Merely thinly normative facts about what is correct or incorrect relative to some standard are not objectionably strange. So, facts about what is what is legal, fashionable, decorous, polite, or grammatical are not strange. Legal claims and fashion claims are thus often true, even in the error theorist's normativity-free universe. We think error theorists should say the same about the epistemic domain. Facts about what is known, epistemically justified, epistemically rational, and the like are only thinly normative, and therefore not strange. Epistemic claims are thus often true, even in the error theorist's ontology. The moral error theory does not entail the epistemic error theory.

Many will recoil at the thought that epistemic normativity might be thin like legal or etiquette normativity. What about the authority platitude? Isn't it just part our ordinary conception of epistemic facts that they imply robust reasons? To view epistemology as only thinly normative is to reject that authority platitude. Consequently, doesn't our proposal fail to capture our conception of epistemic facts and claims? Isn't it, at best, revisionist?

We do not think so. As far as epistemology is concerned, the authority platitude is not, in fact, a platitude. Robust normativity is *not* part of our conception of epistemic facts and claims. A thin conception of epistemic normativity suffices to capture our conception of the epistemic domain. This is because, first, although thin epistemic normativity is incompatible

with the authority platitude in epistemology, it still accommodates a number of closely related platitudes—call this set of closely related platitudes the *epistemic norms* platitudes. Second, although they do not include the authority platitude, the epistemic norms platitudes suffice to capture our conception of epistemic facts and claims. Therefore, thin epistemic normativity suffices to capture our conception of the epistemic domain. Error theorists can thus plausibly deny the authority platitude in epistemology.

4.2 The Epistemic Norms Platitudes

The epistemic evaluations platitude. Essential to the epistemic domain are attributions of positive and negative epistemic evaluations or, as Cuneo calls them, epistemic merits and demerits. Epistemic thought and discourse essentially involve, for example, evaluations of beliefs as knowledge, as epistemically justified, epistemically rational, and the like. It essentially involves, in other words, viewing things like beliefs as epistemically good or bad.

This feature does not require robust normativity, however. Evaluations, merits, and demerits can very well be thinly normative. The ascription of merits and demerits is also essential to legal discourse, etiquette discourse, and other paradigmatic thinly normative domains. Essential to those domains are evaluations of acts as lawful or unlawful, legal or illegal, decorous or indecorous, polite or impolite, stylish or unfashionable, et cetera. Part of making legal claims, etiquette claims, and fashion claims, in other words, is to evaluate them as good or bad from a legal, etiquette, or fashion standpoint. A thin conception of epistemic normativity has therefore no problem capturing the epistemic evaluations platitude.

The importance platitude. Epistemic norms, values, and goals seem to matter. Getting things right and avoiding error is undoubtedly important for human beings. Epistemic standards thus seem to be a very important kind of standards. But this is compatible with such standards being only thinly normative. Thinly normative standards can very well be highly important. After all, we care deeply about many norms and merits that

are also paradigm cases of thinly normative standards. The law, etiquette, and many other conventions are highly important domains. Yet they are not robustly normative.

The deontology platitude. The epistemic domain seems, at least in part, deontological. Epistemology, many think, is essentially about what doxastic attitudes we are epistemically permitted, prohibited, or obligated to have. Talk of permission, prohibition, and obligation appears essential to epistemic discourse.[5] This third platitude is also compatible with our proposal. Just like values and merits, not all permissions and prohibitions are robustly normative. Many thinly normative domains are deontological. The law, etiquette, grammar, games, and many other sets of norms are best put in terms of permissions, prohibitions, and requirements.

The inescapability platitude. You cannot opt out of epistemic evaluations and requirements by not caring about them or by not caring about getting things right. Whether your belief counts as knowledge, as epistemically justified, as epistemically rational, and the like does not depend on what you want or care about. But thinly normative standards can very well be inescapable or desire-independent in that way. Your acts are legal or illegal, decorous or indecorous, and the like whether or not you care about being law-abiding or decorous. Epistemic norms are no different according to our proposal. Even though they are only thinly normative, they apply to you whether or not conforming to them would promote your desires, interests, values, and the like.

The epistemic reasons platitude. Epistemic facts plausibly imply epistemic reasons. If believing that p is epistemically justified or unjustified, then there is an epistemic reason to believe that p or not to believe that p. Moreover, epistemic reasons seem inescapable. Whether there is an epistemic reason for you to form a particular doxastic attitude does not depend on what you care about. Whether or not you care about getting things right or avoiding error, there are epistemic reasons for you not to, for example, engage in wishful thinking.

But this fifth feature of the epistemic domain is also compatible with a thin conception of epistemic normativity. This is because, although they do not entail *robust* reasons, thinly normative facts still imply thin, domain-specific reasons. As Olson points out:

> '[O]ught' and 'reason' are both polysemous terms. […] For example, it seems perfectly fine to say that according to Catholicism there is *reason* to abstain from eating meat on Fridays and to mean by this simply that it is incorrect according to the rules of Catholicism to eat meat on Fridays. […] The claim that there are reasons for writers in English not to split the infinitive might simply mean that splitting the infinitive is inappropriate according to (some) rules of grammar; the claim that there are reasons for male guests to wear a tie at formal dinners might simply mean that this is required by the rules of etiquette. (Olson 2014, 120)

Even when there is no robustly normative reason to do as, for example, the law or etiquette says, it remains true that there are legal reasons to do what the law requires and etiquette reasons to avoid what etiquette proscribes. It is just that in these situations such reasons lack robust normative authority. Our proposal is that epistemic reasons are just like that. While epistemic norms do imply epistemic reasons, they do so in the same, thinly normative sense that, for example, legal norms imply legal reasons. Just like legal reasons, epistemic reasons are not essentially robust reasons.

This explanation also accommodates the inescapability of epistemic reasons. Thin reasons can very well be inescapable and desire-independent. As Olson writes:

> A soldier might not *desire* to comply with the general's order and he might have no other desires whose satisfaction would be promoted by his compliance, but he can still be said to have reasons to comply with the general's order since complying with the orders of those of superior military rank is part of the role of being a soldier. The same goes for chess players and football players; they might not desire to play by the rules and they need not even desire to win. Agents can occupy roles they have no desire to fulfil and engage in activities they have no desire to succeed in. (Olson 2014, 121)

Similarly, there are legal reasons and etiquette reasons for you to avoid illegal and improper behaviour, whether or not you care about the law or etiquette. In the same way, there are epistemic reasons for you to conform to epistemic norms regardless of what you care about. It is just that, just like legal reasons, epistemic reasons are not essentially robust reasons.

The universality platitude. We have been comparing epistemic norms with norms that are clearly not universal or absolute, but rather conventional and relative. Obviously, there is not one framework or code of law or etiquette that is absolutely "correct" or "valid." Things are not illegal or indecorous *period* or absolutely, but only relative to a specific code of law or etiquette. Moreover, facts about what is legal, stylish, or polite are—or are grounded in—facts about social conventions.

Epistemic norms, in contrast, do not seem to be relative and conventional in this way. Rather, much like moral claims, epistemic claims seem to be true universally and absolutely. For most epistemologists, there is not a plurality of equally valid and conventional epistemic frameworks, but rather one absolute and universal system of epistemic standards. It *does* make sense to claim that a belief is epistemically justified, rational, or correct *period*. This is because—again, according to most philosophers—these epistemic norms are plausibly not just a matter of arbitrary conventions that we came up with. Rather, they are discoverable facts that exist independently of such conventions.[6]

Once again, this feature of epistemic facts is compatible with a thin conception of epistemic normativity. Norms can be universal and nonconventional without having robust normative authority. Thinly normative claims, in other words, do not necessarily imply relative and conventional norms. One way in which this can be the case is if the norms in question are grounded in or derived from some fundamental end or standard that is itself absolute and universal. Importantly, this is precisely how many epistemologists construe the epistemic domain. Many of them argue that epistemic norms are those that derive, somehow, from the fundamental goal or norm of *truth* or *knowledge*. This is in line with what Cuneo calls the *substance* platitudes. It is a platitude, Cuneo points out, that

> Propositional attitudes have epistemic merits such as being justified […] in virtue of their being such as to represent reality aright, being likely to represent reality aright, or being such that an agent has done what she ought to in attempting to represent reality aright. (Cuneo 2016, 74)

It is part of our conception of the epistemic domain, in other words, that it has to do with getting things right and not getting things wrong.

If something like this is correct and if truth is universal and not just a matter of convention, then there might very well be only one true system of epistemic norms. Crucially however, none of this entails that this one epistemic system or framework has robust normative authority. It does not mean that there is necessarily a robust reason to achieve the absolute goal or norm of knowledge or truth. The thin conception of epistemic normativity is therefore compatible with epistemic facts being universal, and not relative or conventional.

4.3 Is Something Missing?

A thin conception of epistemic normativity therefore suffices to capture the set of platitudes, which we have labelled the *epistemic norms* platitudes. But do those platitudes suffice to capture our ordinary conception of epistemic facts and claims? Do we *also* need to capture the authority platitude to really capture the epistemic domain? Does our ordinary conception of epistemic facts and claims include the authority platitude *in addition* to the epistemic norms platitudes? Are we missing something platitudinous if we only say that epistemic facts are highly important, inescapable, universal, deontological, and that they imply epistemic reasons and epistemic evaluations?

We do not think so. Our ordinary conception of epistemic facts and claims does not plausibly include the idea that epistemic facts are not only norm-implying, important, inescapable, universal, and so on, but *also* robustly normative. Perhaps that claim would be plausible if it were also platitudinous that there cannot be situations where we have no robust reason to do as epistemic norms require. That is, if it were obvious and uncontroversial that we necessarily have a robust reason to conform to epistemic norms, then perhaps the epistemic norms platitudes would be insufficient and we would also need to capture the authority platitude.

But this is plainly not the case. It is not difficult to imagine cases where, many would agree, conforming to epistemic norms would be so trivial, reckless, or harmful that there would be simply no robust reason to do so.[7] Granted, proponents of robust epistemic normativity have a ready response to such cases: there is still a *pro tanto* robust reason to conform

to epistemic norms in those cases. It's just that it has almost no weight in comparison to reasons against doing so. But while this response is not obviously mistaken, it is not obviously sound either. After all, someone could make the same claim about many norms, including paradigmatic examples of thinly normative ones. For example, in response to cases where the demands of etiquette are completely trivial or reckless, someone could easily say that there is still a robust, *pro tanto* reason to conform to etiquette, but that it has almost no weight. Yet few would be persuaded by such a defence of the robust normativity of etiquette. In the same way, one could legitimately reject that defence in the epistemic case (see, e.g., Côté-Bouchard 2017b; Hazlett 2013; Whiting 2013; Wrenn 2017). It is therefore not platitudinous that there must always be a *robust* reason to conform to epistemic norms, including in cases where doing so is completely trivial or reckless.

More generally, it is not a platitude that epistemic facts and claims are robustly normative *in addition* to being important, inescapable, universal, deontological, and related to epistemic reasons and epistemic value. Capturing what we called the epistemic norms platitudes is sufficient for capturing our conception of epistemic facts and claims.

5 Why Not Say the Same About Morality?

If we are right that the epistemic domain includes the epistemic norms platitudes, but not the authority platitude, then our strategy avoids the revisionism objection; thin epistemic normativity suffices to capture our conception of epistemic facts and claims.

But could this prove too much? Could this be a pyrrhic victory for error theorists who want to avoid the epistemic error theory, but keep the *moral* error theory? After all, one might claim, what we have said against robust epistemic normativity could also be said against the authority of moral facts. That is, one might claim that the same case could be made against robust *moral* normativity and in favour of *thin* moral normativity.

The worry is that our strategy might face a dilemma. Either our case against robust epistemic normativity is persuasive or it is not. If it is not, then our strategy is not successful, and Cuneo's argument remains intact.

If it is, then it may very well undermine the error theorist's thesis that *moral* thoughts and claims essentially aim at representing robustly normative facts. This would be unacceptable for moral error theorists. If moral claims are only about *thinly* normative facts, then their truth does not require the existence of strange normative facts.[8]

We do not think that our case against robust epistemic normativity applies or extends to moral normativity, however. That is, we do not think that the foregoing undermines the moral error theorist's conceptual claim. This is because there remain key differences between the moral and epistemic that make thin epistemic normativity more plausible than thin moral normativity (and robust moral normativity harder to deny that robust epistemic normativity). Put differently, there are differences that make the authority platitude more plausible in ethics than in epistemology.

First, as we have already seen, many would agree that conforming to epistemic norms is not necessarily robustly good and that violating epistemic norms is not necessarily robustly bad. Once again, it is not hard to imagine situations where knowing the truth, avoiding error, and following the evidence would be so trivial, detrimental, or pointless that there would plausibly be no robust value in doing so. Similarly, we can easily conceive of cases where there is plausibly nothing robustly bad in violating epistemic norms. Granted, some have maintained that epistemic successes like true belief have intrinsic robust value.[9] Nevertheless, many epistemologists and value theorists remain unpersuaded and agree that, in some cases, there is simply no robust value in getting things right. But the same idea is much less plausible in the moral case. It is much harder to deny that moral value necessarily constitutes robust value. It seems more plausible, in other words, that there is necessarily something robustly good in doing as morality requires.

Second, recall the point about animals. Non-human animals can meet or fail to meet certain epistemic standards, but they are not subject to moral standards. Moreover, animals are exempt from robustly normative standards. Therefore, a plausible explanation for the exemption of animals from moral standards is that we see those standards as robustly normative.

Conversely, a plausible explanation for the fact that animals are not exempt from epistemic standards is that they are *not* robustly normative. Finally, one feature that is commonly attributed to morality—and to robust normativity more generally—is an *autonomy* from the non-normative. Many think we cannot derive robustly normative conclusions from non-normative premises alone.

One important way in which the autonomy of the robustly normative manifests itself is in the phenomenon of first-person deliberation. From the first-person point of view, it seems that non-normative considerations alone cannot settle for you the robustly normative questions you ask yourself in deliberation. Enoch explains:

> Because only normative truths can answer the normative questions I ask myself in deliberation, nothing less than a normative truth suffices for deliberation. And because the kind of normative facts that are indispensable for deliberation are just so different from naturalist, not-obviously-normative facts and truths, the chances of a naturalist reduction seem rather grim. […] The gap between the normative and the natural, considered from the point of view of a deliberating agent, seems unbridgeable. (Enoch 2011, 80)[10]

Robustly normative deliberation seems autonomous from the non-normative. Yet the same does not seem true of deliberation that is distinctly epistemic. To deliberate is to ask oneself questions like "what should I do?" or "should I ϕ?" But suppose you ask yourself whether you should believe, for example, that your partner is cheating on you. Even though you can conceivably answer this question by determining whether it would be *desirable* or *beneficial* to have that belief, you can also answer it by settling for yourself the question *whether they really are cheating on you*. Importantly, the former kind of doxastic deliberation is not strictly speaking relevant for epistemology and epistemic facts. Answers to the question whether believing that p would be desirable (regardless of the truth of p) are not epistemic claims. Only the latter—the kind of doxastic deliberation that can be answered by settling the question "is p true?"—results in a distinctly epistemic conclusion.

Crucially however, this latter kind of doxastic deliberation is *not* autonomous from the non-normative. It can be settled by non-normative facts alone from the first-person perspective.[11] Once again, suppose you ask yourself "should I believe that my partner is having an affair?" On the epistemic interpretation, this question is settled by answering the question "is my partner really having an affair?" But non-normative premises can suffice to settle that question for you. Therefore, unlike moral deliberation, non-normative facts alone can suffice to settle epistemic doxastic deliberation from the first-person perspective. All you need to settle the question whether you should *epistemically* believe that *p* is to settle the question whether *p* is true and non-normative facts can plausibly suffice to settle this kind of question for you. Epistemic claims, then, do not seem to have the autonomy from the non-normative that is distinctive of robustly normative claims like moral claims.[12]

For these reasons, we do not think that our defence of thin epistemic normativity extends to moral normativity. There remain differences between the epistemic and the moral that make robust moral normativity much harder to deny than robust epistemic normativity. Our strategy does not, therefore, undermine the moral error theory.

6 Conclusion

We have argued that a certain interesting and influential line of objection to the moral error theory rests on a mistake about the commitments of the epistemic realist. It would appear that epistemic standards are like moral standards in some interesting respects, such as their apparent *a priori* status and their independence from human conventions. They are different to moral standards in an important respect, which is that there do not appear to be robust reasons associated with these standards. In this regard, epistemic standards are closer to the standards of etiquette. If it really matters whether we meet them or not, it's not because of the standards themselves but because of some extrinsic reason to care about meeting them. Once we see that epistemic standards are thin, we can see that the chance that the companions in guilt strategy will succeed is thin, too.

Notes

1. For defences of companion in guilt arguments, see Case (forthcoming); Cuneo (2016, 2007); Nobis (2005); Rowland (2013). For critical discussion, see Cowie (2014); Olson (2011, 2014).
2. The labels "thin" and "robust" are from Cuneo (2016). Other labels used in the literature include *strong* versus *weak categoricity* (Joyce 2001), normativity versus mere *norm-relativity* (Hattiangadi 2007), *reason-implying* versus mere *rule-implying* normativity (Parfit 2011), *robust* versus merely *formal* normativity (McPherson 2011), *normative requirements* versus mere *requirements* (Broome 2013), and *irreducible* versus mere *reducible* normativity (Olson 2014).
3. Maria Alvarez pressed the importance of animal knowers on one of us when one of us was too quick to say that justification was required for knowledge. On many standard accounts of knowledge, a belief constitutes knowledge only if the thinker, *inter alia*, is justified in holding her beliefs. We think that this conception of the connection between justification and knowledge is problematic. One idea, worth exploring further, is to think of knowledge as a natural and non-normative property, one that figures in the application conditions of the norms that determine whether a thinker's beliefs are justified. These norms apply only to thinkers that can be held accountable. That's why, for example, if Agnes and I know that dinner is ready but I'm a human and Agnes is a dog, only my belief can be justified. My belief's being justified is down to two things. The first is that it is knowledge or, if you prefer, down to the fact that it stands in some interesting relation to knowledge. The second is that I can be held accountable for meeting a standard like this. (This is where the dog and I differ.) For a defence of this approach to knowledge and justification, see Littlejohn (2018).
4. See Thomson (2008) for a discussion of a hybrid view of the epistemic "ought" according to which a thinker ought to believe only when the belief is epistemically kosher and furthers something non-epistemic.
5. This platitude is, admittedly, more controversial. Some (e.g. Alston 1988; Levy 2007) have argued that deontology has no place in epistemology because belief is involuntary. We set this issue aside, in what follows, and assume that deontology does have a place in epistemology. It is worth noting, however, that our proposal might help resolve the tension between epistemic deontologism and doxastic involuntarism. If

deontic "oughts" imply voluntariness, this will only be true of robustly normative "oughts." Thinly normative deontological claims, in contrast, do not imply "can." The law forbids you to steal, for instance, even if you are an unwilling kleptomaniac. For more on this point, see Côté-Bouchard (2017c).

6. This is, of course, controversial. Many epistemologists defend forms of epistemic relativism and conventionalism. See for instance, Kusch (2002), Neta (2007), Hazlett (2013), Kerr and Carter (2016). See, for example, Boghossian (2006), Carter (2016), White (2007) for criticism.
7. Of course, there might still be *epistemic* reasons to do so and doing so might still be *epistemically* valuable in these cases. But as we have seen, this is compatible with those epistemic reasons and values lacking robust normativity. As an analogy, in cases where there is no robust reason to conform to the law, there are still *legal* reasons to do so; that act still has legal value. See Côté-Bouchard (2017b) for a more extensive discussion of this point.
8. See Finlay (2008, 2014) for a critique along the lines of the latter branch of the dilemma.
9. See e.g. Horwich (2006), Kvanvig (2008), Lynch (2004), Zagzebski (2003). For criticism, see Côté-Bouchard (2017b), Hazlett (2013), Littlejohn (2018), Maguire and Woods (n.d.), Thomson (2008), Whiting (2013), and Wrenn (2017).
10. See also van Roojen (2015, 258).
11. See Hieronymi (2006, 2005) for a helpful discussion on the relationship between non-normative considerations and the closure of doxastic deliberation.
12. For more on the same theme, see Jenkins (2007), Heathwood (2009), Côté-Bouchard (2017a).

References

Alston, William P. 1988. The Deontological Conception of Epistemic Justification. *Philosophical Perspectives* 2: 257–299.

Boghossian, Paul. 2006. *Fear of Knowledge: Against Relativism and Constructivism*. Oxford: Oxford University Press.

Brink, David O. 1992. A Puzzle about the Rational Authority of Morality. *Philosophical Perspectives* 6: 1–26.

Broome, John. 2013. *Rationality through Reasoning*. Chichester: Wiley-Blackwell.

Carter, J. Adam. 2016. *Metaepistemology and Relativism*. Basingstoke: Palgrave Macmillan.

Case, Spencer. forthcoming. From Epistemic to Moral Realism. *Journal of Moral Philosophy*.

Côté-Bouchard, Charles. 2017a. Epistemological Closed Questions: A Reply to Greco. *Manuscrito* 40 (4): 97–111.

———. 2017b. Is Epistemic Normativity Value-Based? *Dialogue* 56 (3): 407–430.

———. 2017c. 'Ought' Implies 'Can' Against Epistemic Deontologism: Beyond Doxastic Involuntarism. *Synthese* 1–16.

Cowie, Christopher. 2014. Why Companions in Guilt Arguments Won't Work. *Philosophical Quarterly* 64 (256): 407–422.

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

———. 2016. Destabilizing the Error Theory. In *Epistemic Reasons, Norms and Goals*, ed. Pedro Schmechtig and Martin Grajner, 71–94. Berlin: De Gruyter.

Enoch, David. 2011. *Taking Morality Seriously: A Defense of Robust Realism*. Oxford: Oxford University Press.

Finlay, Stephen. 2008. The Error in the Error Theory. *Australasian Journal of Philosophy* 86 (3): 347–369.

———. 2014. *Confusion of Tongues: A Theory of Normative Language*. Oxford: Oxford University Press.

Foot, Philippa. 1978. *Virtues and Vices*. Oxford: Oxford University Press.

Garner, Richard T. 1990. On the Genuine Queerness of Moral Properties and Facts. *Australasian Journal of Philosophy* 68 (2): 137–146.

Harman, Gilbert. 1986. *A Change in View*. Cambridge, MA: MIT Press.

Hattiangadi, Anandi. 2007. *Oughts and Thoughts: Rule-Following and the Normativity of Content*. Oxford: Oxford University Press.

Hazlett, Allan. 2013. *A Luxury of the Understanding: On the Value of True Belief*. Oxford: Oxford University Press.

Heathwood, Chris. 2009. Moral and Epistemic Open-Question Arguments. *Philosophical Books* 50 (2): 83–98.

Hieronymi, Pamela. 2005. The Wrong Kind of Reason. *Journal of Philosophy* 102 (9): 437–457.

———. 2006. Controlling Attitudes. *Pacific Philosophical Quarterly* 87 (1): 45–74.

Horwich, Paul. 2006. The Value of Truth. *Noûs* 40 (2): 347–360.
Jenkins, C.S. 2007. Epistemic Norms and Natural Facts. *American Philosophical Quarterly* 44 (3): 259–272.
Joyce, Richard. 2001. *The Myth of Morality*. Cambridge: Cambridge University Press.
———. 2006. *The Evolution of Morality*. Cambridge, MA: MIT Press.
Kerr, Eric T., and J. Adam Carter. 2016. Richard Rorty and Epistemic Normativity. *Social Epistemology* 30 (1): 3–24.
Kusch, Martin. 2002. *Knowledge by Agreement: The Programme of Communitarian Epistemology*. Oxford: Oxford University Press.
Kvanvig, Jonathan. 2008. Pointless Truth. *Midwest Studies in Philosophy* 32 (1): 199–212.
Levy, Neil. 2007. Doxastic Responsibility. *Synthese* 155 (1): 127–155.
Littlejohn, Clayton. 2018. The Right in the Good: A Defense of Teleological Non-Consequentialism in Epistemology. In *Epistemic Consequentialism*, ed. Kristoffer Ahlstrom-Vij and Jeffrey Dunn. Oxford: Oxford University Press 23–47.
Lynch, Michael. 2004. *True to Life: Why Truth Matters*. Cambridge, MA: MIT Press.
Mackie, J.L. 1977. *Ethics: Inventing Right and Wrong*. London: Penguin Books.
Maguire, Barry, and Jack Woods. n.d. Explaining Epistemic Normativity.
McPherson, Tristram. 2011. Against Quietist Normative Realism. *Philosophical Studies* 154 (2): 223–240.
Neta, Ram. 2007. In Defense of Epistemic Relativism. *Episteme* 4 (1): 30–48.
Nobis, Nathan. 2005. Truth in Ethics and Epistemology: A Defense of Normative Realism. Thesis, University of Rochester.
Olson, Jonas. 2011. In Defense of Moral Error Theory. In *New Waves in Metaethics*, ed. Michael Brady, 62–84. Basingstoke: Palgrave Macmillan.
———. 2014. *Moral Error Theory: History, Critique, Defence*. Oxford: Oxford University Press.
Parfit, Derek. 2011. *On What Matters*. Vol. 2. Oxford: Oxford University Press.
Raz, Jospeh. 2014. *From Normativity to Responsibility*. Oxford: Oxford University Press.
Roojen, Mark van. 2015. *Metaethics: A Contemporary Introduction*. Abingdon: Routledge.
Rowland, Richard. 2013. Moral Error Theory and the Argument from Epistemic Reasons. *Journal of Ethics and Social Philosophy* 7 (1): 1–24.

Streumer, Bart. 2017. *Unbelievable Errors: An Error Theory About All Normative Judgments*. Oxford: Oxford University Press.
Thomson, Judith Jarvis. 2008. *Normativity*. Chicago: Open Court.
White, Roger. 2007. Epistemic Subjectivism. *Episteme* 4 (1): 115–129.
Whiting, Daniel. 2013. The Good and the True (or the Bad and the False). *Philosophy* 88 (2): 219–242.
Wrenn, Chase B. 2017. Truth is Not (Very) Intrinsically Valuable. *Pacific Philosophical Quarterly* 98 (1): 108–128.
Zagzebski, Linda. 2003. The Search for the Source of Epistemic Good. *Metaphilosophy* 34 (1–2): 12–28.



Constitutivism About Epistemic Normativity

Christopher Cowie and Alexander Greenberg

Suppose that you possess strong, perhaps decisive, evidence for a proposition that you're considering. It is natural to think that that you thereby possess a *reason* to believe that proposition. What explains this? Why is there such a connection between evidence for (or against) propositions and reasons for (dis)believing them? What, in other words, are the *grounds* of epistemic normativity?

There are a number of different candidate answers to this question. One candidate is that epistemic normativity doesn't have any grounds. According to this view, there are just some basic—that is, ungrounded—truths about what we ought to believe. At the other end of the spectrum is the view that epistemic normativity is grounded in practical normativity.

C. Cowie (✉)
Department of Philosophy, University of Durham, Durham, UK
e-mail: christopher.d.cowie@durham.ac.uk

A. Greenberg
Lady Margaret Hall, University of Oxford, Oxford, UK

Birkbeck, University of London, London, UK
e-mail: alexander.greenberg@philosophy.ox.ac.uk

© The Author(s) 2018
C. Kyriacou, R. McKenna (eds.), *Metaepistemology*, Palgrave Innovations in Philosophy, https://doi.org/10.1007/978-3-319-93369-6_8

According to this view, epistemic reasons for belief are explained by the contribution that believing in light of the evidence makes to prudential, moral or aesthetic ends. Neither of these views is ideal. The former isn't much of an *explanation* at all. It makes epistemic normativity a fundamental constituent of reality.[1] The latter struggles with the fact that believing evidentially supported propositions isn't always of practical value, and the fact that it sometimes seems to be of practical value to have evidentially unsupported beliefs. Of course, defenders of these views have developed sophisticated responses to these basic charges. But the fundamental problems remain. It would be ideal if there were some happy middle-ground; if epistemic normativity were neither ungrounded, nor grounded in the practical. That could allow us to avoid the problems that each approach faces. Interestingly, there is a candidate for occupying this middle ground. The candidate is that epistemic normativity is grounded in the nature of the attitude of *belief*. This view is now often referred to as *constitutivism*, or, more fully, *constitutivism about epistemic normativity*. The idea is that *the nature of belief* explains why have reasons to believe, or disbelieve, propositions in light of evidence possessed for, or against, them.

How is this meant to work? The idea is that *the concept of belief is normative*. This means that for an attitude to fall under the concept "belief" is just for it to stand in normative relations to a certain property (or properties). This property is said to be the *correctness condition* for belief. Belief's correctness condition is typically claimed to be truth; specifically, the truth of the proposition believed. We can therefore express constitutivism about epistemic normativity as follows:

> **TRUTH NORM**: It is a conceptual truth about belief that it is correct to believe a proposition if and only if it is true.[2]

Constitutivism offers an explanation *in terms of the nature of belief* for why there are normative relations between belief and truth. And once this in place, it isn't a big leap—the thought is—to explain why there are normative relations between belief and *evidence* (evidence being an indicator of truth). We now have an account of the grounds of epistemic normativity.

Does this constitutivist explanation of epistemic normativity work? Our focus will not be on whether TRUTH NORM, if accepted, can ground epistemic normativity. The aim of this article is to assess the arguments that have been given for TRUTH NORM. We will ask whether these arguments are persuasive, and we will claim that they are not. We will claim that there is insufficient reason to accept TRUTH NORM. *A fortiori* there is insufficient reason to think that TRUTH NORM is the correct explanation of the grounds of epistemic normativity.

There are a range of different arguments for TRUTH NORM (for a survey, see McHugh and Whiting 2014). As such, one might expect that we would have to proceed in a piecemeal fashion, taking each argument for TRUTH NORM in turn. Our ambition, however, is to take a more unified approach. We aim to present a common strategy. We aim to show that some key arguments for TRUTH NORM fail for a common reason. They fail because TRUTH NORM fails to provide a good explanation of the phenomena that it is meant to explain. More precisely, we will claim arguments for TRUTH NORM falter on a dilemma. The dilemma is as follows:

> **DILEMMA**: TRUTH NORM is appealed to as the best explanation of some phenomenon (the explanandum).
>
> (i) *Either* TRUTH NORM doesn't look like it best explains the explanandum in question.
> (ii) *Or* TRUTH NORM can explain the explanandum, but *only if* it is supplemented by an implausible claim about *following* prescriptions.

Either way, TRUTH NORM fails to provide the best explanation for the phenomenon in question. Given that the ability to provide such an explanation is the reason for positing TRUTH NORM, we shouldn't posit TRUTH NORM.

This is our DILEMMA. We will apply it in two specific arguments for TRUTH NORM: first to an argument based on *deliberative transparency*; second to an argument based on *Moore-paradoxical sentences*. Part of what is interesting about our strategy is that we are drawing together these

seemingly distinct arguments for and against TRUTH NORM in a simple, unifying way. We will begin (Sect. 1) by making some preliminary remarks about how our approach differs from and, we think, improves upon existing opposition to TRUTH NORM. We will then (Sect. 2) show how our strategy applies against arguments for TRUTH NORM based on deliberative transparency. We will then (Sect. 3) show how this strategy applies against arguments for TRUTH NORM based on Moore-paradoxical sentences. We then conclude (Sect. 3.3).

1 What Distinguishes Our Argument from Other Arguments Against TRUTH NORM

The argument that we will develop against TRUTH NORM differs from some of the existing arguments in the literature. In particular there are two important kinds of criticism that have been made of TRUTH NORM that our argument should be distinguished from.

One influential criticism of TRUTH NORM concerns the precise content of the norm that is alleged to be constitutive of belief. There has been considerable debate about whether a truth norm can be formulated in such a way as to avoid implausible prescriptions—such as that one ought to believe ungraspably complex truths and a prescription believe "true blindspots" (Bykvist and Hattiangadi 2007, 2013). We will set this aside however, and grant that the defender of TRUTH NORM has some response to this criticism.

A second important criticism of TRUTH NORM is the "correctness" of beliefs it refers to is not "heavy duty" enough to do any epistemological or metaphysical work (Côté-Bouchard 2016; cf. Bykvist and Hattiangadi 2013, sec. 2; Rosen 2001, 619–21). How our argument differs from this is somewhat complex. While our conclusion is the same—that TRUTH NORM does not best explain the nature of belief—our argument targets a later stage in the dialectic. The arguments we will be rejecting—which claim that TRUTH NORM is required to explain deliberative transparency and Moore's paradox—can be understood as attempts to argue, *contra* the above criticism, that belief's correctness conditions have to be

understood as *genuinely normative*. Our conclusion is that this is not true: these phenomena do not mandate a genuinely normative understanding of belief's correctness conditions.

2 Does TRUTH NORM Explain Deliberative Transparency?

One influential argument for TRUTH NORM is that it best explains the transparency of doxatic deliberation. We will begin by outlining this argument, largely following the presentations of it given by Nishi Shah and David Velleman (Shah 2003, 260–74; Shah and Velleman 2005, 500–02), before criticising it through our DILEMMA (in Sects. 2.1 and 2.2).

When we deliberate as to whether to believe that *p*, we recognise that this question inevitably gives way to the question of whether *p* is true. Furthermore, not only is the deliberative question of whether to believe that *p* answered by the factual question of whether *p* is true, it is *only* answered by that question, to the exclusion of any other considerations, such as whether believing that *p* would be in one's interests. This is what it means to say that deliberation about what to believe is *transparent*. What explains this phenomenon? Arguably the obvious place to look for an explanation is to the existence of a tight relationship between belief and truth, because the phenomenon just is one of a question about *belief* giving way to a question about truth. TRUTH NORM prescribes that we believe *p* if and only if *p* is true. Consequently, when we deliberate about whether to believe *p*, we will—if we follow that norm—simply deliberate about whether *p* is true. The transparency of deliberation is explained. So TRUTH NORM can explain the transparency of deliberation.

This shows that TRUTH NORM is a sensible way of explaining transparency. But why think it is the *best* explanation? The main reason given by its defenders is that it provides a better explanation than alternative *dispositional* accounts of belief. To explain the *exclusive* focus on truth in deliberation, the dispositions specific to belief (that dispositional accounts posit) would need to dispositions to form, maintain, and revise beliefs in

and, crucially, *only in* response to truth-relevant considerations. Call this the *strong dispositional account*.

The fact that a dispositional account of belief would have to be "strong" in this respect in order to explain transparency is the reason why TRUTH NORM is claimed to provide a better explanation. Take a step back. Note that although when we *deliberate* about whether to believe *p* we simply consider whether *p* is true, not *all* of our beliefs are like this. Sometimes, what we believe is determined by our desires and other conative states, in cases of wishful thinking. The key datum, according to defenders of TRUTH NORM, is that wishful thinking doesn't exert its influence *within conscious deliberation*.

This presents a problem for the strong dispositional account. On this account, the products of wishful thinking would not count as beliefs. Now one could respond to this by *weakening* the dispositional connection between belief and truth so as to allow for the products of wishful thinking to count as beliefs, and defend what we can call the *weak dispositional account*. But a defender of this view can no longer explain why, when we consciously deliberate about whether to believe that *p*, we simply deliberate as to whether *p* is true. On her account, it should, in principle, be possible to deliberate about whether *p*, and be moved in those deliberations in part by non-truth related (e.g. conative) considerations. But, *ex hypothesi*, it's not. So this weak dispositional account fails too.

In fact, we can represent the failings of both of the dispositional accounts as follows (Table 1):

Neither of these accounts works because neither can answer "yes" to both questions. This is where the real strength of TRUTH NORM

Table 1 Dispositional accounts

	Explains transparency in deliberation?	Allows for the possibility of (non-deliberative) wishful beliefs?
Strong dispositional account	Yes	No
Weak dispositional account	No	Yes

Table 2 TRUTH NORM

	Explains transparency in deliberation?	Allows for the possibility of (non-deliberative) wishful beliefs?
TRUTH NORM	Yes	Yes

comes in. TRUTH NORM seems able to answer "yes" to both (Table 2):

Why is this? The key point is that TRUTH NORM is a view about the *concept* of belief. This is key because and when we deliberate about what to believe, the concept of belief is playing a role in framing what we do, whereas it isn't in ordinary non-deliberative belief-formation. It is this that allows TRUTH NORM to explain transparency in deliberation, whilst still allowing for the possibility of wishful thought. And so, it is this that makes TRUTH NORM the *best* explanation of transparency in deliberation.

2.1 Transparency in Deliberation: Our DILEMMA

We don't think that this is a good argument for TRUTH NORM. We can show why by appealing to the following DILEMMA:

> **DILEMMA**: TRUTH NORM is appealed to as the best explanation of the transparency of deliberation. But (we argue):
>
> (i) *Either* TRUTH NORM doesn't look like it best explains the explanandum in question.
> (ii) *Or* TRUTH NORM can explain the explanandum, but *only if* it is supplemented by an implausible claim about our motivations to follow prescriptions.

We will make the case quickly before expanding on it in what follows. The suggestion that this argument for TRUTH NORM has a problem with *motivation* was first suggested by Asbjørn Steglich-Petersen (2006), and was further developed by Conor McHugh (2013). But we extend this point, and, crucially, in the following section we attempt to show that is generalises to attempts to use TRUTH NORM to respond to Moore's paradox.

Let's begin with horn (ii). According to defenders of TRUTH NORM, deliberative transparency works as follows. The concept of belief pre-

scribes believing *p* iff *p* is true. Deliberation about what to believe is framed by the concept of belief, according to defenders of TRUTH NORM, because it aims to an answer the question *whether to believe that p*. So, given that it's part of the concept of belief that a belief is correct iff it is true, a deliberator about what to believe will necessarily take themselves to be under this prescription. So they will try to accept the proposition about which they are deliberating iff it is true. This explains why their deliberation is transparent; it explains why the only consideration they regard as relevant is *p*'s truth, and not, for example, whether believing *p* is in their interest.

One of the obvious problems with this story is that it seems to require a very strong connection between (a) taking oneself to be under a prescription to behave in some way, and (b) actually behaving in that way. It seems to require that because deliberators will take themselves to be under a prescription to believe that *p* iff *p* is true, they will actually try to do that. We can label it as follows:

(1) If you take it to be correct for you to V, then you will be motivated to V.

This is a form of "judgement internalism." But it isn't a plausible form. Frequently, we take ourselves to be under a prescription to behave in some way, yet we don't try to behave in that way. So (1) looks like a bad principle. It's far too strong to be true. This is borne out by considering the forms of judgement internalism actually defended by moral philosophers. Those who claim that there is some necessary connection between judging an action to be morally right (or good, correct, or what one ought to do) and being motivated to perform that action typically claim that this connection is, in various ways *defeasible*. Some claim that the moral judgement necessarily results in *some* motivation to act, but not necessarily overriding motivation; others claim that moral judgement necessarily results in motivation to act in *ideal* circumstances, or if one is *rational* (for a survey, see Svavarsdottir 1999, sec. 1). A representative example of this kind of judgement internalism is put forward by Michael Smith:

> [A]gents who judge it right to act in various ways are so motivated, and necessarily so, absent the distorting influences of weakness of the will and other similar forms of practical unreason on their motivation. (Smith 1994, 61)

These qualified versions of judgement internalism are typically defended, as the passage from Smith indicates, because a version of internalism without such qualifications would have the implausible consequence that it would be impossible to judge that something is right but fail to do so out of weakness of the will, deep depression, et cetera. But these kinds of cases of irrationality also show that a principle like (1)—which expresses a simple entailment between judging it to be correct to V and being motivated to V—is implausible.

Can the defender of TRUTH NORM accept that form of internalism expressed in (1) is implausibly strong? This leads us onto horn (i). Perhaps, she might argue, it is possible that we do sometimes take ourselves to be under the prescription of TRUTH NORM, yet fail to follow it. The result is that in some cases we will be able to deliberate as to whether to believe that *p*, yet not be sensitive only to truth-directed related considerations. But, the defender of TRUTH NORM might claim this is ok because deliberation isn't always transparent. This would be true if there are certain cases of self-conscious *epistemic akrasia*, in which someone judges that the evidence suggests that *p* is true, but doesn't believe that *p*. In fact, Pascal Engel defends the claim that the truth norm need not necessarily motivate by appealing to cases of this kind:

> Just as cases of *akrasia* or *accedia* can arise where the agent considers the norm but does not follow it, cases where the norm of truth is considered by the agent but is not followed can arise. (Engel 2007, 199)

If this is correct, then the defence of TRUTH NORM need no longer rely on a very strong connection between the prescriptions that we take ourselves to be under, and our tendency to follow them. This is because the explanandum has now changed; what is now to be explained is a general tendency for deliberation to be transparent. But the problem is that TRUTH NORM now loses its unique explanatory power. A weak dispositional account—one which claims that belief is essentially responsive to be truth-relevant considerations, but not exclusively so—can explain a mere general tendency for deliberation to be transparent. We don't need TRUTH NORM any more.[3]

That's the general structure of the DILEMMA. In horn (ii) a defender of TRUTH NORM is shown to rely on a *very* strong connection between

putting oneself under a prescription to behave in some way, and behaving in that way. And in trying to escape horn (ii), the explanandum (transparency in deliberation) has to be weakened to avoid this commitment. This then lands the defender of truth norm on horn (i) because it results in TRUTH NORM losing its unique explanatory appeal.

What are the best lines of response to our DILEMMA? We will argue that all the obvious moves that might be made by the defender of TRUTH NORM still fall on one of the two horns. Either they involve implausibly strong claims about motivation, or they have the result that TRUTH NORM does not best explain transparency.

2.2 Can the Defender of TRUTH NORM Claim Epistemic *Akrasia* Is Impossible?

One move the defender of TRUTH NORM could make is to claim that the strong version of internalism expressed by (1) is not, as we claim, implausible if it is just restricted to belief. One might argue, according to this line of thought, that (1) is implausible in relation to *action*, because practical *akrasia*—self-consciously acting contrary to what one judges to be correct to do—is a real and familiar phenomenon. However, one could claim that (1) is plausible *if* restricted to belief. One might try and motivate this claim by arguing that epistemic *akrasia*—self-consciously believing contrary to what one judges to be correct—is not in fact possible. This kind of claim is made by Philip Pettit and Michael Smith:

> Imagine that your beliefs run counter to what evidence and fact require. In such a case, your beliefs will not allow those requirements to remain visible because the offending beliefs themselves give you your sense of what is and your sense of what appears to be. You are therefore denied an experience whose content is that you are believing such and such in defiance of the requirements of fact and evidence. (Pettit and Smith 1996, 448)

Pettit and Smith claim that epistemic *akrasia* is impossible because, roughly, one's judgement about the correctness of one's belief that *p* just is one's belief about whether *p*. This provides a response to our DILEMMA because this alleged impossibility of epistemic *akrasia* could be claimed to

support the strong form of internalism expressed in (1) (if restricted to belief). This is because it would show that judging that it's correct to believe that *p* does in fact entail being motivated to believe that *p*.[4]

We don't think this escapes our DILEMMA. Our response is that the claim that epistemic *akrasia* is impossible is an equally contentious and implausibly strong claim about motivation as (1). As such, it still falls foul of horn (i).

Why is Pettit and Smith's claim implausible? While it correct that one's judging that one's belief that *p* is correct or incorrect *typically* results in one forming or revising a belief that *p*, it's quite another thing to claim these are *just one and the same*. This second claim is what denying the possibility of epistemic *akrasia* has to involve; but this second claim is implausibly strong. It is implausibly strong because it rules out certain cases—like Hume's judgement that even though he lacks reason for his beliefs in the external world, he can't help having them—which may be *irrational* but certainly aren't impossible. In this way, the option of denying the possibility of epistemic *akrasia* lands the defender of TRUTH NORM on horn (i) of our DILEMMA, because it relies on a claim about motivation that is just as implausibly strong as the internalism expressed in (1).

2.3 Can the Defender of TRUTH NORM Appeal to a More Plausible Form of Internalism?

The other obvious move for the defender of TRUTH NORM is to claim that their explanation of transparency only needs to appeal to a weaker—and therefore more plausible—version of internalism. But we will show that weaker forms of internalism are still vulnerable to our dilemma; what they gain in plausibility, they lose in their ability to explain transparency.

The clearest way to amend the implausibly strong form of internalism expressed in (1) is as follows:

(2) If you're rational, then: if you take it to be correct for you to V, then you will be motivated to V.

This is a much weaker—and therefore more defensible—account of the relation between taking oneself to be under a prescription and following it.

However, if they are going to accept (2), then defenders of TRUTH NORM will also need to change how they characterise what TRUTH NORM is supposed to explain. Because (2) is restricted to *rational* motivation, it will only be able to explain the transparency of *rational* doxastic deliberation. In other words, the explanandum now becomes the fact that *if you're rational*, then if you deliberate about whether to believe that *p*, your deliberation will be transparent. This could allow defenders of TRUTH NORM to explain deliberative transparency with the weaker version of judgement internalism expressed in (2).

Can moving to (2) save TRUTH NORM from the above DILEMMA? We don't think it can. This is where horn (i) of our DILEMMA bites. The more defensible version of internalism expressed in (2) means that explanandum needs to be re-characterised as the transparency of *rational* doxastic deliberation. But re-characterising the explanandum in this manner means that TRUTH NORM loses its unique explanatory power, because other accounts can equally well explain the transparency of *rational* doxatic deliberation.

Specifically, recall the strong dispositional account, which claimed that belief is essentially regulated by dispositions to be responsive to be formed, maintained, and revised in *and only in* response to truth-relevant considerations. Because the explanandum has now been changed to the fact that *rational* doxastic deliberation is transparent, in order to explain that these dispositions need not be claimed to be essential to belief as such; they only need to be claimed to be essential to *rational* belief. And if such dispositions are essential to rational belief, that can perfectly well explain why rational deliberation about what to believe is motivated by and only by truth-relevant considerations. Furthermore, whereas it was thought to be problematic to think of these strong truth-responsive dispositions as essential to belief *as such*—because it had the result that the products of wishful thinking don't count as belief—it is *not* problematic to think of these strong truth-responsive dispositions as essential to *rational* belief, because the products of wishful thinking aren't *rational* beliefs. In this way, if the explanandum is weakened to the transparency of *rational* doxastic

deliberation, an alternative dispositional account can explain the weakened explanandum just as well as TRUTH NORM. Weakening the explanandum in this fashion has the result, therefore, that it is not best explained by TRUTH NORM.[5]

A second weaker form of internalism is in fact offered by Shah and Velleman. They themselves are wary of the strong view of internalism that their view seems to require. So they offer a weaker alternative in a footnote (2005, 533–34 fn. 40). But the footnoted version of internalism is also vulnerable to our DILEMMA. On the first horn, their weaker alternative internalism is in fact too weak to explain transparency. On the second horn, they could get around this by weakening the conception of transparency that they are trying to explain. But if they do this then TRUTH NORM isn't a better explanation than competitors.

Let us explain why. Shah and Velleman claim that their argument in fact only relies on a weaker form of internalism than (1), which they outline as follows:

> The relevant form of internalism does not require a positive disposition to obey any norm that one applies; what it requires is the lack of a disposition to obey a different norm instead. One cannot genuinely apply the norm of truth to an attitude while simultaneously trying *only* to make it conform to some other, unrelated norm. This form of internalism does not rule out obedience to additional norms compatible with the one applied. One can aim to arrive as quickly as possible at a true cognition with respect to *p*—in which case, one will deliberate in accordance with a norm of speed as well as the norm of truth. (Shah and Velleman 2005, 533–34, italics ours)

We can formulate Shah and Velleman's weaker internalism as follows:

(3) If you take it to be correct for you to V, then you will not be motivated by norms prescribing refraining from V-ing.

(3) is still vulnerable to our dilemma. This is because, according to (3), someone's accepting TRUTH NORM does *not* entail that they are motivated to believe the truth and—crucially—*only* the truth. All that acceptance of TRUTH NORM means, according this version of internalism, is that one cannot be motivated *only* by non-truth-related considerations.

But this is consistent with the possibility that one *can*—even if one accepts TRUTH NORM—be motivated *in part* by non-truth-related considerations (cf. McHugh 2013, 457).

But if this is the case, this seems to be too weak to explain transparency. Why? Because transparency seems to be a phenomenon in which you are motivated *only* by truth-relevant considerations.

Now Shah and Velleman could get around this by denying this conception of transparency. They could say that in ordinary or typical doxastic deliberation, we are only motivated in part by truth-relevant considerations. But if they say this, then TRUTH NORM isn't a better explanation than competitors, in particular, the weak dispositionalist account, and they succumb to horn (i). So Shah and Velleman's own weakened internalism is also vulnerable to our DILEMMA.

Both of these weaker forms of internalism, therefore, suffer from our dilemma. They *either* mean that TRUTH NORM cannot explain transparency, *or* if transparency is re-characterised so that TRUTH NORM can explain it, it no longer *best* explains it. There are, or course, alternative weaker versions of internalism we haven't discussed. One might claim that judging it is correct for one to V entails being motivated to V *somewhat* or *in ideal circumstance*. But there is every reason to think that other weaker versions will be just as vulnerable to horn (i) of our DILEMMA, for predictable reasons. That is, they will require us to weaken the explanandum in such a way that TRUTH NORM no longer uniquely explains it.

3 Does TRUTH NORM Explain the Absurdity of Accepting Moore-Paradoxical Sentences?

In the preceding sections we have explained how our DILEMMA works against the claim that TRUTH NORM is the best explanation of deliberative transparency. We will now extend our basic argumentative strategy in a novel direction. We will argue that our DILEMMA can be used to respond to attempts to use Moore-paradoxical sentences to argue for

TRUTH NORM. The basic structure of our argument mirrors those used above. Attempts to use TRUTH NORM to explain Moore-paradoxical sentences are vulnerable to our DILEMMA. Either they are inadequate explanations of the explanandum, or they must be supplemented by implausibly strong claims about our motivations to follow prescriptions.

Moore-paradoxical sentences are sentences of the following form.

MP: *p*, but I don't believe that *p*.

These sentences can sometimes be true. Suppose that Donald Trump were to say: "man-made climate change is happening, but I don't believe that man-made climate change is happening." This would be true. It would be true because both of its conjuncts are true. The first conjunct is true: man-made climate change is happening. And the second conjunct is true: Trump doesn't believe that it is happening (at least according to his Twitter account).

Although instances of MP can be true, it is often thought that there would be something absurd, incoherent or even impossible with one's accepting—whether believing or sincerely asserting—an instance of it from one's own, first-personal perspective. In sincerely saying, or thinking [man-made climate change is happening, but I don't believe that man-made climate change is happening] Trump would be doing something absurd, incoherent, or perhaps even impossible.

In the philosophical literature on these sentences, the absurdity—in some form or other—is taken for granted. The philosophical trick is to explain its nature and source. The explanation, we know, can't appeal to the inconsistency of the two conjuncts. We know that they're not necessarily inconsistent. So how can we explain it?

One popular kind of explanation appeals to features of *speech*. It is absurd to say, or assert, instances of MP because the first conjunct of MP conversationally implies the falsity of the second or is inconsistent with some conversational principle or maxim, or some principle or maxim governing the nature of assertion itself. Whilst there is a lot to be said for this kind of approach, it faces a problem. The problem is that, *prima facie*, the absurdity of accepting a Moore-paradoxical sentence isn't

restricted to *saying* or asserting it. It's also absurd to genuinely *think* it. Any explanation of the absurdity of accepting Moore-paradoxical sentences needs to respect this. The point is well-put by Michael Huemer:

> Moore's Paradox cannot be fully resolved by appeal to rules governing solely the use of language, because it is easy to construct non-linguistic versions of the paradox. It would be absurd to think to oneself… the thought, that it is raining but that one does not believe this. (Huemer 2007, 144)

Bearing this in mind, it isn't surprising that explanations of the absurdity of accepting instances of MP should draw on the nature of belief. With that in mind, it is interesting to think about TRUTH NORM. Can TRUTH NORM help to explain the absurdity of accepting an instance of MP? You might think (or at least suspect) so. Certainly, there are lots of references in the literature to the relation between the normativity of belief and the absurdity of Moore-paradoxical sentences. For example, both Bernard Williams and Peter Railton claim that Moorean absurdity is a product of the fact that belief "aims at truth" (Williams 1973, 137; Railton 1994, 72–73), and those who defend TRUTH NORM often claim it is a way of unpacking the idea that belief aims at truth (see, e.g., Wedgwood 2002, 267; Engel 2013, 32). It's not hard to get a rough sense that this might be right.

The closest to a developed argument for TRUTH NORM on the grounds that it explains Moore's paradox is given by Thomas Baldwin (2007). Baldwin, however, phrases his explanation in terms of believing that p essentially involving a "normative commitment to truth" rather than in terms of a truth norm. On Baldwin's account, Moorean absurdity is "the absurdity of embracing, within a single assertion or judgment, a commitment both to the truth of p and to one's being committed to the falsehood of p; for even though this act does not entail that one is committed to the falsehood of p, it is absurd to commit oneself to having such a commitment just when one also commits oneself to the truth of p" (Baldwin 2007, 88). For simplicity, and because this idea of a normative commitment to truth is not entirely clear (is it more than just a belief that p is true?), we will consider a parallel explanation which attempts to explain Moorean absurdity explicitly in terms of TRUTH NORM.

3.1 Does TRUTH NORM Help?

TRUTH NORM, recall, states:

> **TRUTH NORM**: It is a conceptual truth about belief that it is correct to believe a proposition if and only if it is true.

Suppose that this is correct. Can it be used to explain what's absurd about accepting MP? Our argument will be that it doesn't—it is vulnerable to our DILEMMA.

Let's begin however by thinking about how TRUTH NORM *might be thought to* explain the absurdity of accepting MP. If we assume TRUTH NORM, then one who accepts MP will be committed by its first conjunct to: [It is correct to believe that *p*]. So, if we assume TRUTH NORM, then one who accepts MP is committed to:

> MP*: It is correct to believe that *p*, but I don't believe that *p*.

So, if TRUTH NORM is true, then the Moore-paradoxical sentence, MP, commits one who accepts it to claiming—simultaneously—that it would be correct to do (or be doing) something, and that one isn't doing that thing. Perhaps *that's* what's absurd about accepting MP: it's absurd to simultaneously accept that it is correct to do (or be doing) something, and that one isn't doing that thing.

Does this really help to explain what's strange about accepting MP? An obvious response is: no, it's not really strange—certainly not incoherent, absurd or impossible—to judge that one is under a prescription to respond in some way, but that one isn't responding in that way. Think about ordinary cases in which the prescription takes the form of an "ought" fact rather than a correctness fact. For example, suppose I judge: [I ought to give money to charity, but I don't]. In making this judgement, I don't seem to be saying anything strange or absurd and certainly not something *impossible*. It seems like a rather normal thing to judge. In fact, I judge it frequently.[6] Contrast it with: [Climate change is happening, but I don't believe it is happening]. This *really does* seem strange. There is a big difference, in that respect, between these two judgements. So, one

might think, it is mistaken to attempt to explain the absurdity of accepting Moore-paradoxical sentences by using TRUTH NORM to change MP into MP*.

Perhaps, however, this is too simple. Perhaps a defender of TRUTH NORM would do better by supplementing the basic structure of explanation (i.e. that MP commits one to MP*) in some way. How could we do that? One might try to claim that there is something *irrational* about judging MP*, and hence MP. This strategy seems to have something to recommend it. It's easy to suspect that there is something irrational in accepting MP*. And if in fact it *is* irrational, then we could explain the absurdity of accepting MP*, and hence MP, in these terms. It would be *impossible for a rational person* to judge MP* and hence MP. TRUTH NORM would have allowed us to explain the absurdity of accepting MP* and hence MP.

There are, however, major problems with this strategy, two of which we will detail. The first problem is that this is the wrong kind of explanation of the absurdity of judging MP. The second, more serious problem, is that it doesn't work as an explanation of the absurdity of judging MP anyway because it fails on our DILEMMA.

3.2 First Problem: Wrong Kind of Explanation

The first problem is really a preliminary. Suppose that it *is rationally impossible* to accept MP. The problem is that it isn't obvious that this is a satisfactory explanation of the absurdity of accepting MP. It doesn't yield the *right kind* of absurdity.

Why not? Simply put: If the problem with believing MP is that it is rationally impossible, then we should expect someone irrational to be able to do it. But I'm not sure that we do expect this. It is not at all obvious that it is possible for an irrational person to believe [*p*, but I don't believe that *p*]. Believing [*p*, but I don't believe that *p*] is, we might think, more like *psychologically* impossible than rationally impossible. Even if you're irrational, you simply won't be able to hold this belief (though of course you may believe that you hold it).

It is useful here to refer back to the discussion of deliberative transparency in the earlier part of the paper. There we noted that it is unsatisfactory

to explain deliberative transparency in such a way that it is *merely rationally impossible* to violate transparency in deliberation. It is not clear that this is the right kind of impossibility. It isn't obviously possible for an irrational person to violate deliberative transparency, nor would it obviously be irrational should someone for whom it is possible violate it. Much the same, I think, is true with Moore-paradoxical sentences.

3.3 Second Problem: Applying Our Dilemma

The second problem is more serious. According to the view we are considering it is irrational to judge MP*. The problem, as noted above, is that it isn't obvious that it *is* irrational. Certainly, we wouldn't want to just *assert* that it is. We would need some reason or argument for thinking that it is. Ideally, we would want a general principle that explains why it is irrational. What might such a principle look like?

The obvious place to look in this area is to some form of *judgement internalism* or *enkratic principle* more broadly. According to a simplified, classic form of judgement internalism (this was already referenced in the discussion of transparency above but we've specified the *scope* here for reasons that will be clear shortly):

(4) If you're rational, then: if you judge it is correct to V, then you will be motivated to V.

This kind of claim is not uncontroversial, but it is often defended in some form or other. Let's suppose it is true. Let's suppose, furthermore, that there are no problems with V-ing standing for believing (as well as acting, as it does in standard formulations). Could it be used to explain why it is irrational to judge MP*, and hence MP? We don't think it could. What we need to explain is why it is irrational to judge: it is correct to believe that p, but I don't believe that p. But (4) won't help us to explain the irrationality of this judgement. (4) would only help us to explain why it is irrational not to believe p, having judged that it is correct to believe p.

Consider an example from the domain of action that can help to make this contrast intuitive. According to classical judgement internalists, it is irrational for one to judge [I ought to give money to charity], and then

not give money to charity. It is a form of practical irrationality. But these same judgement internalists don't claim that it is irrational for one to judge [I ought to give money to charity, but I don't give money to charity]. Classical judgement internalism doesn't entail that this is irrational at all. It might just be a psychologically astute judgement. Similarly for MP. Classical judgement internalism doesn't seem to show that it would be irrational to make the judgement [p, but I don't believe p]. Again, we can think about this in terms of the first horn of our DILEMMA. TRUTH NORM supplemented by (4) isn't a good explanation of the explanandum. It doesn't explain the absurdity of accepting Moore-paradoxical sentences at all.

This argument demonstrates that classical judgement internalism or enkratic principles *don't* show that it is irrational to accept MP*, and hence MP. It also helps us to see what kind of principle we *would* need to do this. The principle would have to look like this:

(5) If you're rational, then: you won't judge [It is correct to V, but I'm not V-ing].

If this general principle were true, then judging MP*, and hence judging MP, would be irrational. As a consequence, we would have some kind of explanation of the absurdity of accepting MP.

So far so good. The problem, however, is that, there's no very good independent reason to accept (5). As we noted above, it doesn't seem irrational to make this judgement in ordinary non-doxastic cases. One might judge: [I ought to give money to charity, but I don't given money to charity]. I'm not thereby be irrational, at least not obviously so. Certainly, there doesn't seem to be anything really strange about this judgement. It's nothing like judging that climate change is happening, but I don't believe it is happening. So this principle is under-motivated.

We can think about this in terms of the second horn of our dilemma. If we supplement TRUTH NORM with (5) we could, potentially, explain the absurdity of accepting MP. But (5) is an implausible claim about the relation between accepting prescriptions and acting on them. So whether TRUTH NORM is supplemented with (4) or with (5), it

fails to escape our dilemma. So TRUTH NORM doesn't explain the absurdity of accepting Moore-paradoxical sentences.

There is a vast literature on Moore's Paradox that we have not discussed. There are responses available to a defender of TRUTH NORM that we do not discuss in detail. We also do not discuss approached to Moore's Paradox in terms of knowledge-based norms on belief, which may be more promising. Our aim has simply been to highlight how our basic strategy—from discussing deliberative transparency—is prima facie applicable when thinking about whether Moore's Paradox supports TRUTH NORM.

4 Conclusion

We have argued that attempts to use TRUTH NORM to explain deliberative transparency and Moore-paradoxical sentences fail on a common dilemma. As a consequence, we don't think that TRUTH NORM is well-supported. If we are right, then TRUTH NORM can't be the basis for a satisfactory middle-ground in metaepistemology.

There is a more general point in the background. Both of the arguments for TRUTH NORM that we have considered—from deliberative transparency and Moore-paradoxical sentences—are really attempts to show that TRUTH NORM does important explanatory work in the philosophy of mind. More specifically, they are attempts to show that TRUTH NORM explains some mental phenomenon that relates what we can (or cannot) believe to the truth of the content of that belief. We think that this *kind* of argument is *bound* to fail. It is bound to fail because, if TRUTH NORM is to provide a uniquely good explanation of the mental phenomenon in question, it will have to make use of an overly strong claim about the relation between our accepting the prescriptions that supposedly constitute belief, and what we subsequently can or cannot believe. Because of this general structure, we are optimistic that our argumentative strategy will generalise. Specifically, we are optimistic that our argumentative strategy will undermine other attempts to show that constitutive norms on attitudes are justified by their role in explaining the impossibility of mental phenomena.

Notes

1. Though for a defence of something close to this view, see Kyriacou (2018).
2. It should be noted that one could claim that a belief is correct just in case it's true without being committed to TRUTH NORM. TRUTH NORM also claims this norm is *constitutive* of belief. We are concentrating on this latter claim, because only it would be able to provide the grounds of epistemic normativity. If belief is subject to a truth norm but not essentially so, that would be a substantial first-order epistemological fact, but it wouldn't be able to ground epistemic normativity. This point is not often clearly made in the literature on truth norms (though see Greenberg 2017, chap. 2; McHugh 2013, 463; Steglich-Petersen 2013, 279, fn. 2).
3. This point is made by Buleandra (2009, 328–31).
4. To our knowledge, no one in the literature on deliberative transparency defends the claim that epistemic *akrasia* is impossible; the closest is Archer (2017), who defends the claim that it is impossible to both judge one has sufficient evidence for p and refrain from believing p. Some also doubt whether judging one's own beliefs to be incorrect suffices for epistemic *akrasia* (Owens 2002), but this does not matter for our purposes. Our concern is whether such a state is possible, not whether it is *akrasia* accurately so-called.
5. McHugh makes a similar point when he argues that a weaker constraint—which he calls "efficacy"—is only a feature of rational doxastic deliberation (2013, 458–62). Our point is more conditional: *if* the internalism expressed in (2) is appealed to, then only the transparency of *rational* doxastic deliberation can be explained by TRUTH NORM, which means it loses *its* unique explanatory power.
6. It will be helpful here to forestall an objection. We are not claiming here that there is nothing strange, absurd, or irrational about the *situation* in which in which I judge that I ought to give charity, but don't do so. We are claiming that there is nothing necessarily strange, absurd, or irrational about a *judgement* with the content [I ought to give money to charity, but I don't], because it might be a psychologically astute judgement about my moral failings. We show how this distinction matters, and undermines the argument for TRUTH NORM, in Sect. 3.3. (We are grateful to Christos Kyriacou in pressing us to clarify this point).

References

Archer, Sophie. 2017. Defending Exclusivity. *Philosophy and Phenomenological Research* 94 (1): 326–341.

Baldwin, Thomas. 2007. The Normative Character of Belief. In *Moore's Paradox: New Essays on Belief, Rationality, and the First Person*, ed. Mitchell S. Green and John N. Williams, 76–89. Oxford: Clarendon.

Buleandra, Andrei. 2009. Doxastic Transparency and Prescriptivity. *Dialectica* 63 (3): 325–332.

Bykvist, Krister, and Anandi Hattiangadi. 2007. Does Thought Imply Ought? *Analysis* 67 (296): 277–285.

———. 2013. Belief, Truth, and Blindspots. In *The Aim of Belief*, ed. Timothy Chan, 100–122. Oxford: Oxford University Press.

Côté-Bouchard, Charles. 2016. Can the Aim of Belief Ground Epistemic Normativity? *Philosophical Studies* 173 (12): 3181–3198.

Engel, Pascal. 2007. Belief and Normativity. *Disputatio* 23: 179–203.

Engel, Pascal. 2013. In Defence of Normativism about the Aim of Belief. In *The Aim of Belief*, ed. Timothy Chan, 32–63. Oxford: Oxford University Press.

Greenberg, Alexander. 2017. The Truth Norm Account of Justification. Thesis, University of Cambridge. https://doi.org/10.17863/CAM.18315.

Huemer, Michael. 2007. Moore's Paradox and the Norm of Belief. In *Themes from G. E. Moore: New Essays in Epistemology and Ethics*, ed. Susana Nuccetelli and Gary Seay, 142–157. Oxford: Clarendon.

Kyriacou, Christos. 2018. From *Moral* Fixed Points to *Epistemic* Fixed Points. In *Metaepistemology—Realism and Anti-Realism*, ed. Christos Kyriacou and Robin McKenna. Basingstoke: Palgrave Macmillan.

McHugh, Conor. 2013. Normativism and Doxastic Deliberation. *Analytic Philosophy* 54 (4): 447–465.

McHugh, Conor, and Daniel Whiting. 2014. The Normativity of Belief. *Analysis* 74 (4): 698–713.

Owens, David. 2002. Epistemic Akrasia. *The Monist* 85 (3): 381–397.

Pettit, Philip, and Michael Smith. 1996. Freedom in Belief and Desire. *Journal of Philosophy* 93 (9): 429–449.

Railton, Peter. 1994. Truth, Reason, and the Regulation of Belief. *Philosophical Issues* 5: 71–93.

Rosen, Gideon. 2001. Brandom on Modality, Normativity and Intentionality. *Philosophy and Phenomenological Research* 63 (3): 611–623.

Shah, Nishi. 2003. How Truth Governs Belief. *The Philosophical Review* 112 (4): 447–482.

Shah, Nishi, and David Velleman. 2005. Doxastic Deliberation. *The Philosophical Review* 114 (4): 497–534.

Smith, Michael. 1994. *The Moral Problem*. Oxford: Blackwell.

Steglich-Petersen, Asbjørn. 2006. No Norm Needed: On the Aim of Belief. *The Philosophical Quarterly* 56 (225): 499–516.

———. 2013. The No Guidance Argument. *Theoria* 79 (1): 279–283.

Svavarsdottir, Sigrun. 1999. Moral Cognitivism and Motivation. *The Philosophical Review* 108 (2): 161–219.

Wedgwood, Ralph. 2002. The Aim of Belief. *Philosophical Perspectives* 16: 267–297.

Williams, Bernard. 1973. Deciding to Believe. In *Problems of the Self*, ed. Bernard Williams, 136–151. Cambridge: Cambridge University Press.

Correctness and Goodness

Allan Hazlett

In *Truth and Truthfulness* (2002), Bernard Williams articulates and defends a familiar idea:

> [T]ruth has an internal connection with beliefs and assertions. In a sense, truth figures in this connection as a value … The assessment of beliefs and assertions as true is a favourable one. These facts involve valuations in terms of truth. In one direction, this takes us a long way in the direction of truth as a value. (Williams 2002, 84)

But he qualifies this idea with the following argument:

> In another direction, however, all this takes us nowhere at all. If we define the belief-assertion-communication system in the narrowest functional terms, the story so far tells us something of how that system works … But as soon as a *question* comes up, whether an individual or group should on a given occasion continue to work the system—whether, for instance, they

A. Hazlett (✉)
Department of Philosophy, Washington University at St Louis, St Louis, MO, USA
e-mail: allanhazlett@gmail.com

© The Author(s) 2018
C. Kyriacou, R. McKenna (eds.), *Metaepistemology*, Palgrave Innovations in Philosophy, https://doi.org/10.1007/978-3-319-93369-6_9

> should tell a lie—the sense we have recovered so far in which truth is a value gets us nowhere. Indeed … people could not tell a lie unless there were assertions, and utterances would not be assertions unless they were expected to be true, and they would not be expected to be true if they did not, a lot of the time, express true beliefs. All this may be correct, and … people may know that it is correct, but it does not answer their question. They know that they can tell a lie, as every day millions of people do, and the heavens will not fall; if the heavens were going to fall, they would have fallen already. (Williams 2002, 84–85; cf. Shah 2003, 458–59)

Here I would like to defend and elaborate on this argument. What Williams' argument shows is that the fact that truth has an internal connection with belief and assertion does not entail that truth is a value—it does not entail that true beliefs and true assertions are good.[1] The way I will put this is that correctness does not entail goodness. It follows from this that correctness is not a species (or kind) of goodness, and that being correct is not a way of being good.

Here is the plan for the paper: I will articulate the relevant concept of correctness (Sect. 1), argue that correctness does not entail goodness (Sect. 2), consider the idea that correctness is a species of relative goodness (Sect. 3), discuss the connection between correctness and reasons (Sect. 4), and conclude with a brief statement of the significance of my conclusions (Sect. 5).

1 Representational Correctness

What do we mean when we say that truth has an internal connection with belief and assertion? We mean the same as when we say that truth is the "aim of belief" or the "norm of belief" and that truth is the "aim of assertion" or the "norm of assertion." But what do we mean when we say those things?

It is sometimes suggested—and it is suggested by the invocation of "norms"—that this internal connection has something to do with obligation or requirement (cf. Boghossian 2003, 37; Gibbons 2013; Hawthorne 2004, 23; Wedgwood 2002, 268; Williamson 2000, 11, 241). Consider, for example:

i. You ought to believe that *p* if and only if it is true that *p*.
ii. You ought to assert that *p* if and only if it is true that *p*.

But this is clearly not the familiar epistemological sense of "ought," since on that sense there are cases in which you ought to believe something false (e.g., when you have misleading evidence), there are cases in which you ought not believe something true (e.g. when you have no evidence for it), there are cases in which you ought to assert something false (e.g., the case of the murderer at the door), and there are cases in which you ought not assert something true (e.g., when it is something that you do not believe, or when asserting it would be irrelevant, or tactless, or dangerous). Appealing to obligation and requirement in this context, therefore, risks further obscuring an already obscure *explanans*. Let us therefore leave aside obligation and requirement in our search for an articulation of the internal connection between truth and belief and assertion.

We do better when we recognise that belief and assertion are both species of representation that can be correct or incorrect. The relevant senses of "correct" and "incorrect"—call these *representational correctness* and *representational incorrectness*—seem to me to be primitive, so there will be no defining or giving a reductive analysis of these terms. But we can say something to indicate what primitive concepts we mean here.

For "correct" we might as well have said "right" or "accurate" and for "incorrect" we might as well have said "wrong" or "inaccurate" or "mistaken" or "in error." A correct representation is an accurate representation, one that gets things right; an incorrect representation is an inaccurate representation, one that gets things wrong. With representations of this kind there is always the possibility—in the sense of the conceivability, in the sense that we can make sense of the very idea—of correctness and incorrectness. Consider Wittgenstein's (1969) rhetorical question in defense of the idea that "hinges," like "External things exist," are not instances of knowledge: "What would a mistake be like here? And have I any *clear* idea of it?" (1969, sec. 17; cf. secs 51, 54, 70–75, 138, 194–96)

It is imperative, for reasons that will emerge in a moment, that we distinguish the present concepts of correctness and incorrectness from the concepts of obligation, permission, and prohibition. That a

representation is correct does not entail that it is obligatory or permitted (nor does the latter entail the former); that a representation is incorrect does not entail that it is prohibited (nor does the latter entail the former).

If this is right—that belief and assertion are species of representation that can be correct or incorrect—we can articulate the internal connection between truth and belief and assertion. Here is what I have in mind:

iii. Believing that *p* (i.e., the belief that *p*) is correct if and only if it is true that *p*, and incorrect if and only if it false that *p*.
iv. Asserting that *p* (i.e., the assertion that *p*) is correct if and only if it is true that *p*, and incorrect if and only if it is false that *p*.

Note that if we assume the principle of bivalence, these two principles entail the following two principles:

v. Believing that *p* (i.e., the belief that *p*) is correct if and only if it is true that *p*, and incorrect otherwise.
vi. Asserting that *p* (i.e., the assertion that *p*) is correct if and only if it is true that *p*, and incorrect otherwise.

To put this another way, truth is the correctness condition for belief and assertion.

It is sometimes said that truth is the *constitutive* correctness condition for belief, that is, that it flows from the nature or essence of belief that truth is its correctness condition (see e.g. Shah 2003). The present formulation does not commit us to that additional claim, although talk of an "internal" connection suggests it. And it certainly does not commit us to the additional claim that the fact that truth is the standard of correctness for belief *exhausts* the nature or essence of belief.

We can now see why it was imperative to distinguish the present concepts of correctness and incorrectness from the concepts of obligation, permission, and prohibition. There can be correct beliefs that are neither obligatory nor permissible (e.g., true beliefs based on insufficient evidence); there can be incorrect beliefs that are not prohibited (e.g., false beliefs based on misleading evidence); there can be correct assertions that are neither obligatory nor permissible (e.g., saying something true that

you do not believe, or saying something true but irrelevant, or tactless, or dangerous); there can be incorrect assertions that are not prohibited (e.g., lying to the murderer at the door).

I said that belief and assertion are species of representation that can be correct or incorrect. There are other species of this genus. Depictions can be correct or incorrect, depending on whether they are true or false to their subjects, whether they depict things as they are or depict things as other than they are. And perceptual experiences can be correct or incorrect, depending on whether they are—for lack of better words—accurate or inaccurate. These are all representations that can be correct or incorrect.

There is a broad sense of "representation," on which there are representations that cannot be correct or incorrect. Consider the state or activity of merely considering a proposition, without believing it or disbelieving it. It is fair to say that in so doing you represent that proposition, but your representation is not the sort of thing that could be correct or incorrect. Representations that can be correct or incorrect—including beliefs and assertions—constitute a species of representation, in the broad sense of "representation."

The difference between representations that can be correct or incorrect and other representations comes out when we consider borderline cases. Take, for example, imagination. Imagination is a hard case for two reasons. First, some episodes of imagination can be criticised for inaccuracy, as when a paranoid or phobic person cannot stop imagining some horrible calamity; but other episodes, such as instances of daydreaming, cannot be criticised on such grounds. Second, you might think—although it is not clear whether this is right—that imagination represents its object as possible—such that the imagination of something impossible is an instance of incorrect imagination. This reveals something important about representations that can be correct or incorrect, which is that their basic form is the representation of *x* as *y*, for some *x* and *y*. This explains why desire is another hard case: defenders of the "guise of the good" thesis argue that desire represents its object as good, while critics of that thesis follow Hume in arguing that desires are not representations of anything as anything, and therefore cannot be correct or incorrect. Or consider pain: is pain a representation of some part of the body as injured,

and thus a species of representation that can be correct or incorrect, or merely a qualitative feeling, with no such representational content?

Borderline cases involving particular individual things are even harder. Consider the paintings of Mark Rothko. Are they representations that can be correct or incorrect? You might argue that they are instances of "abstract expressionism," that what is expressed is some emotion, and that expression is a species of representation that can be correct or incorrect. But what would the correctness condition for Rothko's paintings be? Some kind of correspondence between them and Rothko's emotions? Or between them and some type of emotion that he was attempting to represent? Alternatively, you might reject (with Rothko) the interpretation of his paintings as instances of "abstract expressionism." If Rothko were to tell us, as he sometimes did, that his paintings were windows, does that suggest that they are depictions of windows (and thus representations that can be correct or incorrect)? Or that the paintings themselves, either literally or metaphorically, are windows (and thus not representations at all)? The fact that these questions do not have obvious answers is part of what makes Rothko's paintings interesting.

I have offered the claim that truth is the correctness condition for belief and assertion as an articulation of the ideas that truth is the "aim of belief" or the "norm of belief" and that truth is the "aim of assertion" or the "norm of assertion." Some contemporary philosophers argue that knowledge, rather than truth, is the "aim of belief," "norm of belief," "aim of assertion," or "norm of assertion" (Benton 2011; DeRose 2002; Gibbons 2013; Hawthorne 2004, 23; McHugh 2011; Williamson 2000, 11, chap. 11; cf. Hawthorne 2004, 85–91; Hawthorne and Stanley 2008). Whatever they mean by that, they cannot be talking about representational correctness, since the following are completely implausible:

vii. It is correct to believe that *p* if and only if you know that *p*, and incorrect otherwise.
viii. It is correct to assert that *p* if and only if you know that *p*, and incorrect otherwise.

The fool who believes and asserts that they will win the lottery does not know that they will win the lottery, even if they will win the lottery, but

in the event that they will win the lottery, their belief and their assertion are correct (i.e., right, accurate), and not incorrect (i.e., wrong, mistaken, in error). There may be some sense in which knowledge is the "aim of belief," "norm of belief," "aim of assertion," or "norm of assertion." But this does not threaten the claim that truth is the correctness condition for belief and assertion.

2 Correctness Does Not Entail Goodness

In this section I will argue that correctness does not entail goodness. To put this thesis another way, and to clarify what I mean, it is not the case that, necessarily, if *x* is correct then *x* is *pro tanto* good. Something is *pro tanto* good if and only if it is good in some way or to some degree; it is characteristic of *pro tanto* goodness that it can be outweighed by other values. In what follows, "goodness" and its cognates will always refer to *pro tanto* goodness.

There are familiar cases of true beliefs that would not be all-things-considered best for the believer, and I have defended the existence of such beliefs elsewhere (Hazlett 2013, chaps. 2–3). After an improbable come-from-behind win against the Cleveland Cavaliers in 2016, Boston Celtics guard Avery Bradley observed: "If you play this game, you should believe you're going to win until the horn goes off." That includes, of course, those games in which the improbable does not happen, and you lose. In such games, a true belief (that you are going to lose) would not be all-things-considered best for you. But you might argue that this is not yet a case of correctness without goodness, on the grounds that the true belief that you are going to lose, although it would not be not all-things-considered best for you (in virtue of the confidence that you would be afforded by believing otherwise), would be good for you—whether you are going to lose is, after all, something that you (presumably) care about, and something that you (so we can easily imagine) want to know. However, in some cases, it is not plausible that true belief would be good for you. Consider trivial questions which are both subjectively uninteresting and objectively insignificant, for example, the question of the number of dots on the ceiling of some bus station. Suppose a traveller

absent-mindedly counts the 548 dots on the ceiling of the bus station (cf. Grimm 2009, 247–49; Sosa 2003, 156). They couldn't care less about how many dots there are, and it is utterly unimportant how many dots there are. If they arrive at the correct answer, and believe that there are 548 dots, it would not benefit them at all. Their belief, although correct, would not be good. There is not some prudential value of true belief that is outweighed, as in the case of Bradley's confidence; there is no prudential value there to begin with.

We have focused on the prudential value of true belief. You might object that the goodness entailed by correctness is not prudential goodness, but rather "epistemic" goodness. Your true belief, in the case described, so you might argue, is not good for the believer, but it is "epistemically" good—it enjoys that species of goodness that attaches distinctively to true beliefs in virtue of their being true. Now if this just means that your true belief is correct (Sect. 1), that is conceded. But if this means more than that, if this means that there really is something good about your belief, then I think we have to ask: what is supposed to be good about your knowing some trivial proposition, something you do not care about, something you do not want to know, and something totally insignificant? We cannot be satisfied merely by the idea that the goodness here is "epistemic"; that is just to give a label to the supposed goodness that we are trying to understand. If someone were to say that some (seemingly worthless) saucer of mud was good, we would want to know what was supposed to be good about it, and if they were to appeal to some analogue of "epistemic" goodness—to "that species of goodness that attaches distinctively to saucers of mud, in virtue of their being saucers of mud"—we would not be satisfied by their explanation. And if we insist at this point that there must be *something* good about your belief, because it is after all correct, that is begging the question. But apart from that question-begging line of reasoning, we have nothing to support the claim that your belief is good.[2]

None of this is to deny that there are true beliefs that are non-prudentially valuable—but the present case is not plausibly a case of that kind. The non-prudential value of true belief, like the prudential value of true belief, seems conditional, attaching to some true beliefs but not to others (cf. Hazlett 2013, chap. 4).

I conclude that correctness does not entail goodness: there are beliefs that are correct but not good.

There are analogous cases of assertion. There are true assertions that are not all-things-considered best, as in the case of the murderer at the door, although you might argue that true assertion in this case would be good—good, perhaps, for the murderer, or something like that. But we can take away any elements that may make it seem like there is something good about true assertion, in that case: consider the irrelevant assertion of a claim about the number of dots on the bus station ceiling by the aforementioned traveller. Again, this does not seem like a case where the goodness of your assertion is outweighed, and we are left wanting some explanation of what is good about your assertion.

You might object that the argument so far conflates the goodness of believing or asserting the truth about whether *p* with the goodness of believing or asserting anything at all about whether *p* (cf. Steglich-Petersen 2011). You might concede that there are cases—like the cases suggested—in which it would not be good to believe or assert whether *p*, but argue that, even in such cases, if you *do* believe or assert whether *p*, it is better that your belief or assertion be true, and thus, in that sense, that it would be good to believe or assert the truth. I disagree. When the question of whether *p* is utterly trivial, it is not better to believe or assert the truth about whether *p*. Compare two travellers who both absent-mindedly count the dots on the bus station ceiling, who come to different conclusions: the first traveller comes to believe that there are 548 dots, and the second traveller 549. What is supposed to be better about the first traveller's belief? Suppose that, upon arriving at their destinations, each traveller announces the result of their count: "There were 548 dots on the bus station ceiling," the first traveller says, while the second traveller says, "There were 549 dots on the bus station ceiling." Again: what is supposed to be better about the first traveller's assertion? In both the case of their beliefs and of their assertions, the difference between them is utterly unimportant, meaningless, insignificant—in other words, it does not seem to be a difference in value.

This is not to deny that there are cases in which, although it would not be all-things-considered best to believe or assert whether *p*, it would be better to believe or assert the truth about whether *p*, if you do believe or

assert whether *p*. You would like to learn something about Borneo right now, but you have an obligation to pick up your children from school—and so your learning about Borneo right now would not be all-things-considered best. If you were to decide to violate your obligation, it would be better to consult an authoritative book about Borneo, as opposed to engaging in ungrounded speculation or fantasy. But the case of the travellers is not like that. There is, we imagine, something good about your forming beliefs about Borneo; but there is nothing good about the travellers forming beliefs about the number of dots on the bus station ceiling.[3]

You might argue that, although correctness does not entail goodness, correctness is *generally*—or *normally*, or *typically*, or *usually*, or *most of the time*—good, and that cases of correctness without goodness are *exceptional*—or *abnormal*, *atypical*, or *unusual*.[4] However, it is unclear whether even this more modest claim about the value of correctness is true. This depends, for one thing, on whether and in what sense trivial beliefs and assertions are exceptional. This much seems plausible: the beliefs and assertions of the two travellers (above) are exceptional for the triviality of their contents, by contrast with most of our actual beliefs and assertions. But is this evidence that correctness is generally good? What may confuse us here is that we are generally reliable at forming correct beliefs that are good and making correct assertions that are good. We seek out interesting or significant truths and assert what is relevant or useful. But it would be a mistake to conclude from this that correct representations are generally good—at least if this is understood as a claim about correctness *per se*, as opposed to a claim about our capacity for making correct representations that are good.

The case of pictures (cf. Sect. 1) is illustrative here. Some contexts call for correct (i.e., accurate) pictures: consider sketches of criminal suspects used by the police, historical paintings, the drawings made of court proceedings, or portraits of politicians and monarchs.[5] But in many unexceptional contexts, accuracy is irrelevant. It simply does not matter if a casual doodle is accurate or inaccurate. Or consider the use of pictures in advertising. In its late nineteenth century and early twentieth century ad campaigns, the Quaker Oats Company used drawings of William Penn in their advertisements. “Here you see the picture of William Penn,

standard bearer of the Quakers and of Quaker Oats," one newspaper ad begins, "He gazes benignly out upon each and every man, woman and child who reads this sage advertisement." We can certainly ask whether the printed "picture of William Penn" was accurate. But that seems beside the point; these depictions were not made for the purpose of historical documentation, but to sell oats. At least in the case of pictures, it is clear that correctness has its uses, but there is no special connection between correctness and goodness. The same, I have suggested, in the cases of belief and assertion.

3 Goodness *qua* Representation

A natural objection to the argument that correctness does not entail goodness (Sect. 2) is that true beliefs are good *qua* belief, that true assertions are good *qua* assertion, and that, in general, correct representations are good *qua* representation.[6] In that sense, you might argue, correctness entails goodness—goodness *qua* representation. Now, as with the appeal to "epistemic" goodness, the claim that true beliefs are good *qua* belief may just be an alternative way of saying that truth is the correctness condition for belief (cf. Sect. 1), and in that case it should be conceded. However, I shall argue here that nothing more than this should be conceded.

Goodness *qua* representation is a species of what is sometimes called "relative goodness." What this means is that, when we attribute goodness *qua* representation—for example, when we say of some belief that it is good *qua* belief—we treat "good" as what Peter Geach (1956) calls an "attributive adjective": that a belief is good *qua* belief does not entail that it is good *simpliciter*. Goodness *qua* representation, therefore, does not entail "absolute goodness." However, you might think that it is nevertheless a species of goodness, such that correctness entails goodness.

That would be a mistake. The best way to see why that would be a mistake is to consider the essential connection between goodness and reasons. Goodness, I submit, is sufficient for reasons (letting φ range over any and all things that can be done, i.e., "doings," including actions, but also believing, desiring, intending, and so on):

> **Goodness is Sufficient for Reasons (GSR)**: Necessarily, if your φing would be *pro tanto* good, then there is a *pro tanto* reason for you to φ.[7]

Therefore, necessarily, if there is no *pro tanto* reason for you to φ, then it would not be *pro tanto* good were you to φ.[8] (Note that this is neutral on the order of explanation between goodness and reasons.) GSR explains why we can sometimes conclude that someone's attribution of goodness is insincere when they seem insensitive to the corresponding reasons. "I recognise the value of switching to carbon-neutral methods of production," insists the CEO of a corporation notorious for carbon dioxide emissions. But there is no evidence that the CEO does recognise that value. Such evidence would take the form of their doing something to promote a switch to carbon neutral methods of production, for example, investing part of the corporation's research and development budget in lowering carbon dioxide emissions. But the CEO in fact does nothing. Given GSR, we reasonably conclude, in this case, that the CEO does not really believe that switching to carbon-neutral methods of production is good.

The problem for the appeal to goodness *qua* representation is that relative goodness is insufficient for reasons: from the fact that it would be relatively *pro tanto* good were you to φ, it does not follow that there is a *pro tanto* reason for you to φ. As Geach puts this point, "if I call a man a good burglar or a good cut-throat I am certainly not commending him," although "one can imagine circumstances in which these descriptions would deserve to guide another mans' choice (for example, if a commando leader were choosing burglars and cut-throats for a special job)" (1956, 36–37). We can conclude from this that "calling a thing a good A does not influence choice unless the one who is choosing happens to want A" (Geach 1956, 37). Relative goodness is, therefore, insufficient for reasons (cf. Thomson 2008, 14–17).

From these two premises, we can conclude that relative goodness is not a species of goodness: goodness is sufficient for reasons (i.e., GSR); relative goodness is insufficient for reasons; therefore, relative goodness is not a species of goodness. Relative goodness *per se* does not give us any reason to do anything—it does not *per se* matter to us, and thus has all the characteristics of something that is *per se* worthless: unimportance, meaninglessness, and insignificance.

An alternative way of seeing this point is to note that, if we could rescue the idea that correctness entails goodness by appeal to the notion of goodness *qua* representation, then we could just as well rescue the ideas (Sect. 1) that you ought to believe that *p* if and only if it is true that *p* and that you ought to assert that *p* if and only if it is true that *p*. We could say that these claims are true, when "ought" is interpreted as "obligation *qua* representation." But such a "relative obligation" would be no obligation at all—or, at least, it would be no obligation such that we would have any reason to oblige. Perhaps we would be saying something true by saying that a cut-throat has an obligation *qua* cut-throat to cut people's throats, but only a cut-throat would treat such an "obligation" as a reason to cut people's throats.

4 Correctness and Reasons

I have argued that correctness does not entail goodness (Sect. 2). I'll now argue that this conclusion undermines a *prima facie* plausible account of a species of reasons for belief and assertion. The account I have in mind says, roughly, that the fact that truth is the correctness condition for belief and assertion entails that there is a reason for you to believe or assert the truth. However, consider some trivial question, for example, the question of the number of dots on the bus station ceiling (Sect. 2). You might plausibly think that, even though truth is the correctness condition for belief, because this question is uninteresting and insignificant, there is no reason for you to believe anything about the number of dots, and so, in that sense, no reason to believe the truth about the number of dots. And the same, *mutatis mutandis*, in the case of assertion. So, if there is a sense in which the fact that truth is the correctness condition for belief entails that there is a reason for you to believe or assert the truth, it must be this: the fact that truth is the correctness condition for belief and assertion entails that, *if you are going to believe or assert something*, then there is a reason for you to believe or assert the truth. We can think of such a reason as a *conditional* reason, and thus articulate the idea like this:

Correctness Entails Reasons (CER): The fact that truth is the correctness condition for belief and assertion entails that you have a conditional reason to believe or assert the truth.

We could alternatively have articulated this idea in terms of contrast: given that truth is the standard of correctness for belief and assertion, you have a reason to believe or assert the truth, rather than the false—but not necessarily a reason to believe or assert the truth, rather than nothing at all.

My argument assumes a controversial premise: that goodness is not only *sufficient* for reasons (Sect. 3), but also *necessary* for reasons, in other words:

Goodness is Necessary for Reasons (GNR): Necessarily, if there is a *pro tanto* reason for you to φ, then your φing would be *pro tanto* good.[9]

Thus, for every reason there is a corresponding instance of goodness. As with GSR, GNR is neutral on the order of explanation between goodness and reasons. It is thus compatible with a "teleological" or "value-based" account of reasons, on which reasons to φ are explained by the fact that φing would be good—consider, for example, the consequentialist idea that right actions have good consequences, and thus that right actions are instrumentally good, or Kant's idea that right actions are instances of good will, and thus that right actions are non-instrumentally good—and with "buck-passing" accounts of value, on which the goodness of φing is explained by the fact that there are reasons to φ. (It is also compatible with "fitting attitude" or "dispositional" accounts of value.) In any case, you cannot have a reason to do something unless there is something good about your doing so.

Why accept GNR? First, GNR explains why goodness is what we cite when we are trying to explain why we have reason to do something. "Vote for Kang," you propose; "Give me some reasons," I reply—and you then proceed to argue that voting for Kang would be good in various ways: you enumerate Kang's virtues, defend the sensibility of his policies, suggest the prosperity and security that a Kang administration would bring about, appeal to the symbolic power of casting a vote for Kang,

explain that it is my sacred duty to vote for Kang, and so on. Without such arguments, you would not have even attempted to give me reasons to vote for Kang. Second, GNR is suggested by the fact that thinking you have a reason to do something requires seeing some way in which it would be good to do so. Unless I am convinced that voting for Kang would be good in some way, I cannot come to think that I have any reason to vote for Kang. If you concede that voting for Kang will have neither good consequences nor be good in itself—and, moreover, that it would not be good in any way—than you have failed to give me a reason to vote for Kang.

We are now in a position where my argument can be articulated: correctness does not entail goodness (Sect. 2); goodness is necessary for reasons (i.e., GNR); therefore, the fact that truth is the correctness condition for belief and assertion does not entail that you have a (conditional) reason to believe or assert the truth, in other words, CER is false. From the fact that it would be correct to φ, it does not follow that it would be good to φ; therefore, given GNR, from the fact that it would be correct to φ, it does not follow that there is a reason for you to φ. We can agree that truth has an internal connection to belief and assertion; we can agree that truth is the correctness condition for belief and assertion; we can even agree that all and only true beliefs are good *qua* belief and that all and only true assertions are good *qua* assertion (Sect. 3). But none of this, by itself, gives us even a conditional reason to believe or assert the truth.

You might object that, although goodness is necessary for *practical* reasons (i.e., reasons for action), it is not necessary for *theoretical* reasons (i.e., reasons for belief). You might concede that, necessarily, if there is a *pro tanto* reason for you to φ, then it would be good were you to φ, when, and only when, φing is *acting*, and not, in particular, when φing is *believing*. Reasons for belief, so the argument might go, unlike reasons for action, are not connected to goodness in this way—for example, evidence that p constitutes a *pro tanto* reason for you to believe that p, regardless of whether it would be good were you to believe that p, and indeed regardless of how things stand with goodness in general. However, it is *prima facie* puzzling why goodness would be necessary for practical reasons and not for theoretical reasons. Consider the idea that, whenever there is a practical reason for you to φ, this is explained by the fact that it would

be good were you to φ. If such an explanation is always required in the case of practical reasons, why isn't such an explanation required in the case of theoretical reasons? Alternatively, consider the idea that, when there is a practical reason for you to φ, this explains why it would be good were you to φ.[10] If the existence of practical reasons generates value, why doesn't the existence of theoretical reasons generate value? The familiar explanation of the supposed difference between practical and theoretical reasons in the contemporary literature appeals to the idea that belief, unlike action, is involuntary (cf. Adler 2002, chap. 2; Kelly 2003; McHugh 2012a, b; Moran 2001, chap. 2; Owens 2003; Shah 2003; Shah and Velleman 2005; Steglich-Petersen 2009, 2006; Williams 1970). But if this is what explains the difference between practical reasons and theoretical reasons, theoretical reasons emerge as an unusual posit. Consider a machine whose essential function—that which it was designed to do, and that which makes it the kind of machine that it is—is to sort red marbles and green marbles; the green marbles go to the left and the red marbles go to the right. Suppose now a green marble comes down the chute. There is perhaps a sense in which the machine has a reason to send the marble to the left: that is where the green marble is supposed to go, after all. But it is equally natural to say that because the sorting behaviour of the machine is involuntary, it does not really have reasons to do one thing or another—the greenness of the marble may *cause* the machine to send it to the left, but this is not the machine's *reason* for sending the marble to the left. The involuntariness of belief does not so much suggest a difference between practical reasons and theoretical reasons as that "theoretical reasons" are not genuine reasons in the first place. (This difficulty has been with us at least since Hume proposed a "sceptical" explanation of our beliefs as determined by innate principles of our animal nature.) However, the fact that belief is involuntary seems to get us no closer to understanding how there could be genuine theoretical reasons that do not entail goodness. The more our beliefs are like the behaviour of the marble-sorting machine, the more it seems that "theoretical reason" is just a *façon de parler*; the less our beliefs are like the behaviour of the marble-sorting machine, the more it seems that theoretical reasons must entail goodness, in the manner of practical reasons.

5 Conclusion

We began with Williams' argument that the fact that truth has an internal connection with beliefs and assertions "takes us nowhere at all" when it comes to, for example, deciding whether to lie. The way I would put Williams' point is that correctness does not entail goodness, which is necessary for reasons. The fact that truth is the correctness condition for belief and assertion does not give us a reason to prefer true belief and assertion to false belief and assertion. And this leaves open the possibility that we have no such reason at all.

My conclusions here are significant for epistemology in at least two ways. First, if CER is false (Sect. 4), then one familiar kind of account of "epistemic normativity" is undermined, namely, the kind of account that explains "epistemic reasons" by appeal to the premise that truth is the "aim of belief." Second, if correctness is not a species of goodness (Sect. 2), then we cannot appeal to the *per se* goodness of correct representation when justifying intellectual policies and practices. Should you inquire carefully, as opposed to recklessly? Should we correct the errors in the outdated textbooks used by our public schools? Should we favour the establishment of an intellectual ethics that is most likely to promote knowledge and understanding? The *per se* goodness of correct representation cannot be part of our answer to these questions.[11]

Notes

1. The idea that true beliefs and true assertions are good is a familiar one, but for contemporary articulations of the (*per se*) value of true belief, see Horwich 2006 and Lynch 2009.
2. The same, *mutatis mutandis*, when it comes to the appeal to intrinsic goodness.
3. You might object that this focus on goodness ignores the possibility that, although correctness doesn't entail goodness, incorrectness entails badness. As William James observed, the goodness of truth and the badness of falsehood are distinct values, in principle capable of being given separate treatments. However, it seems to me that the case that incorrectness

does not entail badness is just as strong as the case that correctness does not entail goodness. Just as there is nothing intuitively good about the belief and assertion of the first traveller, there is nothing intuitively bad about the belief and assertion of the second traveller.

4. Cf. the view (Hazlett 2013, chap. 1) that true belief is generally (normally, typically) better for the believer than false belief.
5. I set aside photographs here; their status as representations is controversial.
6. However, for an argument against this premise, see Thomson 2008, 111–12.
7. "Reasons" here, and in what follows, refers to reasons that are normative (as opposed to motivating), objective (i.e. they do not depend on your evidence or beliefs), and external (i.e. they do not depend on your pro-attitudes).
8. We could just as well put this point in terms of "normativity," following Christine Korsgaard, who writes that "ethical standards are *normative*," such that "when I say that something is good I am recommending it as worthy of your choice" (1996, 8–9).
9. Note that, like GSR, this is neutral on order of explanation.
10. As on a "buck passing account of value".
11. Thanks to Anne Baril for comments on a draft of this material.

References

Adler, Jonathan. 2002. *Belief's Own Ethics*. Cambridge, MA: MIT Press.

Benton, Matthew. 2011. Two More for the Knowledge Account of Assertion. *Analysis* 71 (4): 684–687.

Boghossian, Paul. 2003. The Normativity of Content. *Philosophical Issues* 13 (1): 31–45.

DeRose, Keith. 2002. Assertion, Knowledge, and Context. *Philosophical Review* 111 (2): 167–203.

Geach, P.T. 1956. Good and Evil. *Analysis* 17 (2): 33–42.

Gibbons, John. 2013. *The Norm of Belief*. Oxford: Oxford University Press.

Grimm, Stephen. 2009. Epistemic Normativity. In *Epistemic Value*, ed. Adrian Haddock, Alan Millar, and Duncan Pritchard, 243–264. Oxford: Oxford University Press.

Hawthorne, John. 2004. *Knowledge and Lotteries*. Oxford: Oxford University Press.

Hawthorne, John, and Jason Stanley. 2008. Knowledge and Action. *Journal of Philosophy* 105 (10): 571–590.

Hazlett, Allan. 2013. *A Luxury of the Understanding: On the Value of True Belief*. Oxford: Oxford University Press.

Horwich, Paul. 2006. The Value of Truth. *Noûs* 40 (2): 347–360.

Kelly, Thomas. 2003. Epistemic Rationality as Instrumental Rationality: A Critique. *Philosophy and Phenomenological Research* 66 (3): 612–640.

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

Lynch, Michael. 2009. The Values of Truth and the Truth of Values. In *Epistemic Value*, ed. Adrian Haddock, Alan Millar, and Duncan Pritchard, 225–242. Oxford: Oxford University Press.

McHugh, Conor. 2011. What Do We Aim at When We Believe? *Dialectica* 65 (3): 369–392.

———. 2012a. Belief and Aims. *Philosophical Studies* 160 (3): 425–439.

———. 2012b. Epistemic Deontology and Voluntariness. *Erkenntnis* 77 (1): 65–94.

Moran, Richard. 2001. *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton: Princeton University Press.

Owens, David. 2003. Does Belief Have an Aim? *Philosophical Studies* 115 (3): 283–305.

Shah, Nishi. 2003. How Truth Governs Belief. *Philosophical Review* 112 (4): 447–482.

Shah, Nishi, and David Velleman. 2005. Doxastic Deliberation. *Philosophical Review* 114 (4): 497–534.

Sosa, Ernest. 2003. The Place of Truth in Epistemology. In *Intellectual Virtue: Perspectives from Ethics and Epistemology*, ed. Linda Zagzebski and Michael DePaul, 155–180. New York: Oxford University Press.

Steglich-Petersen, Asbjørn. 2006. No Norm Needed: On the Aim of Belief. *Philosophical Quarterly* 56 (225): 499–516.

———. 2009. Weighing the Aim of Belief. *Philosophical Studies* 145 (3): 395–405.

———. 2011. How to Be a Teleologist about Epistemic Reasons. In *Reasons for Belief*, ed. Asbjørn Steglich-Petersen and Andrew Reisner, 13–33. Cambridge: Cambridge University Press.

Thomson, Judith Jarvis. 2008. *Normativity*. Chicago: Open Court.

Wedgwood, Ralph. 2002. The Aim of Belief. *Philosophical Perspectives* 36 (s16): 267–297.

Williams, Bernard. 1970. Deciding to Believe. In *Problems of the Self*, written by Bernard Williams, 136–151. Cambridge: Cambridge University Press.

———. 2002. *Truth and Truthfulness: An Essay in Genealogy*. Princeton: Princeton University Press.

Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press.

Wittgenstein, Ludwig. 1969. *On Certainty*. Oxford: Oxford University Press.



The Genealogy of Relativism and Absolutism

Martin Kusch and Robin McKenna

1 Introduction

In this paper we bring together two strands of epistemological theorising that have lived separate lives: the controversy around epistemic relativism (e.g. Boghossian 2006) and the debate over the genealogy of knowledge. The latter refers to a method first suggested in Edward Craig's *Knowledge and the State of Nature* (1990), and then further developed in Bernard Williams' *Truth and Truthfulness* (2002) and Steven Reynolds' *Knowledge as Acceptable Testimony* (2017). There are different ways to interpret the genealogical method. On our reading (Kusch and McKenna forthcoming), genealogy comes into its own when used to explain intuitions underlying different epistemological theories. The explanation in question is functional: the *explanans* are ubiquitous human needs; and the primary *explananda* are social practices and institutions of gathering and sharing information about the natural and social world. Specific epistemic

M. Kusch (✉) • R. McKenna
Department of Philosophy, University of Vienna, Vienna, Austria
e-mail: martin.kusch@univie.ac.at; rbnmckenna@gmail.com

© The Author(s) 2018
C. Kyriacou, R. McKenna (eds.), *Metaepistemology*, Palgrave Innovations in Philosophy, https://doi.org/10.1007/978-3-319-93369-6_10

concepts such as *knowledge*, and the intuitions involved in their usage, are theorised as arising in this context.

Recent years have seen a flurry of publications for and against different forms of relativism. As far as epistemic relativism is concerned, the most influential discussion is Paul Boghossian's *Fear of Knowledge* (2006). Boghossian attacks what he regards as unacceptable forms of epistemic relativism in Richard Rorty's *Philosophy and the Mirror of Nature* (1981) and the Sociology of Scientific Knowledge (=SSK) (e.g. Barnes and Bloor 1982). Boghossian's discussion of relativism foregrounds two ideas: that the relativist denies the possibility of absolute standards of epistemic justification ("non-absolutism"); and that she assumes a plurality of epistemic systems (of epistemic standards). At least in some cases these epistemic systems license incompatible verdicts on epistemic justification ("pluralism").

A second relativism-debate has emerged over the last fifteen years in philosophical semantics. One key question concerns the semantics of knowledge attributions. "Contextualists" (e.g. DeRose 2009) hold that the meaning of "to know" (and its cognates) varies with the context of utterance; "invariantists" hold that "to know" has the same meaning in all contexts, and that knowledge attributions are true or false absolutely (e.g. Brown 2006; Rysiew 2001). "Relativists" (e.g. MacFarlane 2014) hold that, while "to know" has the same meaning in all contexts, knowledge attributions are true or false only relative to the "context of assessment."

We shall bring genealogy to bear on both of these debates, though we focus on the first. We shall call the (anti-)relativism at issue in Boghossian "B-(anti-)relativism" and the (anti-)relativism at issue in MacFarlane "M-(anti-)relativism." We begin by giving a brief explanation of the genealogical method. We then give genealogical accounts of the intuitions underlying B-(anti-)relativism and M-(anti-)relativism. Up to this point, our treatment of all sides to these disputes will be even-handed: the genealogist seeks to explain the intuitions underlying each view, not to undermine them. In the final section we investigate whether the genealogist can (or should) maintain this neutral stance. Our tentative conclusion is that genealogy is closer to B-relativism than to absolutism.

2 Genealogy—A Primer

In *Was Wir Wissen Können* Craig situates his project in close proximity to two unlikely bedfellows: Wittgenstein and natural science (1993, 37). Wittgenstein is an ally since he opposes conceptual analysis in terms of necessary and sufficient conditions, studies the function of concepts, and introduces the category of family-resemblance concepts. Craig's project has affinities with natural science in its method of hypothesis testing, the search for explanation, and focus on evolution. We would add *model-building* to the list: the building of simplified models of complex target systems.

Craig's model-construction has two stages: the first focuses on the epistemic state of nature, that is, a small community of language-using humans, engaging primarily in face-to-face communication, who are co-operative, dependent upon one another for information, and of unequal skills and talents. The central question is: Why would a concept like *knowledge* be introduced under these idealised—simplified and distorted—conditions?[1] Craig answers that people in this situation have a salient need, to wit, to pick out and flag good informants. And the concept used to flag good informants is the core—or one central aspect—of *knowledge*.

In the state of nature, individuals depend upon one another for information. Distinguish between the inquirer, who needs information that they are currently unable to obtain themselves, and the informant, who offers such information. Inquirers must be able to separate good from bad informants. And it is natural to assume that meeting this need will involve concepts. Assume that the concept of a *protoknower* is the central conceptual tool for dealing with this problem; our ancestors used this concept as a tag for good informants. Which conceptual components should *protoknower* contain? Craig's answer is that the concept *protoknower* (whether *p*) comprises these elements:

(i) being as likely to be right about *p* as the inquirer's current needs require;
(ii) being honest;

(iii) being able to make the inquirer believe that *p*;
(iv) being accessible to the inquirer here and now;
(v) being understandable to the inquirer; and
(vi) being detectable as a good informant concerning *p* by the inquirer.

To elaborate briefly on (vi), the inquirer needs to find indicator-properties that she can detect and that correlate closely with holding a true belief, or telling the truth, as to whether *p* (1990, 25, 135). "Being at the top of a tree" might be such a property for some inquirers in the state of nature when *p* is the proposition that a tiger is approaching the village. Usually more than one property will be involved. The properties that make Fred a medical protoknower are not one but many.

Craig is adamant that (i) to (vi) are not necessary and sufficient conditions. While all these elements are present in prototypical situations, the concept has a use even when some elements are missing. Finally, *protoknowledge* differs from *knowledge* in that: (a) only the former is closely tied to testimony; (b) *protoknowledge* is indexed to the capacities and needs of specific inquirers (1990, 90); (c) *protoknowledge* can only be ascribed to others, not to oneself; and (d) protoknowledge is not undermined by accident or luck.

Craig goes to great lengths to show that his model of the epistemic state of nature passes the test of (what the philosophy of scientific models calls) "external validation." He does so by arguing that his model predicts and explains several features of our concept(s) of *knowledge* that have been identified in various philosophical theories. These theories are often seen as excluding one another, but Craig thinks that his model can partially vindicate all of them: they contradict each other only if we overgeneralise them.

This brings us to the second half of the genealogical just-so story: the hypothetical social-historical narrative that takes us from *protoknowledge* to *knowledge*. Craig speaks of this development as a process of "objectivisation" of *protoknowledge*. Key steps in objectivisation are the following. First, *protoknowledge* comes to be used in self-ascription. In response to the question "who knows whether *p*?" group members start to investigate their own indicator-properties. Second, inquirers begin to recommend informants to others. This can be done in a helpful manner only if the

perspectival or indexical character of *protoknowledge* is weakened. The recommended informant must be good in the eyes of both the recommender and the recipient of the recommendation. Further movement in this dimension—recommending an informant to ever more inquirers—makes protoknowledge increasingly harder to get. The endpoint is the idea of "someone who is a good informant as to whether *p* whatever the particular circumstances of the inquirer … That means someone with a very high degree of reliability, someone who is very likely to be right—for he must be acceptable even to a very demanding inquirer" (1990, 91). And a very demanding inquirer will not accept epistemic luck or accident. Third, inquirers begin to use "being recommended" as an indicator property. This move dilutes the original detectability requirement. Inquirers begin calling someone a "protoknower" even when none of the original "natural" indicator-properties is in sight. Fourth, in the context of group action inquirers cease to care whether the needed information is accessible to them as individuals; they are satisfied if it is accessible to someone in the group. As a result they will speak of "protoknowledge" even outside the context of testimony. The process of objectivisation ends with our concept of *knowledge*: "The concept of knowing … lies at the objectivised end of the process; we can explain why there is such an end, and why it should be found worth marking in language" (1990, 90–91).

The second stage of Craig's model construction adds a *dynamic* dimension. The dynamic model takes the epistemic state of nature as its starting point and tracks how the concept *knowledge* would evolve and diversify as the simplifications and distortions of the state of nature are removed step by step. This suggests that the dynamic model is really a form of de-idealisation.

The dynamic model too needs to pass muster as far as external validation is concerned. Craig suggests that it correctly predicts, or at least makes sense of:

- contexts with very high epistemic standards (1990, chap. 10),
- intuitions about lottery propositions (1990, chap. 11), and
- our conflicting intuitions about epistemological scepticism (1990, chaps. 12–13).

As Craig emphasises more clearly in (2007) than in (1990) or (1993), this is not to be taken as a historical thesis: the epistemic state of nature is not a historical period "like the Pleistocene." It is rather a ubiquitous and important type of social-epistemic situation that one is likely to find in all human communities, past and present (2007, 191). This suggests that what the two models present as different stages in the historical development of *knowledge* are really two different types of situation that we experience from day-to-day. In some situations, we are still in the state of nature; in other situations, we are at various stages of the process of objectivisation. Note however that this interpretation of Craig shows that his talk of a "core" of *knowledge* might be misleading: if the other uses co-exist, why assume that "flagging good informants" is more fundamental than the other uses? Or put differently, why assume that the right model for the conceptual development is an avalanche rather than a phylogenetic tree? The avalanche model suggests a small conceptual stone rolling down the snowy (semantic) mountain, in the process putting on layer after layer of further conceptual features. The phylogenetic tree is without a core. We do not think of *homo erectus* as the core or essence of *homo sapiens* just because *homo sapiens* developed out of *homo erectus*. Of course, to keep with the analogy we here assume counterfactually that *homo erectus* might still be alive today.

Up to this point our exposition of genealogy has focused on Craig. But for certain purposes it is important to complement his account with the developments introduced by Williams (2002) and Reynolds (2017).

From Reynolds we adopt an account of how we collectively impress upon each other the need to testify only what one knows. His guiding thought is that it would be exceedingly laborious for inquirers to keep track of the different degrees of reliability of several informants concerning numerous topics of interest. Clearly, it would be better for inquirers if, prior to volunteering their testimony, informants checked whether their views are appropriately based on the right kind of experience, appropriate norms of epistemic conduct, and good judgment. In short, it would be best if informants internalized the correct testimonial norms (2017, 50). Reynolds suggests that we could collectively bring it about that informants engage in time-consuming epistemic self-monitoring via reward and punishment. When informants tell the truth in relevant ways,

we repay them with praise and reciprocation. When informants don't, we sanction them with criticism and lack of co-operation. This practice, Reynolds continues, would eventually lead to the emergence of a special vocabulary "to express this approval and disapproval and to indicate the conditions that tended to lead to it" (2017, 57). Our concept of *knowledge* is part of this vocabulary.

Williams adjusts Craig's genealogical method in one crucial respect. According to Williams, Craig's "imaginary genealogy" needs to be complemented by "real genealogy," that is, by an engagement with historical and cultural contingent realities. Although Williams does not put it in these terms, one can read him as offering an imaginary and real genealogy of the *social institution* of testimony. Thus Williams shares a method with Craig, but he is more interested in the *values*, *motivations* and *virtues* that underpin testimony than in *concepts*. According to Williams, the central virtues of testifiers are accuracy (a disposition to seek the truth and to report it) and sincerity.

Unlike Craig, Williams does not assume that individuals in the state of nature are co-operative and eager to offer information to others. Williams recognises that the institution of testimony is a collective good. Individuals who are rational in a purely self-interested way will try to free-ride: they will seek to obtain accurate and sincere testimony from others without offering anything in return. After all, collecting useful information usually involves costly "investigative investments" (2002, 88). How is the problem of collective action solved? The core of Williams' solution to the free-rider problem is the suggestion that accuracy and sincerity (and with them the institution itself) must come to be regarded by community members as *shared intrinsic*—rather than as merely *instrumental*—*values* (2002, 90). For community members to have trust in others' reports, they must be convinced that accuracy and sincerity are non-negotiable. And this implies that these values are not—except under extreme circumstances—weighed against, and possibly outweighed by, other interests and values (2002, 91). Williams holds that values come in socially shared webs and systems, that intrinsic values occupy a central position in such webs or systems, and that values make sense only in their essential relations to one another (2002, 92). Williams' view of intrinsic values relates directly to his distinction between imaginary and real genealogies. While

no community can exist without values like accuracy and sincerity, different communities embed such values in different "wider range[s] of values." And what these wider ranges of values are, "varies from time to time and culture to culture" (2002, 93).

Williams teaches that the collective good of the testimonial institution can exist only if the virtues of accuracy and sincerity (amongst others) become intrinsic values. But how can communities make sure that their members look upon these virtues as intrinsic values? Williams has an answer to this problem as well. He tells us that "people may be discouraged or encouraged, sanctioned, shamed, or rewarded with respect to this" (2002, 44); that a crucial motivation is the "fear … of disgrace in one's own eyes, and in the eyes of people whom one respects and who one hopes will respect oneself" (2002, 116), that "the motivations of honour and shame play an important part" (2002, 120), and that the structure "of mutual respect and the capacity for shame in the face of oneself and others, is a traditional, indeed archaic, ethical resource, but it is still very necessary" (2002, 121). In a nutshell, Williams' idea is that a system of sanctioning—of honouring and dishonouring—encourages community members to constitute and maintain the institution of testimony and thus the virtues of accuracy and sincerity.

Putting this in different terms, sanctioning in support of the collective good works through the "deference-emotion system" (Scheff 1988). The precondition of this system is our emotional need to continuously monitor how others treat and think of us. We respond to our assessments of this treatment by changing our position on an internal scale that ranges from pride to shame. When we believe that others treat us with deference, when we believe that others honour us, we feel pride (and related feelings) and move ourselves up on the pride-shame scale. When we suspect that such deference and honouring are missing, we tend to feel bad about ourselves and slide downwards towards the shame end of the scale. This emotional dependence on others is exploited by the deference-emotion system. The granting or withholding of deference constitutes a subtle system of social sanctions, a system that we barely notice. And the operation of the deference-emotion system is inseparable from our ongoing conversation about the collective good. Working with and through

the deference-emotion system, this conversation continuously re-establishes the importance of the collective good in everyone's mind.

Finally, we can connect Williams' and Reynolds' proposals as follows. Attributions of (proto-)knowledge (and their cognates) play a key role in the collective action that constitutes the institution of testimony. They do so by *honouring* informants. In keeping with the genealogical method, let us first see how the link between protoknowledge attributions and honouring works under the simplified conditions of the state of nature. To publicly apply the concept *protoknower* to someone is not only to classify them as a reliable source of information, it is also to honour them, or to encourage others to do likewise. To classify someone as a *protoknower* is to praise them for their contribution to the institution of testimony, and thereby for their contribution to the well-being of the community itself. After all, the community cannot survive without the institution of testimony. By means of attributions of protoknowledge members of the community honour good informants for contributing their part to the existing and flourishing of the community. *Mutatis mutandis*, withholding or denying protoknowledge is a way of censoring and dishonouring. It is to mark someone as not willing, not worthy or not able to participate in the constitution of the collective good, and thus as not fit to be a (working) part of the group. In that sense, to deny someone *protoknowledge* is to expel them from the group.

Our suggestion connects Craig's focus on conceptual needs of the inquirer with Williams' emphasis on the motivational problems of the informants. By using the concept of *protoknowledge* to both classify and honour informants, the inquirer manages to serve two key goals at once: the goal of tagging good informants for future reference (to herself and others), and the goal of motivating community members to make, or keep making, investigative investments.

3 A Genealogy of B-(Anti-)Relativism

To offer a (Craigean) genealogy of an epistemological view is to identify and explain the intuitions underwriting the view. The explanation tells us why these intuitions are natural given the context of a community of

humans operating with and maintaining the social institution of testimony. We now turn to the task of offering such explanations for five views: B-absolutism, B-relativism, invariantism, contextualism, and M-relativism. We begin with B-absolutism.

It is helpful to distinguish between B-absolutism and "monism." Monism with respect to a domain D is the view that there is only one set or system of norms operating in D. The system might be more or less complex, and contain more or less distinct principles. But monism holds that all these principles cohere with one another, and do not—when interpreted correctly, or applied singly or in combination—contradict one another. The opposite of monism is "pluralism." Absolutism takes a crucial step beyond monism: it holds that the one and only system of norms has the property of being uniquely and absolutely correct. Non-absolutist monists refuse to take that step. They acknowledge that there is only one system yet do not draw absolutist conclusions. They might offer an evolutionary explanation why we have the system we have, or they might hold that it is a sort of convention (cf. Hazlett 2014). We are interested in the intuitions that support either one, or both, of these views.

We begin with monism. Here Craig's account of objectivisation is relevant. As objectivisation progresses, the standards for useful testimony come to be set in ways that increasingly abstract from the particular circumstances of inquirers. In the process, testimonial norms concerning correct investigative investments, appropriate levels of accuracy and sincerity, get more demanding. To make sure that the information gained and testified to is of use to an ever-wider range of differently situated inquirers, informants must be motivated to meet ever higher demands of reliability. This development is clearly inseparable from the emergence of strongly monistic ways of thinking about testimony. What *every* inquirer, however situated, needs, can no longer vary from one context to another. Objectivisation pushes all parties involved into assuming—at least tacitly and on an intuitive level—that there is but one epistemic system of norms.

We can complement the Craigean genealogy of monism with themes from Reynolds and Williams. Recall the idea that the social institution of testimony is maintained and secured via the deference-emotion system. For this to work, the system of norms used for evaluating epistemic

performances had better not to be excessively complicated and context-sensitive. If it were complicated and context-sensitive, it would be beyond the ken of most non-specialists. And, at least under the simplified conditions of the initial genealogical setting, epistemic specialists do not feature in the story. Monistic intuitions—giving support to one epistemic system binding all—are the obvious solution to this problem of evaluation.

This intertwining of epistemic norms and the deference-emotion system can also be used to give a genealogy of absolutism. Adherence to the one and only epistemic system needs to be secured by continually re-establishing its unique value and significance in everyone's mind. Community members are likely to go about this task by giving the values underwriting their epistemic system a central place in their overall web of values. These values therefore come to be regarded as intrinsically valuable, and as beyond all instrumentalist weighing up of different goals. The endpoint is that the epistemic system is itself seen as absolute—as not relative to anything else.

Let us now turn to B-relativism. How might pro-B-relativist intuitions live alongside the absolutist intuitions mentioned above? Again, we will distinguish between two steps: the step towards a plurality of norms, and the further step to denying B-absolutism.

Craig's imaginary genealogy starts with differently situated inquirers focusing exclusively on what is useful to them, taken individually. At this stage, the needs of each individual are the standard against which the utility of information is measured. This clearly is a pluralist stage. On our understanding of Craigean genealogy, this stage is never completely left behind. Even after objectivisation, we still often assess testimony in terms of what we want in our own very specific contexts. In other words, our highly idiosyncratic needs and the idealised abstract needs of the assumed very demanding inquirer exist side by side. Sometimes they overlap, often they do not. Of course, the two endpoints of the scale—the wholly idiosyncratic and the wholly universal—are both idealisations. Most epistemic evaluations involve the informational needs of different groups of varying size and composition. To be a competent social actor is to have acquired the competence to judge which testimonial norms are relevant for which group. Subject matter, the required degree of reliability, the

costs attached to false beliefs, and much else, are relevant variables. The result will be a variety of testimonial norms.

Admittedly, a *variety* of norms is still compatible with monism. We get pluralism only once this variety is partitioned into distinct sets such that different sets give incompatible answers to the same epistemic questions. One system allows one to judge that, in light of Jones' investigative investment concerning *p*, Jones is a good informant as to whether *p*. Another system compels one to say that despite his investigative investment Jones is *not* a good informant in saying *p*. It is plausible to think that the idea of such distinct epistemic systems goes hand in hand with a specific social change: to wit, the forming of distinct epistemic communities with a clear sense of their distinctive shared individual, or collective, goals or needs. There is nothing in Craig's, Reynolds' or Williams' genealogies that makes the emergence of such subgroups intelligible. But we need not add much to the original versions to create the space for this obvious idea. After all every social theory of the development of human societies emphasises the crucial role of division of labour. And division of epistemic labour is tantamount to a plurality of epistemic communities with distinct standards.

Again, it is helpful to connect this to our collective operation of the deference-emotion system. Consider what might happen when we sanction someone by refusing to accept their testimony, and chastise them for their lack of investigative investment. This will often lead to an aggressive encounter where criticisms fly back and forth. The resulting conflict might endanger social cohesion and peace. The bad informant might well refuse to correct their ways, and be unwilling to share their information with us even on later occasions where their investigative investment is, by our lights, exactly at the required level.

The potential social explosiveness of distrust creates pressure to help the other save face. Face saving might take different forms. We might try to find mitigating circumstances: perhaps the testifier gained their (bad) information in ways we find excusable. Or we might try to find ways of holding that, by their lights, they were justified: perhaps the information was obtained from a usually reliable source, or they were relying on a source they were *prima facie* entitled to trust since they had no evidence suggesting otherwise. Perhaps our testifier acquired a system of epistemic

norms different from ours by trusting their elders, which is presumably how we acquired our system of norms. We therefore have a variety of ways to avoid sanctioning them. We might let things rest and accept that, at least for certain questions, they should not be our first port of call.

Epistemic division of labour and epistemic face-saving incline epistemic agents towards the intuition that there are different epistemic systems at least some of which might give conflicting verdicts on the same issues. The move from this pluralistic stance to the denial of absolutism is of course a further step. It is not obvious that it is inevitable; though once pluralist intuitions are in place, and strong, it may be irresistible.

4 A Genealogy of M-(Anti)-Relativism

We can also give a Craigean genealogical explanation of the intuitions underlying three standard views of the semantics of knowledge attributions: invariantism, contextualism and (M-)relativism.

A brief explanation of these views will suffice here. There are two dividing lines in the literature on the semantics of knowledge attributions. The first concerns the meaning of "to know" (and its cognates). Contextualists hold that uses of "to know" mean different things in different contexts, in much the same way as uses of indexicals like "I." In contrast, invariantists and M-relativists hold that "to know" is univocal: it means the same thing in all contexts. The second dividing line concerns whether knowledge attributions express propositions that are absolutely true or false. While invariantists and contextualists disagree about the contents of the propositions expressed by knowledge attributions, they agree that these propositions are absolutely true or false. In contrast, M-relativists hold that the propositions expressed by knowledge attributions are only true or false relative to a context of assessment. Thus, a knowledge attribution may be true as assessed relative to one context, but false as assessed relative to another.

In the literature one can find advocates of all three positions appealing to Craigean genealogy (see Kelp 2011 for invariantism, Henderson 2009 for contextualism and MacFarlane 2014, chap. 12 for M-relativism). Recall that we are currently interested in genealogical explanations of the

intuitions underlying philosophical views, not genealogical *vindications*. But the accounts given by Kelp, Henderson and MacFarlane give good indications of how to generate genealogical explanations of their respective positions.

We can start by noting that contextualism fits perfectly with some aspects of Craig's genealogy. It seems clear that whether a subject is a good informant depends on and varies with the context. If I'm talking to a friend over lunch about Isla's whereabouts last night and I have good but not conclusive evidence that she was at the party I'll volunteer myself as an informant on her whereabouts. However, if I'm giving a statement to the police and I have the same evidence I'll not volunteer myself as an informant on her whereabouts. This suggests that, because what one will require of a good informant depends on and varies with the context, what it means to say someone "knows" must depend on and vary with the context too. Now, objectivisation clearly complicates this picture. But recall Craig's remark that, in some situations, we are still in the state of nature. One way of thinking of objectivisation (suggested by Henderson, and by our understanding of Craig) is as pushing us towards co-ordinating our uses of "to know" in the majority of situations, but allowing us to use the word in ways that are appropriate to our particular situation when the need arises (e.g. when the stakes are particularly high).

However, contextualism fits badly with the central role of testimony in Craig's (and Reynolds' and Williams') genealogy. If uses of "to know" mean different things in different contexts, then it is not helpful to report that someone else knows (Hawthorne 2004). I may be told that Isla said that Morven knows the bank is open, but unless I also know what epistemic standards Isla was using, this information is useless. Compare: I may be told that somebody said "I am tired" but, unless I know who said this, this is useless if I want to keep track of who is tired. This suggests that, if we are to have a functioning testimonial practice, we need a univocal semantics for "to know."

Enter the M-relativist. MacFarlane thinks that M-relativism can utilise both the aspects of Craig's genealogy that seem to support contextualism and the aspects that seem to support invariantism. The contextualist highlights the fact that we may require different things of good informants in different situations and reasons that these requirements will lead

us to use "to know" in different ways in different situations. But this fact might equally well be taken to support M-relativism: while words like "to know" (or "good informant") mean the same thing in all contexts of utterance, whether it is true that someone "knows" (or is a "good informant") is relative to the context of assessment. MacFarlane buttresses this point by arguing that relativism is more plausible than contextualism because it places less demands on our memory. He says that, if M-relativism were true, "[t]here would be no need to store a standard with each knowledge attribution, because all of the knowledge attributions would be evaluated in relation to the current standard" (2014, 312). To hammer this point home MacFarlane even offers an evolution of assessment sensitivity: it may well have been that "once upon a time, 'knows' behaved just as contextualists say it does" (2014, 317). But as social interaction increased and knowledge attributions were exchanged ever more widely across situation and standards, it simply became too tedious to keep track of the standards attached to each attribution. And thus speakers drifted towards the M-relativist understanding.

And yet, while there is a plausible genealogical explanation of why we should expect to have intuitions supporting M-relativism, there is also a plausible story to be told that goes the other way. There are grounds for doubting whether M-relativism is really compatible with the role of testimony in a genealogical account. Consider Reynolds' idea that informants need to internalise the correct testimonial norms. This is difficult to fit with M-relativism: the M-relativist claims testimonial norms require us to assert only propositions that are true relative to our present context of assessment (see MacFarlane 2014, chap. 12), but if we follow these norms our testimony may be of little use to inquirers in different contexts of assessment. Further, M-relativism poses more general problems for our testimonial practices (Rysiew 2012). We often rely on reports to the effect that someone knows that *p*. But, for the M-relativist, these reports will be true relative to some contexts of assessment and false relative to others. Should the inquirer rely on these reports? Either the M-relativist has to hold that we shouldn't, or they have to allow that we need a way of keeping track of which contexts of assessment reports are true relative to. And thus the seeming advantage of M-relativism over contextualism disappears.

5 Genealogy and Neutrality

Let us take stock. We have argued that a genealogical account can be given of the intuitions underlying five philosophical views: absolutism, B-relativism, invariantism, contextualism and M-relativism. Some of these views cannot be adopted together. For instance, absolutism and B-relativism are incompatible. This prompts three questions about the relationship between (Craigean) genealogy and relativism:

(i) Can genealogy maintain a neutral stance towards these views?
(ii) Should genealogy maintain a neutral stance?
(iii) If genealogy cannot (or should not) maintain a neutral stance, which of these should it favour?

We take each question in turn.

Question (i): Genealogy is primarily a *methodology*. Its aim is to make sense of different epistemological intuition-pumps as natural responses to very general social-epistemic situations. On Craig's account, *knowledge* was introduced in order to serve certain deep-seated human needs related to these situations. But it is far from obvious that there should be only one single way in which these needs can be met. The different genealogical rationales for both absolutism and B-relativism (or for the various views about the semantics of knowledge attributions) make this point vivid: incompatible philosophical positions can be understood as motivated by the same set of general needs. Moreover, in providing genealogical explanations for different, or even opposed, intuitions, we are not necessarily taking a stand on the correctness or rationality of these intuitions. We can understand something as a response to a need without taking the need (or response) to be rational. For instance, you can view a politician's provocative statements as a response to a need for attention without regarding this response as rational, or as a good thing.

A comparison with the "impartiality" and "symmetry" principles of the Strong Programme of the Sociology of Scientific Knowledge (SSK) can help strengthen this point (Barnes and Bloor 1982; Bloor 1991). The Strong Programme "would be impartial with respect to truth and falsity, rationality or irrationality, success or failure" and it "would be

symmetrical in its style of explanation. The same types of cause would explain say, true and false beliefs" (Bloor 1991, 7). Thus the proponent of SSK offers the same general types of explanations of theory acceptance in science regardless of whether the respective theories are true or false by our lights. For instance, SSK gives symmetrical sociological explanations of the beliefs of all sides in controversies like the dispute between Boyle and Hobbes over the "Spring of air" or the possibility of a mechanically produced vacuum (Shapin and Schaffer 1985).

Genealogy does something similar, albeit at a different level. Where SSK tends to focus on *particular* and *contingent* historical controversies, the genealogist starts from an *abstract*, *idealised* and *simplified* social scenario—the epistemic state of nature—and then imagines how human needs related to these scenarios lead to the development of one or more concepts. Both forms of analysis are impartial and symmetrical.

And yet, this comparison prompts a worry. Isn't genealogy just an empirically unconstrained form of philosophical speculation? Why engage in genealogy when we have SSK (and related forms of social theory and anthropology)?

At this point we dig in our genealogical heels. The worry is over-stated. First, genealogy is not alone in thinking through abstract and simplified social scenarios. Social theory, the abstract and highly theoretical end of social science, does likewise. Take game theory, for example. Few social scientists are inclined to dismiss game theory out of hand just because it abstracts away from details about the social world. Second, genealogy is not entirely unconstrained by empirical information. In Craig's version, we build on information about supposedly universal human needs. In Williams's version, we add further assumptions about human practices, collective goods, free-riding, values, commitments, and much else. True, Craigean genealogy does not study actual, concrete, dated events. But that is because it investigates ubiquitous and repeatable *types* of events. Finally, recall Williams' insistence that genealogy ultimately has to "make contact" with actual historical writing. Genealogy is not complete until it has identified the actually endorsed values underlying our epistemic institutions, until it has pinpointed the actual values to which accuracy and sincerity are linked. Going beyond Williams' own writings, one might argue, for instance, that Steven Shapin's (1994) study of Boyle's scientific

testimony brilliantly brings out the importance of gentlemanly values and conventions surrounding the reporting of facts (Kusch 2009).

Question (ii): Let us assume that there is a sense in which genealogy can be neutral—it is an explanatory method for making sense of conflicting intuitions. Alas, this does not yet show what exactly would be *valuable* about neutral genealogy. Maybe, if it is to have any interest, a genealogy needs to take sides. The general worry here is that, while an explanation of where our intuitions came from may be of historical or sociological interest, it isn't of philosophical interest.

We are not convinced. If philosophy is anything, it is an exercise of critical self-reflection. And surely, it must be part and parcel of such self-reflection to ask where our (most deeply held) intuitions come from. Their origins may be important for all sorts of reasons. For instance, if Williams is right, understanding the origins of our institution of testimony is crucial for understanding the twin values of sincerity and accuracy. More generally, if we see social epistemology as in some sense continuous with social theory, we surely must care about the social origins of our epistemological intuitions and institutions.

An analogy might be helpful. We see the role of genealogy as being roughly akin to the role of experimental philosophy when it asks which psychological mechanisms and processes influence our judgements. The experimental philosophers do not just claim that it would be interesting to find out what drives our judgements; they claim that their findings are philosophically important. Take, for instance, psychological work on the cognitive processes underlying our judgements about whether subjects "know" in bank cases (and other similar cases). Some of this work explains intuitions appealed to by contextualists in support of their semantics for knowledge attributions (see e.g. Gerken 2013; Nagel 2008, 2010). The underlying idea is that epistemology is continuous with psychology, and that therefore the epistemologist must care about the psychological origins of our epistemological intuitions. We want to insist that the same goes for epistemology and social theory.

One might object that, while this analogy may clarify the philosophical relevance of genealogy, it does so at the cost of compromising its neutrality. After all, Gerken and Nagel's work is generally seen as an argument *against* a contextualist account of the semantics of knowledge

attributions. We grant that this is both Gerken's and Nagel's intention. But we reject the inference against neutrality. Why should it be assumed that we could not use the experimental method to explain pro-invariantist intuitions too?

Nevertheless, we do not insist on the complete neutrality of genealogy regarding the five positions at issue in this paper. Instead, we shall ask: if genealogy has to take sides, which position should it favour?

Question (iii): We are going to focus on the absolutism/B-relativism dispute. The genealogist who "takes a side" in this dispute seems to have (at least) four options:

(a) The genealogical explanation debunks both absolutism and B-relativism (both are "empty ideas", cf. Unger 2014).
(b) The explanation leads to philosophical quietism (the question of the correctness of these views never arises).
(c) The explanation vindicates both absolutism and B-relativism.
(d) The explanation vindicates one view, but not the other.

We think the genealogist should plump for (d), and this vindicates B-relativism over absolutism.

The problem with (a) is that it conflates *origin* and *validity*. It assumes that just because our intuitions have their origins in contingent social settings, our intuitions are worthless. The accusation of worthlessness falls back on this very argument. (This seems like an instance of the genetic fallacy. Cf. Srinivisan 2015).

The problem with (b) is that it is unstable. On one rendering, it collapses into option (a): there is nothing to say because both views have been debunked—they were both wrong, and it was a mistake to have the dispute in the first place. On another reading, it amounts to a form of "second-order" relativism: there is nothing to be said since both views "have their place." But it is hard to see how a resolution of the absolutism/B-relativism dispute that involves second-order relativism could possibly be even-handed. Thus the second reading of (b) is really tantamount to a version of (d).

Option (c) is problematic for the same reason. Absolutism and B-relativism can't both have their place: absolutism is committed to non-absolutist views being thoroughly mistaken.

This leaves alternative (d). There are two ways of arguing that, if genealogy takes the fourth option, it will end up debunking absolutism. The first is relatively direct. Start with Williams' solution to the free-rider problem: in order to stop free riders taking advantage of the institution of testimony, we impress upon each other the need to regard accuracy and sincerity as intrinsic, non-negotiable values. Whatever the merits of this solution, it clearly involves a form of anti-realism about the norms of accuracy and sincerity. On Williams' view, these norms are "of our making", rather than "there anyway." That these norms are there anyway is very much at the heart of the absolutist's credo. Something similar applies to genealogy more generally. Whenever we offer an account of the genealogy of our epistemic concepts and norms, we are presenting them as human constructs, anchored in contingent human needs, subject to the causal influences of biological and cultural evolution. This type of analysis jars with the idea of independent epistemic truths (cf. Street 2009). Admittedly, there is a conceptual difference between anti-realism and B-relativism: while relativism entails anti-realism, the reverse is not true. Nevertheless, anti-realism is incompatible with absolutism. And thus genealogy debunks absolutism. While not being a direct argument *for* B-relativism, it does remove B-relativism's main competitor from the contest.

There is also a second, less direct way in which genealogy weakens absolutism. One common line of argument in defence of absolutism—whether in epistemology or in ethics—is that it is the *default* view: the view that we should adopt unless we are given good reason not to (see Cuneo 2007). Part of the reason why absolutism is taken to be the default view is that, if absolutism is true, we can explain a wide range of common intuitions. For example, we can explain why epistemic norms, such as norms of accuracy and sincerity, seem so attractive. Genealogy decisively blocks this route. Genealogical, symmetrical explanations of absolutist *and relativist* intuitions put all of these intuitions on a par. None of these intuitions have the special rights of the firstborn. None of them are the default. Hence the relativist need not fight the uphill battle of challenging allegedly absolutist common sense. Neither side is common sense alone. Both are—up to a point. Given absolutists' reliance on the burden of proof, the genealogical redistribution of this burden helps B-relativism.

Finally, where does this leave the invariantism, contextualism and M-relativism dispute? For the reasons just given, it is hard to see how a genealogical account of this dispute could be entirely neutral. But which side does it favour? Answering this question requires getting clear on the relationship between M-relativism and B-relativism: to the extent that M-relativism fits naturally with B-relativism, the argument above offers reasons to think that genealogy favours M-relativism. But this is a complicated issue, which we lack the space to address here.

6 Summary

In this paper we have brought together two important strands in contemporary epistemology: the debate around epistemic relativism and the genealogy of knowledge. We have proposed—what we believe to be—an original rendering of the latter, and have applied it to the former. We have identified the general human needs and practices that give rise to the intuitions underlying absolutism, B-relativism, invariantism, contextualism and M-relativism. And we have tried to maintain—albeit in a tentative fashion—that genealogy is a natural ally to B-relativism.[2]

Notes

1. We use italics for concepts throughout.
2. Our work on this paper was assisted by funding from the ERC Advanced Grant Project "The Emergence of Relativism" (Grant No. 339382).

References

Barnes, Barry, and David Bloor. 1982. Relativism, Rationalism and the Sociology of Knowledge. In *Rationality and Relativism*, ed. Martin Hollis and Steven Lukes, 21–47. Oxford: Blackwell.

Bloor, David. 1991. *Knowledge and Social Imagery*. 2nd ed. Chicago: Chicago University Press.

Boghossian, Paul. 2006. *Fear of Knowledge: Against Relativism and Constructivism*. Oxford: Oxford University Press.

Brown, Jessica. 2006. Contextualism and Warranted Assertibility Manoeuvres. *Philosophical Studies* 130 (3): 407–435.

Craig, Edward. 1990. *Knowledge and the State of Nature: An Essay in Conceptual Synthesis*. Oxford: Oxford University Press.

———. 1993. *Was Wir Wissen Können*. Frankfurt am Main: Suhrkamp.

———. 2007. Genealogies and the State of Nature. In *Bernard Williams*, ed. Alan Thomas, 181–200. Cambridge: Cambridge University Press.

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

DeRose, Keith. 2009. *The Case for Contextualism: Knowledge, Skepticism, and Context*. Vol. 1. Oxford: Oxford University Press.

Gerken, Mikkel. 2013. Epistemic Focal Bias. *Australasian Journal of Philosophy* 91 (1): 41–61.

Hawthorne, John. 2004. *Knowledge and Lotteries*. Oxford: Oxford University Press.

Hazlett, Allan. 2014. Expressivism and Convention-Relativism about Epistemic Discourse. In *Naturalizing Epistemic Virtue*, ed. Abrol Fairweather and Owen Flanagan, 223–246. Cambridge: Cambridge University Press.

Henderson, David. 2009. Motivated Contextualism. *Philosophical Studies* 142 (1): 119–131.

Kelp, Christoph. 2011. What's the Point of 'Knowledge' Anyway? *Episteme* 8 (1): 53–66.

Kusch, Martin. 2009. Testimony and the Value of Knowledge. In *Epistemic Value*, ed. Duncan Pritchard, Adrian Haddock, and Alan Millar, 60–94. Oxford: Oxford University Press.

Kusch, Martin, and Robin McKenna. forthcoming. The Genealogical Method in Epistemology. *Synthese*.

MacFarlane, John. 2014. *Assessment Sensitivity: Relative Truth and Its Applications*. Oxford: Oxford University Press.

Nagel, Jennifer. 2008. Knowledge Ascriptions and the Psychological Consequences of Changing Stakes. *Australasian Journal of Philosophy* 86 (2): 279–294.

———. 2010. Knowledge Ascriptions and the Psychological Consequences of Thinking about Error. *Philosophical Quarterly* 60 (239): 286–306.

Reynolds, Steven. 2017. *Knowledge as Acceptable Testimony*. Cambridge: Cambridge University Press.

Rorty, Richard. 1981. *Philosophy and the Mirror of Nature*. Princeton: Princeton University Press.

Rysiew, Patrick. 2001. The Context-Sensitivity of Knowledge Attributions. *Noûs* 35 (4): 477–514.

———. 2012. Epistemic Scorekeeping. In *Knowledge Ascriptions*, ed. Jessica Brown and Mikkel Gerken, 270–293. Oxford: Oxford University Press.

Scheff, Thomas J. 1988. Shame and Conformity: The Deference-Emotion System. *American Sociological Review* 53 (3): 395–406.

Shapin, Steven. 1994. *A Social History of Truth: Civility and Science in Seventeenth-Century England*. Chicago: University of Chicago Press.

Shapin, Steven, and Simon Schaffer. 1985. *Leviathan and the Air-Pump: Hobbes, Boyle and the Experimental Life*. Princeton: Princeton University Press.

Srinivasan, Amia. 2015. The Archimedean Urge. *Philosophical Perspectives* 29 (1): 325–362.

Street, Sharon. 2009. Evolution and the Normativity of Epistemic Reasons. *Canadian Journal of Philosophy* 35 (s1): 213–248.

Unger, Peter. 2014. *Empty Ideas: A Critique of Analytic Philosophy*. Oxford: Oxford University Press.

Williams, Bernard. 2002. *Truth and Truthfulness: An Essay in Genealogy*. Princeton: Princeton University Press.



Reasons Primitivism and Epistemic Expressivism

Teemu Toppinen

What are reasons? Perhaps not much can be said in response by way of giving an analysis. Here's T. M. Scanlon in his *What We Owe to Each Other*:

> I will take the idea of a reason as primitive. Any attempt to explain what it is to be a reason for something seems to me to lead back to the same idea: a consideration that counts in favor of it. 'Counts in favor how?' one might ask. 'By providing a reason for it' seems to be the only answer. (1998, 17)

Derek Parfit advances a similar view in his *On What Matters*:

Earlier versions of this paper were presented at the University of Helsinki and at the 2015 Foundations of Normativity Workshop in Edinburgh. I thank the terrific audiences on these occasions—especially: Debbie Roberts (my commentator in Edinburgh), Matthew Chrisman (who, with Allan Hazlett, also deserves thanks for organising the workshop), Jaakko Hirvelä, Richard Holton, Mike Ridge, and Geoff Sayre-McCord—for their very helpful comments.

T. Toppinen (✉)
Department of Philosophy, University of Helsinki, Helsinki, Finland
e-mail: teemu.toppinen@helsinki.fi

© The Author(s) 2018
C. Kyriacou, R. McKenna (eds.), *Metaepistemology*, Palgrave Innovations in Philosophy, https://doi.org/10.1007/978-3-319-93369-6_11

> It is hard to explain the concept of a reason, or what the phrase 'a reason' means. Facts give us reasons, we might say, when they count in favour of our having some attitude, or our acting in some way. But 'counts in favour of' means roughly 'gives a reason for'. Like some other fundamental concepts, such as those involved in our thoughts about time, consciousness, and possibility, the concept of a reason is indefinable in the sense that it cannot be helpfully explained merely by using words. (2011, 31)

Scanlon and Parfit endorse, then, the following view:

(Reasons primitivism) The concept of a normative reason is unanalysable.

Reasons primitivism, in this sense, goes perhaps smoothly together with certain normative and metaphysical claims. We could strengthen it as follows:

(Reasons primitivism*) The concept of a normative reason is unanalysable, and reasons properties and facts form a basic category of normative properties and facts.[1]
(Reasons primitivism**) The concept of a normative reason is unanalysable, and reasons properties and facts form a basic category of properties and facts.

Scanlon and Parfit accept reasons primitivism in these stronger senses, but someone could accept the conceptual claim, Reasons primitivism, and yet reject the way in which Parfit and Scanlon package it together with non-naturalist (or "quietist") normative metaphysics, as in Reasons primitivism**. According to Parfit and Scanlon, the property of being a reason is not a natural property, and the facts concerning what is a reason for what are not natural facts. That is, roughly, reasons properties and facts are not, on Parfit's and Scanlon's views, properties and facts of the kind that could be studied in empirical sciences. This stronger view is what people standardly mean by "reasons primitivism."[2]

In this paper, I shall be discussing Michael Smith's (2017, 2015) recent, and to my mind interesting, case against the primitivist view.[3] The

argument can be read as attacking the conceptual claim of reasons primitivism, as well as a metaphysical claim. I will mostly focus on a version that is directed at the conceptual claim. I start by explaining Smith's argument (Sect. 1). I then proceed to suggest that reasons primitivists can deal with this argument if they accept a suitable *expressivist* view about judgements concerning normative reasons (Sects. 2 and 3). In the final section of the paper (Sect. 4), I briefly note that Smith's argument could be given also a metaphysical twist.

1 Smith's Argument(s)

Smith presents his case against reasons primitivism in somewhat different ways in different places (Smith 2015, 2017). Here's a lengthy quotation from his paper "Parfit's Mistaken Metaethics":

> But how plausible is the idea that the relation of 'being a reason for' is unanalyzable? Focus on reasons for belief. […] Though the idiom isn't Hume's, we can understand Hume's famous endorsement of *deductive* reasoning, and attack on *inductive* reasoning, as tantamount to the suggestion that we can analyze the reasons relation in the case of belief […]. The analysis, according to Hume, is to be given in terms of, *inter alia*, the concept of *entailment* […]. 'The fact that *p*' provides a reason for believing a proposition *q*, Hume seems to have thought, if and only and because *p entails q*.
>
> Hume's view is, of course, much maligned, but what's maligned isn't the background assumption that we can analyze the concept of a fact's being a reason for belief in other terms. What's maligned is rather his restrictive tying of the concept of a reason for belief to the concept of entailment […]. If we reject this Humean idea, and extend our concept of a reason to include cases where certain facts provide evidence for the truth of other propositions, without entailing them, then we might suppose instead that the analysis of a reason is to be given in terms of some more liberal theory […].
>
> What would such a theory look like? […] Such a theory would presumably draw on the theory of perception, on the one hand, and the theory of

> epistemic rationality, on the other […]. Here is not the place to go into details […]. But the very fact that we have Hume's own reductive theory of reasons for belief in terms of entailment as a model, together with the fact that we can specify the materials in terms of which an alternative to Hume's theory might be given, makes the idea that we should embrace [reasons primitivism], at least in the case of reasons for belief, premature to say the least. We should proceed on the assumption that the concept of a reason for belief can be analyzed.
>
> […] this provides us with excellent reasons to reject [reasons primitivism] more generally. For how could the concept of a reason for a desire […], or an action, be metaphysically fundamental, if the concept of a reason for belief is not? […] The concept of a reason would be a ragbag, like the concept of a-number-or-a-dog. On the plausible assumption that our concept of a reason isn't such a ragbag, the only conclusion to draw, at least *pro tem*, is that [reasons primitivism] is quite generally false. (2017, 100–02)

Someone who finds reasons primitivism initially plausible seems unlikely to be persuaded by what Smith says here. It is true that Hume's analysis is too restrictive, and undoubtedly more relaxed and consequently more plausible reductive analyses could be offered. But it seems like a considerable leap to conclude from this that reasons primitivism isn't very plausible, or that we should proceed on the assumption that it is false.

Chris Heathwood argues, perhaps in a similar spirit, for a reductive account of claims concerning epistemic reasons, reasonableness, and justification. He suggests that *something like* the following might be true (Heathwood 2009, 89–92):

> "*p* is reasonable for S to believe" means the same as "*p* is likely, given S's information."

Attempts to analyse moral concepts in entirely descriptive terms plausibly fail Open Question style tests: "This is good" doesn't seem to mean the same as "This is something we desire to desire," as we may coherently accept "This is something we desire to desire, but this is not good." By contrast—or this is what Heathwood proposes—"This is likely, given my

evidence, but it's not reasonable for me to believe it" *does* have "an air of incoherence about it." (2009, 90) Heathwood emphasises that even if we don't have the intuition that this is self-contradictory, we at least lack the intuition that it's not. This contrasts with the moral case, and might be taken to indicate that the prospects for a reductive analysis of epistemic concepts are much sweeter than those of a reductive account of moral concepts. While we may have doubts about the extensional adequacy of Heathwood's toy proposal, the problem doesn't seem to lie in the seeming impossibility of a reductive analysis. On the assumption that "is reasonable to believe" and "is likely, given the information available" *are* necessarily coextensive, we may accept the analysis of the former in terms of the latter. Or so Heathwood suggests.[4]

We could perhaps understand Smith's argument, above, along the same lines. In another recent paper, he writes:

> Even though Hume's own account of how reasons for belief count in favour of believing in terms of entailment is too restrictive […] his account does appear to be right in all its essentials. Facts that are reasons for belief do seem to count in favour of believing by *entailing*, or if not by entailing then by *providing evidence* for, or by *making probable*, or by *lending weight to*, or by *being part of the best explanation of* the truth of the propositions for which they are reasons. Though it may be controversial which of these is the correct analysis of the reason relation—and which philosophical analyses aren't controversial?—it is not controversial that the relation in question has something to do with truth-conduciveness. (2015, 37)

This seems to be very much in the spirit of Heathwood's suggestion. But we could also read the argument somewhat differently. A reasons primitivist should plausibly admit that, when it comes to reasons for belief, reasonhood necessarily has something to do with truth-conduciveness. But she might suggest that this can be explained without offering a reductive analysis of the concept of a reason.

Let's first consider an unsuccessful suggestion that might seem tempting. This suggestion is that a reason for believing that p necessarily is a consideration that provides evidence, or makes it more likely, that p, but that this is so as a matter of substantive normative fact. Just as a utilitarian

might hold that moral rightness necessarily has something to do with happiness maximisation while also accepting a view on which the concept of rightness is unanalysable, a reasons primitivist might hold that reasons for belief necessarily have something to do with truth-conduciveness while also holding onto her reasons primitivism.

Smith finds this response unsatisfactory. That reasons for a belief bear on the truth of the relevant belief is "non-negotiable", Smith says, and "to say that the facts that are reasons for belief favour belief is to say that they in some way support the truth of the beliefs which they favour." (2015, 38) I take it that Smith's idea, here, is that the link between reasons for belief and truth-conduciveness is not just a substantial normative link, but holds as a matter of conceptual necessity. If that's right, then an appeal to substantive normative facts won't do in the present context. If there were a conceptual connection between moral rightness and happiness maximisation, a moral non-naturalist could not explain this by appealing to a substantive, utilitarian, moral principle. Likewise, if it is a conceptual truth that reasons for belief favour beliefs by being truth-conducive, a mere appeal to substantive normative claims won't allow us to explain why this would be so.

So, it seems that at the heart of Smith's argument is a demand that a certain conceptual truth be explained. We could perhaps formulate the argument as follows:

1. (Strong evidential constraint) The fact that *p* is a reason to believe proposition *q* only in case *p* provides evidence, or makes it more likely, that *q* is true.
2. Premise 1 is a conceptual truth.
3. The only (or best) explanation for the truth of premise 2 is that the concept of a reason for belief is analysable in terms that explain the conceptual connection between reasons for belief and evidence. So:
4. Reasons primitivism is not true of the concept of a reason for belief.
5. The concept of a reason is not "a ragbag, like the concept of a-number-or-a-dog." So:
6. For any *x*, reasons primitivism is not true of the concept of a reason for *x*.

In the next section, I will challenge premise 3. However, before engaging with that issue, I should flag two more assumptions in this argument that I'm happy to grant. The first is that we can provide a naturalistic or descriptive specification of what it is to provide evidence, or to make it more likely, that something is the case.[5]

The second assumption is that the truth of premise 2 requires (some non-trivial) explanation. It's not completely clear that this is so. Some might be tempted to say that while the strong evidential constraint doesn't flow from the analysis of the concept of a reason for belief, it nevertheless is a constraint on being a competent user of this notion that one uses it in accordance with this constraint. To the extent that this seems to be so, a natural reaction to someone who would fail to obey this constraint would be to think that she hasn't adequately grasped the relevant concept. And this, some might be tempted to say, is all that it takes to explain the status of the strong evidential constraint as a conceptual truth.[6]

This does not seem satisfying to me. If a natural reaction, on the part of competent speakers, to someone who seems to be failing in thinking or talking in accordance with the strong evidential constraint is that this someone doesn't fully grasp the concept of a reason for belief, this may give us reason to believe *that* the strong evidential constraint is a conceptual truth. But giving such a reason does not amount to explaining *why* the strong evidential constraint would be a conceptual truth. It would be natural to think that an explanation of the relevant sort would have to flow in some way from an informative account of the concept of a reason for belief. But I shall not press this issue here. Also, the first way of understanding Smith's argument, mentioned above, does not require that a specific conceptual truth be explained. (Finally, we could also give the argument a metaphysical twist—see Sect. 4.)

2 Epistemic Expressivism and Evidential Constraints

It is quite natural to suppose, as Smith does, that an interesting conceptual truth concerning reasons for belief, such as the strong evidential constraint, should be explained with reference to the analysis of the concept

of a reason. Where else could we look? We might hope to explain what needs explaining here by giving an expressivist account of reasons-for-belief thought and talk.[7] I take expressivism to be the following view (see Dreier 2004):

(Expressivism) For any normative belief, *p*, and subject, S, if S believes that *p*, this is so in virtue of S's being in a (broadly) desire-like state, and not in virtue of S's being doxastically related to a normative property.[8]

Expressivism explains the state of mind of having a normative belief without appealing to what this belief is about. So, an expressivist account of beliefs concerning reasons for belief explains the nature of these beliefs without appealing to the idea that these beliefs somehow relate the believer to the property of being a reason for belief. By giving such expressivist account of the relevant beliefs, we might hope to explain certain interesting features of reasons-for-belief thought—for example, the evidential constraint—without relying on an analysis of the concept of a reason.

The idea isn't unfamiliar. For instance, it is widely believed to be a conceptual truth that the normative supervenes on the natural or the descriptive. And it is also pretty widely believed that an expressivist account of normative discourse would allow us to explain why this is so. The idea would be, very roughly, that given the functions of normative thought and talk, the states of mind expressed by normative sentences are such that someone who is in these states of mind is thereby committed to accepting the supervenience thesis. To think that something, *x*, is good or supported by reasons, say, is to see *x* as having some natural or descriptive properties, and to (say) approve of *x* in the light of its having these properties. And so, anything, *y*, which is identical to *x*, as far as its natural or descriptive properties are concerned, must be taken to be identical to *x* also with regard to its normative features.[9] I shall next outline a way of using a broadly similar strategy in explaining why there would be a conceptual evidential constraint on judgements concerning reasons for belief.

According to a familiar expressivist story about judgements concerning what we there would be reason to *do*, or what we ought to do, such judgements consist in desire-like states of mind such as a sentiment of

approval, or a plan to act in certain ways. A natural way of extending this idea to judgements concerning reasons for belief would be to say that judgements concerning what there would be reason to believe consist in some desire-like states of mind such as a plan to form beliefs in certain ways. Something like this is suggested by Allan Gibbard. On his account, to think that the fact that the litmus paper turns red is a reason to believe that the solution is acid is to plan, "for the contingency of seeing red, to believe the solution acid." (2012, 171)[10]

Here's a worry, though. This kind of story cannot explain why we are constrained to treat only considerations that bear on the truth of some proposition as reasons for believing that proposition. For it seems as if we might just as well plan to believe that *q* for the contingency of it being amusing to believe that *q*, while also thinking that the fact that believing that *q* would be amusing would not bear on the truth of *q*. It seems like we could plan to have beliefs for "pragmatic reasons." And so it seems like we could, on Gibbard's account, believe, without any conceptual confusion, that such pragmatic considerations are reasons for believing things. This would contrast with the idea that Smith's strong evidential constraint is a conceptual truth.[11]

Gibbard could answer that a plan, for the contingency of it being amusing to believe that *q*, to believe that *q* is not a plan *to believe* that *q*, but rather a plan *to cause oneself to believe* that *q*. And one could suggest, moreover, that it is not possible to plan to believe that *q* on pragmatic grounds.[12] But why would this be so? If we can plan to believe things on evidential grounds, then why couldn't we plan to believe things on pragmatic grounds—say, on the grounds that believing in certain ways would be fun?

The idea of planning beliefs sounds a bit troubling, anyway. Usually, we talk of planning for things that are under our control in relatively direct ways. I might plan to have some herb infusion on the grounds that "in a sip, balance returns and the day ahead looks a whole lot brighter" (that's what they promise my mixture of organic peppermint, fennel and rose will do to me), and then set off for the philosophy department's coffee room, executing this plan in the light of my beliefs. I don't similarly plan to believe propositions on the grounds that they are true, or on the grounds that believing them would be amusing, and then go on to form

these beliefs in the light of suitable instrumental beliefs. I don't believe at will; believing ain't acting.[13]

Gibbard (2012, 171–72) grants all this, and suggests that plans to believe are to beliefs as plans to act are to willing. We don't believe at will, but neither do we will at will. Plans for action and plans for belief tend to ensue in willing and belief, respectively. But the relation between planning and willing seems different from that between planning and believing. Plans ensue in willing to perform some particular actions with the help of some instrumental beliefs. We do not, however, form beliefs (in part) thanks to believing that we can reach our goals by believing in the relevant ways.

My proposal is that the expressivist should understand beliefs about reasons for belief (or Gibbard's "plans") as *inferential dispositions* of sorts. Philip Pettit (1993, chap. 1) suggests a three-fold division between intentional states—a division between beliefs, desires (broadly construed), and what he calls "inferential habits." Inferential habits, or, as I shall say, inferential dispositions, are desire-like in having a "world-to-mind" direction of fit: a desire (or a plan) that q "aims" or functions to bring it about that q; an inferential disposition "aims" or functions to bring it about that one's beliefs or credences are related in a certain way. However, whereas desires are *belief-channelled*, inferential dispositions are not. If I plan to perform a certain action, this can only lead me to perform the action in question given that I have a suitable belief concerning the ways in which my plan can be realised. By contrast, if I'm disposed to infer that q from it being the case that p and that if p then q, this can lead me to believe that q without any belief to the effect that by believing that q I can reach my inferential goal (see Pettit 1993, 19).

So, perhaps the expressivist should say something *very roughly* along the following lines:

(Reasons for intention/action) A belief that the fact that p is a reason to intend to (or to perform) φ consists in a desire-like state (e.g., a desire or a plan).

(Reasons for belief) A belief that the fact that p is a reason to believe that q consists in an inferential disposition to raise one's credence in q given p.

These theses would need a great deal of refinement—something I'm not offering here.[14] Not just any old belief-forming disposition can constitute a belief concerning reasons for belief, though. Plausibly, we must be able to make a three-fold distinction between (1) merely conforming to certain regularities or patterns of belief-formation, (2) inferential dispositions that involve being in some suitable sense guided by some rule (e.g., *modus ponens*, or some inductive rule), and (3) inferential dispositions that involve being guided in one's belief-formation by thoughts concerning what there is reason to believe.

So, as a first stab (or better: swing), I suggest that beliefs concerning reasons for belief consist in inferential dispositions which involve further dispositions to go "Oops!" and to revise one's beliefs upon failures to form beliefs in accordance with the relevant patterns (cf. Gibbard 2012, 173), and are perhaps also accompanied by dispositions to accept certain normative sentences. So, an inferential disposition to move from a belief that *p* to a belief that *q* only amounts to thinking that *p* is a reason to believe that *q* when this disposition is coupled with a disposition to have a suitable negative response to a failure to believe that *q* upon believing that *p*, a disposition eventually to come to believe that *q* upon such failures, and perhaps also a disposition to assent to certain normative sentences. Clearly this will require much clarification and elaboration. Also, even if it is, on this account, easy to distinguish between mere conformity to rules and inferential dispositions, drawing the line between inference that doesn't involve normative beliefs and inference that does may turn out to be tricky. I don't know if that's very bad, though. That line does seem potentially quite fuzzy. Anyway, the idea would be that as we add to the complexity of the relevant dispositions—building upon a disposition to conform to a certain rule—we at some point arrive at an inferential disposition that also constitutes a normative belief.

The thought, now, is that if beliefs about reasons for belief consist in inferential dispositions of roughly this kind, then perhaps this allows us to explain why there is an evidential constraint on such beliefs. Why would this be so? We may first note that any inferential disposition is a disposition to move from one mental state to another, where being in the first makes it *intelligible* to be (or "rationalizes" being) in the other, and the causal link between the two states is of the "right kind" (in order to

rule out deviant causal chains that don't count as inference). I'm assuming here that intelligibility and what it is for a causal chain to be of the right kind can be specified in naturalistic or descriptive terms. Someone who manifests inferential dispositions to move from one set of beliefs and desire-like states to another necessarily exhibits, by and large, certain kinds of patterns among her mental states. However, these patterns can be specified in naturalistic or descriptive terms. Or so I shall assume here. (We are now operating on the assumption that belief, which plausibly is essentially and conceptually linked to certain kinds of inferential dispositions, can be understood non-normatively.)[15]

So, inferential dispositions must by and large exhibit certain intelligible patterns. Now, let's consider the kinds of inferential dispositions that I have suggested expressivists might say beliefs concerning reasons for belief consist in. For example, let's consider an inferential disposition to raise one's credence in *q* upon believing that *p*, where this disposition also involves a further disposition to go "Oops!" in case one fails to raise one's credence in *q* upon believing that *p*, and, under such circumstances, to reconsider the matter as to whether *q*, where this tends to ensue in finally readjusting one's credences in a suitable way. Call such inferential dispositions "inferential dispositions with property F" or "F-dispositions." And let us say that an inferential disposition is "based" on (or "sensitive" to) a certain consideration when a belief in this consideration triggers the disposition in question. It seems quite plausible that in order for us to be able to interpret someone as a relatively complex believer, or as a believer with F-dispositions, we must take her inferential dispositions of this kind to be largely sensitive to, or based on, what she takes to bear on the truth of her beliefs. Is it possible for us to understand someone as being disposed, in the relevant sort of way, to believe that *q* as a result of her believing that *p*, and not see the truth of *p* as being in any way relevant to whether *q* is true? That is, is it possible for us to understand someone as being disposed to believe that *q* upon believing that *p*, to go "Oops!" in case one fails to do this, to readjust one's credence in *q* upon such failure, and yet not take the truth of *p* to bear on the truth of *q*?

Perhaps this is possible. But let us first suppose it isn't. That is to say, let us suppose, for a moment, that we are constrained to interpret any believer with F-dispositions in such a way that her F-dispositions to

believe that *q* are based only on considerations that she takes to bear on the truth of *q*. Let us now, moreover, suppose that the expressivist account of beliefs concerning reasons for belief, suggested above, is correct. If this were so, this would explain why any believer having inferential dispositions with property F would be committed to thinking that only considerations that bear on the truth of *q* can be reasons for believing that *q*.

Would this amount to explaining why Smith's strong evidential constraint is a *conceptual* truth? So it seems. At least, on this view, anyone who formed beliefs concerning what there is reason to believe would thereby be committed to forming these beliefs in accordance with the relevant constraint. The explanation would also be wholly compatible with reasons primitivism. The concept of a reason could be taken to be primitive. It would just be that anyone who could be taken to be applying this primitive concept would be bound to apply it in accordance with the strong evidential constraint. So, given the preceding (big) assumptions—that is, given that F-dispositions must be sensitive exclusively to evidential considerations, and that a suitable form of expressivism is true—Smith's challenge to reasons primitivism can be dealt with.

Perhaps our first assumption—the assumption that any believer's inferential dispositions with property F are bound to be based on evidential considerations—wasn't credible, though. (Naturally, the expressivist assumption is extremely controversial, too, but we're sticking with it in the present context). Perhaps the constraints on interpreting someone as a believer are not nearly as strict as this first assumption would have it. It is not obvious that we cannot imagine someone whose F-dispositions to believe that *q* are based on considerations that do not seem to her to bear on the truth of *q*. How should we understand the relation between a believer's inferential dispositions with property F and sensitivity to evidential considerations? Here are some options:[16]

(Evidential only) Any believer's set of inferential dispositions (with property F) to believe that *q* contains only inferential dispositions that are based on considerations that seem to the believer to bear on the truth of *q*.

(Non-evidential) Any believer's set of inferential dispositions (with property F) to believe that *q* may contain only inferential dispositions

that are based on considerations that do not seem to her to bear on the truth of *q* (e.g., considerations having to do with the benefits of believing that *q*).

(Evidential mostly) Any believer's set of inferential dispositions (with property F) to believe that *q* contains mostly inferential dispositions that are based on considerations that seem to the believer to bear on the truth of *q*, but may also contain some inferential dispositions that are based on considerations that do not seem to her to bear on the truth of *q*.

(Evidential mostly*) Any believer's set of inferential dispositions (with the property F) to believe that *q* contains mostly inferential dispositions that are based on considerations that seem to the believer to bear on the truth of *q*, but may also contain some inferential dispositions that are based on considerations that do not seem to her bear on the truth of *q*, if the inferential dispositions that are based on considerations that seem to the believer to bear on the truth of *q* *suffice* to incline her to believe that *q* rather than not-*q*.

As noted above, Evidential only may seem too restrictive. Let us imagine, for instance, a great reward being offered for believing that *q*. Might it not be conceivable that there would be a creature who, upon believing that a great reward is associated with believing that *q*, would thereby be disposed to believe that *q*, as well as to react negatively, reconsider, and revise her belief as to whether *q* under circumstances in which she doesn't do so? Perhaps this is conceivable. However, Non-evidential seems implausibly permissive. Might it be conceivable that some believer's inferential dispositions with property F were predominantly like this—that is, not based on what the believer takes to bear on the truth of her relevant beliefs? This, too, seems implausible. So, perhaps something like Evidential mostly or Evidential mostly* is on the right track. There seems to be at least a "necessarily mostly" relation between any believer's inferential dispositions with property F and the sensitivity of such inferential dispositions to evidential considerations.

Let us suppose that this is right. And now, let us combine this thought with the expressivist account of beliefs concerning reasons for belief, suggested above. This combination of ideas would not allow us to explain

why Smith's strong evidential constraint is a conceptual truth. But it would allow us to explain a weaker constraint such as the following:

(Weak evidential constraint) Most, or at least a sizeable proportion, of the reasons for believing things provide evidence, or make it more likely, that these beliefs are true.

For suppose that having an inferential disposition, with property F, to raise one's credence in *q*, given *p*, amounts to thinking that the fact that *p* is a reason to believe that *q*. And suppose, moreover, that, as we have assumed, any believer's inferential dispositions with property F must be predominantly based on what the believer in question takes to be evidence, or to make it more likely, that *q* is true. It would then follow that any believer with F-dispositions would be committed to accepting the weak evidential constraint. The concept of a reason for belief remains (potentially) unanalysable; it's just that any thinker who applies this primitive concept necessarily applies it in accordance with the weak constraint.

Securing the status of the weak evidential constraint as something that any believer is committed to accepting would not be insignificant. Smith's strong evidential constraint is pretty strong, after all. If it seems to us that a believer's set of inferential dispositions may contain some dispositions that are based on non-evidential considerations, it also might seem quite plausible that someone could coherently reject the strong constraint. We could still hold on to the idea that Smith's strong constraint is *true*. The thought would just be that someone who, perhaps mistakenly, rejects this constraint does not thereby manifest any conceptual failure. Perhaps allowing for this possibility would even be a virtue in an account of beliefs concerning reasons for belief. There are, after all, very smart people who have thought about reasons for belief long and hard and yet come to reject strong evidential constraints.[17] It might be considered a plus if these people turned out to be guilty of a substantial mistake, at most, rather than a conceptual one.

Also, it seems that a Smith-style argument could be run also with a weaker constraint. We can replace premise 1 with the following, and the rest of the argument still makes perfect sense:

1*. (Weak evidential constraint) Most, or at least a sizeable proportion, of the reasons for believing things provide evidence, or make it more likely, that these beliefs are true.

However, given the availability of a suitable expressivist account, we may again reject premise 3. And, as noted above, the result is again secured in a way friendly to reasons primitivism.

This kind of expressivist view would also allow us to respond to Smith's argument understood as similar to the one presented by Heathwood (see Sect. 1). This line of argument relies, again, on the idea that our conceptual intuitions fail to classify sentences such as "This is likely, given my evidence, but it's not reasonable for me to believe it" as coherent. But this wouldn't be surprising, given that complex believers such as us would be bound to have certain kinds of inferential dispositions, and so to give considerable weight to evidential considerations in their thinking about reasons for belief.

3 A Worry

According to my expressivist proposal, concepts of a reason for intention and a reason for belief are explained in terms of different types of states of mind. Does not this make the account of the concept of a reason unacceptably rag-bag?

I don't think it does. On the proposed view, beliefs concerning reasons for intentions consist in desire-like states such as desires or plans, whereas beliefs concerning reasons for belief consist in certain kinds of inferential dispositions. So, the states with reference to which the concepts of a reason for intention and a reason for belief are explained really are of different types. However, desire-like states and inferential dispositions are similar in that they both have a "world-to-mind" direction of fit: a desire (or a plan) that *q* "aims" to bring it about that *q*; an inferential disposition "aims" to bring it about that one's credences are related in a certain way (Pettit 1993, chap. 1). As both judgements concerning reasons for intentions and judgements concerning reasons for belief consist in states that

function to guide attitude-formation, it makes sense that they both can be said to concern reasons.

This explanation for the unity of reason concepts is not available to those who reject expressivism and would rather like to understand reasons-beliefs as representations of reasonhood. A defender of such "cognitivist" account of reasons-beliefs, too, could hold that all reasons-beliefs essentially function to guide attitude-formation. However, on a cognitivist account, this attitude-guiding function of a reasons-belief would be explained by what the belief is about, and not the other way around. So, the cognitivist cannot explain the unity of reason concepts by appealing to their attitude-guiding function, but must rather appeal to what the beliefs are about. On the cognitivist view, the unity must spring from the nature of the favouring relation itself.

It may still seem like the meaning of "a reason for" isn't, on this account, strictly speaking the same in "a reason for intention" as it is in "a reason for belief." An expressivist explains the meanings of sentences in terms of the types of states of mind that they express, and, on the current proposal, sentences concerning reasons for intentions and reasons for beliefs do express different types of states of mind, respectively.[18]

Stephen Kearns and Daniel Star (2009, 220–21) present some arguments for the idea that reasons for intentions and reasons for belief "are of a kind," which might be taken to bear on rejecting the strict unity of the concept of a normative reason. First, they point out that we use the same word "reason" for both sorts of considerations. I have already suggested that this would be understandable, given that both kinds of reasons-beliefs consist in states the function of which is to guide our attitude-formation.

Second, they note that the word "reason" behaves similarly both in connection with reasons for intentions (or actions) and reasons for beliefs in that we can talk of both reasons *to* intend (or act) and reasons *to* believe. I take it that this, too, would be unsurprising given the similarity between the states of mind at issue.

Third, Kearns and Star suggest that sentences such as "That the ground is wet is a reason to believe it is raining and to intend to take an umbrella" seem to apply the very same reason concept to both intentions and beliefs. However, nothing of significance seems to be lost even if we admitted

that, strictly speaking, what this sentence says is that the fact that the ground is wet is a reason to believe it is raining and a reason* to intend to take an umbrella. Given the similarity between the functions of the states expressed by claims concerning reasons and reasons*, and given that we can talk about both reasons to believe and reasons* to intend, it would not seem too surprising that sentences such as "That the ground is wet is a reason to believe it is raining and to intend to take an umbrella" would come out fine.

Finally, Kearns and Star rely on the thought "that we can weigh up reasons to act and epistemic reasons to believe against each other," and that this "would not be possible if they were not of a kind." (2009, 220) Perhaps, then, the possibility of weighing reasons for action against reasons for belief shows that a single concept of a reason operates on both actions and beliefs. Here is one of Kearns and Star's examples:

> […] imagine a professional high-jumper who must jump higher and higher every time he jumps in a particular prestigious competition if he is to succeed in winning a medal. The high-jumper is aware that he could think carefully about the height of the jump on each occasion he jumps and thus come to form a well-justified belief about whether or not he is going to succeed on each particular occasion. However, he is also aware that paying attention to this kind of reason for belief may be something that will guarantee that he fails to successfully make certain jumps. He has a reason to ignore the evidence concerning whether or not the high-jump is too high for him, given his track-record, and this is the reason he has to win the competition. A reason for action trumps a reason for belief, over and over again. (2009, 220–21)

I do not find this to be a persuasive example of a case where reasons for action would outweigh reasons for belief. The man has a reason to make certain jumps, and this reason grounds a reason for him to avoid paying attention to whether he is likely to make them. But these reasons do not seem to weigh against whatever reasons there are for him *to believe* that he will or will not make the jumps. They only count against his *paying attention* to those reasons. In fact, there seems to be no reason for the high-jumper to pay attention to what reasons there are for him to form the

relevant beliefs. His practical reasons all weigh against doing so. His epistemic reasons, on the other hand, may count in favour of his believing that he will make the jump or in favour of his believing that he will not. Plausibly, these reasons cannot be weighed against the man's reasons for action.

4 Going Metaphysical?

Could an interesting Smith-style argument be devised, which relied on the idea that the evidential constraint holds as a matter of metaphysical necessity? Smith himself notes, on one occasion of formulating his argument, that he "will not fuss about the differences between real definition, conceptual analysis, and metaphysical reduction." (2017, 100) Consider the following:

i. (Strong evidential constraint) The fact that *p* is a reason to believe proposition *q* only in case *p* provides evidence, or makes it more likely, that *q* is true.
ii. Premise 1 holds as a matter of metaphysical necessity.
iii. The only explanation for the truth of premise 2 is that the property of being a reason for belief is analysable in terms that explain the metaphysically necessary connection between reasons for belief and evidence. So:
iv. Reasons primitivism is not true of the property of being a reason for belief.
v. The property of being a reason is not "a ragbag, like the property of being a-number-or-a-dog." So:
vi. For any *x*, reasons primitivism is not true of the property of being a reason for *x*.

This argument would not work against reasons primitivism understood as a view about the concept of a reason, but it might have bite against Reasons primitivism** (see Sect. 1).

I do not have the space to consider the plausibility of this argument here. I will just note that understood this way, the argument would seem

to rely on an instance of the interesting challenge—for the non-naturalist—to explain the necessary connections between the natural and the normative (see, e.g., McPherson 2012). I have elsewhere defended the idea that it will be helpful, in the face of this challenge, for the non-naturalist to "go quasi"—that is, to adopt an expressivist, quasi-realist position, which will allow us to see the relevant necessity claims as first-order normative claims that can then be explained (or left unexplained) as such (see Toppinen 2018).

Notes

1. This formulation is inspired by that offered in Kauppinen (2015).
2. The one-star formulation of reasons primitivism allows for the possibility of giving a metaphysically reductive account of normative properties and facts in terms of natural properties and facts (see, e.g., Schroeder 2007). None of the formulations given above imply the claim that all other normative (or evaluative) concepts, properties and facts would be analysable in terms of reasons.
3. Worries in the same spirit are aired in passing also in Kearns and Star (2009). For other recent criticisms of reasons primitivism or developments of alternatives to it, see Broome (2013, chap. 4); Finlay (2014, chap. 3); Howard (2019); McHugh and Way (2016); Thomson (2008, chaps. 8–9); Väyrynen (2011).
4. For criticism, see Cuneo and Kyriacou (forthcoming); for Heathwood's response, see his chapter in this volume.
5. For discussion, see Heathwood (2009) and Cuneo and Kyriacou (forthcoming).
6. This kind of line is taken, for example, in Dreier (1992, 21), and Olson (2014), with regard to the issue of explaining why the supervenience of the normative on the natural (or the descriptive) would be a conceptual truth.
7. Perhaps there is yet another place to look. We could perhaps suggest that while the concept of a *reason* is to be taken as primitive, the evidential constraint is built into the concept of a *belief* in that it is true, as a matter of conceptual fact, that we ought to believe in accordance with our evidence, and thus only take something, p, to be a reason to believe something, q, if p is evidence, or makes it more likely, that q (on normativity of belief, see, e.g., McHugh and Whiting 2014).

8. I mean to understand "(broadly) desire-like" broadly enough in order to encompass desire-plus-belief states, and higher-order states of being in some such complex states, such that behave in broadly desire-like ways (see Schroeder 2013; Toppinen 2013; Ridge 2014).
9. See, e.g., Blackburn (1993); Gibbard (2003, chap. 5); Hare (1952, chap. 10).
10. See Gibbard (1990, chap. 3, 2003, chap. 11, 2012, chap. 8).
11. Nishi Shah's (2003) criticism of Gibbard seems to be partly based on this point.
12. This seems to be how Shah (2003) suggests Gibbard could respond to this particular worry.
13. See also Kyriacou (2012).
14. Just to mention one fairly big issue, expressivists would need to have something to offer in relation to Frege-Geach style worries. Perhaps they should go hybrid/relational (cf. Schroeder 2013; Toppinen 2013; Ridge 2014). Now, this is rough, and *extremely preliminary*, but perhaps to think that *p* is a reason to believe that *q* is, roughly, to (a) have inferential dispositions to raise one's credences on certain sorts of grounds (or to have certain ways of forming beliefs), and (b) to believe that raising one's credence in *q* on the grounds of believing that *p* would fit the bill (amount to forming a belief in this way). To believe that there's no reason to believe that *q* could then be a matter of having these inferential dispositions (ways of forming beliefs) and of believing that such dispositions don't manifest in raising one's credence in *q* (that a belief that *q* cannot be formed in these ways). For more sophisticated expressivist treatments of judgements concerning epistemic justification or knowledge, see Chrisman (2007), Field (2009), Kappel (2010), Kyriacou (2012) and Ridge (2007). For discussion of how expressivists should understand reasons for action thought, see Ridge (2014, chap. 4) and Sinclair (2016).
15. For the idea that the minimal rationality constraints on interpretation can be understood naturalistically or descriptively, see, e.g., Glüer and Wikforss (2009). For discussion of deviant causal chains, see, e.g., Mele (2003) (which defends a causalist account of action) and Wedgwood (2006).
16. This selection of options has been inspired in part by McHugh (2013, 2015).
17. For defences of "pragmatic" reasons for belief, see, e.g. Howard (2019), Leary (2017) and Reisner (2008).

18. The talk of explaining meaning in terms of states of mind is not very precise. If expressivism is advanced as a *metasemantic* thesis (see, e.g., Chrisman 2016 and Ridge 2014), it is perhaps possible to have a unified account of reasons-talk and thought at some level despite the fact that the explanation for why this unified account works the way it does would be somewhat disunified.

References

Blackburn, Simon. 1993. Moral Realism. In *Essays in Quasi-Realism*, ed. Simon Blackburn, 111–129. Oxford: Oxford University Press.

Broome, John. 2013. *Rationality Through Reasoning*. Chichester: Wiley-Blackwell.

Chrisman, Matthew. 2007. From Epistemic Contextualism to Epistemic Expressivism. *Philosophical Studies* 135 (2): 225–254.

———. 2016. *The Meaning of "Ought": Beyond Descriptivism and Expressivism in Metaethics*. Oxford: Oxford University Press.

Cuneo, Terence, and Christos Kyriacou. forthcoming. Defending the Moral/Epistemic Parity. In *Metaepistemology*, ed. Conor McHugh, Jonathan Way, and Daniel Whiting. Oxford: Oxford University Press.

Dreier, James. 1992. The Supervenience Argument Against Moral Realism. *Southern Journal of Philosophy* 30 (3): 13–38.

———. 2004. Metaethics and the Problem of Creeping Minimalism. *Philosophical Perspectives* 18 (1): 23–44.

Field, Hartry. 2009. Epistemology Without Metaphysics. *Philosophical Studies* 143 (2): 249–290.

Finlay, Stephen. 2014. *Confusion of Tongues: A Theory of Normative Language*. Oxford: Oxford University Press.

Gibbard, Allan. 1990. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Cambridge, MA: Harvard University Press.

———. 2003. *Thinking How to Live*. Cambridge, MA: Harvard University Press.

———. 2012. *Meaning and Normativity*. Oxford: Oxford University Press.

Glüer, Kathrin, and Åsa Wikforss. 2009. Against Content Normativity. *Mind* 118 (469): 31–70.

Hare, R.M. 1952. *The Language of Morals*. Oxford: Oxford University Press.

Heathwood, Chris. 2009. Moral and Epistemic Open-Question Arguments. *Philosophical Books* 50 (2): 83–98.

Howard, Christopher. 2019. The Fundamentality of Fit. In *Oxford Studies in Metaethics Vol. 14*, ed. Russ Shafer-Landau. Oxford: Oxford University Press.

Kappel, Klemens. 2010. Expressivism About Knowledge and the Value of Knowledge. *Acta Analytica* 25 (2): 175–194.

Kauppinen, Antti. 2015. Favoring. *Philosophical Studies* 172 (7): 1953–1971.

Kearns, Stephen, and Daniel Star. 2009. Reasons as Evidence. In *Oxford Studies in Metaethics Vol. 4*, ed. Russ Shafer-Landau, 215–242. Oxford: Oxford University Press.

Kyriacou, Christos. 2012. Habits-Expressivism About Epistemic Justification. *Philosophical Papers* 41 (2): 209–237.

Leary, Stephanie. 2017. In Defense of Practical Reasons for Belief. *Australasian Journal of Philosophy* 95 (3): 529–542.

McHugh, Conor. 2013. Normativism and Doxastic Deliberation. *Analytic Philosophy* 54 (4): 447–465.

———. 2015. The Illusion of Exclusivity. *European Journal of Philosophy* 23 (4): 1117–1136.

McHugh, Conor, and Jonathan Way. 2016. Fittingness First. *Ethics* 126 (3): 575–606.

McHugh, Conor, and Daniel Whiting. 2014. The Normativity of Belief. *Analysis* 74 (4): 698–713.

McPherson, Tristram. 2012. Ethical Non-Naturalism and the Metaphysics of Supervenience. In *Oxford Studies in Metaethics Vol. 7*, ed. Russ Shafer-Landau, 205–234. Oxford: Oxford University Press.

Mele, Alfred R. 2003. *Motivation and Agency*. New York: Oxford University Press.

Olson, Jonas. 2014. *Moral Error Theory: History, Critique, Defence*. Oxford: Oxford University Press.

Parfit, Derek. 2011. *On What Matters*. Oxford: Oxford University Press.

Pettit, Philip. 1993. *The Common Mind: An Essay on Psychology, Society and Politics*. New York: Oxford University Press.

Reisner, Andrew. 2008. Weighing Pragmatic and Evidential Reasons for Belief. *Philosophical Studies* 138 (1): 17–27.

Ridge, Michael. 2007. Epistemology for Ecumenical Expressivists. *Proceedings of the Aristotelian Society, Supplementary Volumes* 81: 83–108.

———. 2014. *Impassioned Belief*. Oxford: Oxford University Press.

Scanlon, Thomas M. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.

Schroeder, Mark. 2007. *Slaves of the Passions*. Oxford: Oxford University Press.
———. 2013. Tempered Expressivism. In *Oxford Studies in Metaethics Vol. 8*, ed. Russ Shafer-Landau, 283–314. Oxford: Oxford University Press.
Shah, Nishi. 2003. How Truth Governs Belief. *Philosophical Review* 112 (4): 447–482.
Sinclair, Neil. 2016. Reasons, Inescapability and Persuasion. *Philosophical Studies* 173 (10): 2823–2844.
Smith, Michael. 2015. Evaluative Judgements, Judgements about Reasons, and Motivations In *Motivational Internalism*, edited by Gunnar Björnsson, Caj Strandberg, Ragnar Francén Olinder, John Eriksson, and Frederik Björklund, 27–43. New York: Oxford University Press.
———. 2017. Parfit's Mistaken Metaethics. In *Does Anything Really Matter? Essays on Parfit and Objectivity*, ed. Peter Singer, 99–120. Oxford: Oxford University Press.
Thomson, Judith Jarvis. 2008. *Normativity*. Chicago: Open Court.
Toppinen, Teemu. 2013. Believing in Expressivism. In *Oxford Studies in Metaethics Vol. 8*, ed. Russ Shafer-Landau, 252–282. Oxford: Oxford University Press.
———. 2018. Non-Naturalism Gone Quasi. In *Oxford Studies in Metaethics Vol. 13*, ed. Russ Shafer-Landau. Oxford: Oxford University Press.
Väyrynen, Pekka. 2011. A Wrong Turn to Reasons? In *New Waves in Metaethics*, ed. Michael Brady, 185–207. Basingstoke: Palgrave Macmillan.
Wedgwood, Ralph. 2006. The Normative Force of Reasoning. *Noûs* 40 (4): 660–686.

Part III

Beyond the Realism/Anti-realism Divide

What Anti-realism About Hinges Could Possibly Be

Annalisa Coliva

1 Introduction

In a number of writings, I have been developing a form of hinge epistemology inspired by Wittgenstein's remarks in *On Certainty* (1969). The key take-home message is that justification and knowledge—to the extent that knowledge is thought of as, roughly, justified true belief—always take place within a system of assumptions. For instance, for Wittgenstein, it is only by taking for granted that the Earth has existed for a long time that we can consider fossils as bearing on the determination of the *specific* age of the Earth. If we doubted that the Earth had existed for a long time, we could no longer consider fossils as being evidentially significant for that specific inquiry, since we could no longer take for granted that their formation took place on the Earth throughout a long amount of time. Their presence would then be compatible with alternative hypotheses such that the Earth was created just a little while ago, replete with all the

A. Coliva (✉)
Department of Philosophy, University of California, Irvine, Irvine, CA, USA
e-mail: a.coliva@uci.edu

© The Author(s) 2018
C. Kyriacou, R. McKenna (eds.), *Metaepistemology*, Palgrave Innovations in Philosophy, https://doi.org/10.1007/978-3-319-93369-6_12

fossils we have found on it, which were created on some other planet or in some sci-fi lab, for instance. Wittgenstein called "hinges" propositions like "The Earth has existed for a very long time," and thought of them as conditions of evidential significance—that is, as what needs to stay put for the justification of specific empirical propositions to be possible at all. He also had a very wide conception of hinges. In his view, they ranged from very general ones like "The Earth has existed for a very long time," to more specific ones like "Nobody has ever been on the Moon"; from impersonal ones, like the previous ones, to personal ones, like "This is my hand" and "My name is AC"; and also from universal ones, like "Cats don't grow on trees," to only contextual ones, like "Here is my hand." Furthermore, he thought of them as playing a rule-like role not just with respect to the conditions of evidential significance but also of meaningful discourse, including meaningful doubt. Thus, for instance, if one really doubted that there is a hand here (while holding up one's hand in front of one's nose), then it would no longer be clear what one would mean by "hand." Hinges thus became Wittgenstein's ammunition in stock against various forms of scepticism. On the one hand, by relying on their meaning-constitutive role, they were used against Cartesian scepticism to claim—well before Putnam's brains in a vat—that if we were dreaming we would also be dreaming that our words have a meaning. So we could not meaningfully entertain the hypothesis that we were dreaming (or that we were brains in a vat, or that we were living in the Matrix, or that we were being deceived by an evil demon) after all. On the other hand, by relying on their role of conditions of evidential significance, they were used against forms of Humean scepticism, to claim that for doubts to get any purchase at all, hinges had to stay put. In particular, that real, empirical and specific doubts could only be raised by taking for granted the existence of physical objects, and the broad well-functioning of our perceptual and cognitive faculties. Thus, to conclude that "There is an external world," "My senses are broadly well-functioning" or "My cognitive faculties are broadly reliable" are not rationally held because they cannot be empirically justified would depend on failing to realise that these propositions themselves contribute to the very possibility of there being empirically justified beliefs and doubts at all. To doubt of them would thus be self-defeating, as it would make doubt ultimately impossible.[1]

In my systematic—as opposed to historical—work on hinge epistemology, I have proposed a different view of hinges, while retaining the gist of what I consider Wittgenstein's main insight in *On Certainty*. First, I have restricted my attention to what Dretske used to call "heavy-weight assumptions." That is, very general propositions like "There is an external world," "Our sense organs are broadly well-functioning," "Our cognitive faculties are broadly reliable."[2] Thus, I have restricted the scope of hinges to those that are constitutive of basic epistemic practices or methods of inquiry, broadly conceived. For one could not even start investigating a mind-independent physical reality if one did not take for granted that there are physical objects, that one's sense organs by and large correctly represent them, and that one's cognitive faculties allow one to form generally correct beliefs about them. Moreover, I have not made any play with the idea that these heavy-weight assumptions could be less than propositional in nature, like rules are generally thought of as being.[3] Finally, I have not considered hinges as meaning-constitutive rules, mainly because it does not seem impossible to think of cases in which one has acquired the meaning of the relevant words by being in contact with their referents, thus being able to ostensively define their meanings and make correct applications of those terms, and then be placed, unbeknownst to one, in a sceptical scenario radically disconnected from the external world, while still retaining—at least for the time being—the initial meanings and concepts. In such a scenario, asking oneself whether one is dreaming, or whether one is brain-envatted, or whether one is living in the Matrix, et cetera, would not seem immediately to entail a violation of the rules of semantic significance. Thus, however short-lived that radical sceptical doubt could be, it would not be altogether unintelligible. Furthermore, it is advisable not to make one's anti-sceptical strategy hostage to the success of this or that theory of meaning. Hence, I haven't pursued any semantic anti-sceptical strategy centered on the role of hinges as meaning-constitutive rules.

Retaining the propositionality of hinges and making them the content of a propositional attitude of acceptance brings with it several problems, however.[4] One might think that it would re-open the door to Cartesian scepticism, since it would no longer allow one to claim that imagining radically sceptical scenarios would ultimately be self-defeating as this

would entail a violation of the conditions for meaningful discourse. This, however, is not the main problem, in my view. For one could still make other moves to counter that form of scepticism. In particular, by denying the unconditional validity of Closure, one could defend the idea that we know that there is a hand here, even if we do not know that there is an external world, or that we are not dreaming, or that we are not brains in a vat, et cetera. One could then defend that position from the charge of condoning abominable conjunctions by noticing that,[5] though not an instance of knowledge or of justified belief, there being an external world, not dreaming, or not being a brain in a vat are rational assumptions we are entitled to make insofar as they are constitutive of epistemic rationality. Of course, one may object to the extended rationality view I proposed, according to which epistemic rationality extends beyond those empirical beliefs that are epistemically justified to the very conditions of possibility of that kind of justification. This is not the place to rehearse possible criticisms and to respond to them.[6] Rather, in my view, the more serious issue, and a very intricate philosophical question indeed, is to understand the metaphysical status of hinges. One way of presenting that challenge is by phrasing it semantically, as hinging—pun intended—on the kind of truth hinges are supposed to enjoy. That is to say, when we say "'There is an external world' is true," what kind of truth are we attributing to the proposition expressed by the sentence "There is an external world"? Is it robustly correspondentist, or evidence-dependent, or, finally, minimalist?[7] This is the topic of the present paper. As is apparent, it is a theme at the interface between epistemology and metaphysics, with the potential of re-shaping reflection in each of these areas—particularly in epistemology. For commitment to realism seems to be a tenet of the discipline. That is, it is usually taken for granted that there are physical objects—understood, in a Humean fashion, as objects existing continuously even when not perceived, and as not dependent for their existence on anyone's minds. What epistemology then pursues is either an explanation of how we can form justified or even knowledgeable beliefs about such mind-independent entities, once we start out with the Quinean "meager input" of experience (Quine 2008, 297); or else, an understanding of how we may claim that we are within our epistemic rights in having those beliefs, *vis-à-vis* various kinds of sceptical paradoxes. The former

is the project pursued by naturalised epistemologies, the latter the one pursued by normative ones. Quine too, however, was well aware that the two projects were not in opposition to one another, but rather complementary. For he recognised that what is the tenet of naturalized epistemology is in fact what normative epistemology tries rationally to redeem.

As a methodological point, in the following I will be exploring the issue of the metaphysical status of hinges not only by performing a semantic ascent, but also by bringing in considerations that fall within the normative project in epistemology. For what is worth, let me just record here, with no space to argue for it in any detail, that, in a spirit of division of cognitive labour, I consider the naturalised project more of a project for cognitive sciences nowadays, rather than for philosophy. Thus, it will be left aside in the following.

2 Realism About Hinges: A Correspondentist Conception of Hinges' Truth?

The traditional realist view has it that a sentence S, expressing a proposition *p*, is true iff it is the case that *p* (or it is a fact that *p*). Thus, S is true iff *p* is a state of affairs that obtains and whose obtaining is prior and independent of its being the object of a representation (linguistic or otherwise). Accordingly, the English sentence "There is an external world" would express the proposition that there is an external world and it would be true iff there were such an external world. That is, if its existence were prior and independent of our thinking or talking about it. It is because there is such a mind-independent world, then, that if we say (or think) "There is an external world" what we say (or think) is true. What we would thereby say (or think) would be true *because* it would correspond to how things are in the world, independently of us. The world is what it is, it is structured the way it is, and when we talk (or think) about it, what we say (or think) is true just in case it mirrors or reflects that outer entity. In the usual terms of the *Euthyphro* contrast (cf. Wright 1992, 108–39), for a realist it is *because* there is an external world that, when we say (or

think) there is, what we say (or think) is true; not the other way around. That is, it is not *because* we hold it true that there is an external world that it is the case that there is.

Now, as remarked at the end of the previous section, in epistemology the realist side of the *Euthypro* contrast is usually taken for granted. The question then is how we can ascertain or claim that we are within our epistemic rights in taking it for granted that there is an external world. The problem of philosophical scepticism, in my view, just dramatises this issue. If it takes the form of a Cartesian paradox, it challenges us by claiming that if we were dreaming (or were in any of the other Cartesian scenarios that posit a broad disengagement from reality) we would have no evidence we could legitimately appeal to, to reassure ourselves that there is indeed a world of mind-independent physical entities. For that evidence would be compatible with its being dreamt of (or otherwise produced in abnormal ways).[8]

If it took the form of a Humean paradox, in contrast, scepticism would challenge us by considering that in order to be able to justifiably hold any of our ordinary empirical beliefs concerning specific mind-independent physical objects, based on our perceptual experiences, we should already take it for granted that there is an external world, which is broadly manifested to us through our sense organs. Thus, we cannot obtain a justification to believe in the existence of the external world by running anything like a Moorean argument, as that kind of argument would be question-begging.[9]

Now, it would take us too far afield to rehearse the long history of the attempts to meet these challenges, even if we were to confine ourselves to just its most relevant episodes in contemporary philosophy (for an overview see Coliva 2012). We have quickly seen how a semantic strategy could try to respond to Cartesian scepticism and the reasons why it does not seem ultimately satisfactory. Serious concerns can be raised for contextualist strategies too (see for instance DeRose 1995) which, at any rate, are designed to allow for the legitimacy of sceptical doubts, albeit only in philosophical contexts. Strategies inspired by Robert Nozick's (1981) and Fred Dretske's (1970) work on conditional analyses of knowledge would then make play with counterfactual considerations regarding knowledge (if not justification). While they are considered by several philosophers to

be problematic as analyses of knowledge as such, when assessed with respect to their anti-sceptical potential, they would all seem to suffer from the same kind of very general problem. Namely, let us concede that the possible world where there are physical objects and we perceptually interact with them and the ones where a Cartesian sceptical scenario obtains are far apart. Yet, the question remains as to how we can tell which one, among these possible worlds, is the actual one. Positing that it is the former rather than the latter is philosophically impotent *vis-à-vis* the Cartesian sceptical challenge. By those lights, that is just a mere posit, which reveals our prejudice, preferences, or habits—at any rate, nothing for which we could provide any kind of epistemic justification.

In the remainder of this section, let me just say a few words against two prominent strategies meant to counter the Humean sceptical paradox. The first is Jim Pryor's (2000, 2004), which relies on a liberal (or dogmatic) understanding of the structure of empirical justification, whereby just by having a hand-like experience, absent defeaters, one would thereby possess a justification for "Here is a hand," which would then transmit to the conclusion of a Moorean-argument, namely "There is an external world." Among the various problems that beset this strategy (for criticisms see Coliva 2015, 20–28, 58–64), the one which immediately meets the eye is that if we conceive of experiences in an internalist fashion, as Pryor does, there seems to be nothing in the experience itself to make it more likely that we are indeed seeing a hand, rather than having an impression as of a hand, while being ourselves brains in a vat, say (cf. White 2006 and Wright 2007). If so, then, the strategy does not meet the Humean sceptical challenge. For it does not have the resources to claim that it is indeed the case that our experiences as such, absent defeaters, and unaided by collateral assumptions regarding their provenance, give us a justification to believe that there is a hand here. Absent such an immediate justification, then, the strategy does not have the resources to rehabilitate the epistemic legitimacy of a Moorean anti-sceptical argument.

One might then think that the problem with the liberal counter is not so much the structure of epistemic justification it proposes, as the conception of perceptual experience it relies on. Disjunctivist strategies, proposed by John McDowell (1982, 1995) and, more recently, by Duncan

Pritchard (2016), can be seen as capitalising on this intuition. Accordingly, if we are perceiving a hand, then there is a hand there and we have a justification to believe there is (or even have knowledge of that). If there is no hand, then we are not perceiving in the first place. Yet, there is no highest common factor between perceiving a hand or merely hallucinating, or seeming to see one. If we are in the favourable disjunct, we then possess an immediate justification for (or even knowledge of) "Here is a hand," which transmits to the conclusion of the Moorean argument. Thus, we have a direct reply to the Humean sceptical challenge.

As simple and pleasing as this strategy may be, two immediate worries arise, however. First, it is hostage to a conception of perception that is not sacrosanct, and indeed that is challenged by contemporary work on the psychology of perception.[10] Now, it is true that, methodologically, I think philosophy and psychology should each take care of themselves. Yet, I would be very worried if a philosophical theory about one or the other of our cognitive faculties were not compatible with our best science about that very faculty. True, science may evolve and compatibility between it and a given philosophical view be re-established, but if it doesn't (or if it hasn't yet) we can't simply dismiss its findings with a shrug of the shoulders. Prudence would thus recommend suspending judgement and looking for alternative philosophical strategies. Second, one may happily agree that if we perceive a hand, then we have an immediate justification for (or knowledge of) "Here is a hand" (which could then transmit to "There is an external world" via the Moorean argument). Yet that is just a conditional claim, and the question remains as to how to discharge its antecedent. In particular, how can we tell if we are right now perceiving, as opposed to merely having an experience as of a hand (see Wright 2002)?

Finally, let us acknowledge the interdependence of these very sceptical paradoxes and the realist conception of hinges such as "There is an external world." A moment's reflection suffices to see that the Cartesian paradox trades on the distance, posited by realism, between the mind and the world. The world is what it is, independently of our minds, and if so, why be so arrogant to think that it should be accessible to us or that it should be accessible to us just as is? Whence the play with all those scenarios which are supposed to save appearances (literally), while making salient

to us the discomforting possibility that things (in themselves—one would be tempted to say), are rather different, thus leaving us with the impossibility of ascertaining whether they are one way or the other.

Now, true, the Humean paradox does not make play with far-fetched sceptical scenarios. Yet it tells us that all empirical justification and knowledge take place within a system of assumptions—paramount among them "There is an external world"—for which we cannot have independent justification (or which we cannot have independent knowledge of). Yet it is precisely such independence that would make them justified, if we could somehow bridge the gap. Conversely, it is because such independence is built into the problem that it becomes impossible to redeem our justification for those very assumptions. For, recall, we are looking for an epistemic justification for "There is an external world." That is, a justification which allows us to believe that that is the case (or that it is more likely the case) and therefore that allows us to claim that our representation of the world as populated by mind-independent objects is true because that is exactly how the world is. Yet, if that justification is supposed to come from the senses, hence from our minds, how could it in and of itself certify that it is in keeping with a mind-*independent* reality?

In short, both forms of sceptical paradox make play with the supposed gap between, and independence of mind and world and display the difficulty of reconquering the world, once that distance is taken as given.

3 Anti-realism About Hinges I: An Evidence-Dependent Conception of Hinges' Truth?

One might then be tempted to embrace the opposite, anti-realist side of the Euthyphro contrast and construe it, as has been the case for comic and taste discourse, as grounded in our human reactions (see Wright 1992 for this strategy). For example, "This joke is funny" would not be true because the proposition expressed by that sentence corresponds to a mind-independent reality that already carves up the world into jokes that are funny and jokes that are not. Rather, it is true because—given our

own reactions—we judge that that joke is funny, while other ones are not. Similarly, on this account, "There is an external world" would not be true because it corresponds to a mind independent reality. Rather, it would be true because, based on our evidence, we judge that that is the case. That is, it looks to us as if there are objects that persist even when they are not perceived by us, or that seem to afford coherent perceptions over time and across various sensory modalities. We therefore have reasons to believe that they exist independently of us.

The problem with this strategy, however, is that it is supposed to ground the truth of "There is an external world" in our own evidence in favor of it. Yet, that evidence would be sensory evidence. Now, as we just saw in connection with Cartesian and Humean scepticism, on the one hand, that evidence would be compatible with its being produced in abnormal ways, which do not depend on any causal interaction with physical objects. On the other, we have already pointed out how, in the vein of Humean scepticism, its role as evidence which justifies us in believing that there is a hand, where we seem to see it, for instance, owes to the collateral assumption that it be produced in what we take to be the normal way. That is, through causal, sensory interaction with a world populated by mind-independent physical objects. Thus, the problem facing a supporter of the anti-realist side of the Euthyphro contrast is that the evidence we have cannot justify, in and of itself, the belief in the existence of an external world. Investigating the disanalogy between comic discourse and discourse about the very existence of an external world will prove instructive. So let me address it in the remainder of this section.

With comic discourse, it is readily admissible that we are somehow recording our own subjective reactions to jokes, for instance. Thus, the provenance of truth, when we ascribe it to the proposition expressed by "This joke is funny," can easily be considered to be from our own minds. With "There is an external world," however, there is no such readily available sense according to which we would thereby be recording our own subjective reactions—unless one wished to go idealist, or phenomenalist. Yet, as long as one takes "There is an external world" at face value, as affirming the existence of mind-independent physical objects, the prospects of an evidence-dependent account of its truth look dim.

To put it in terms familiar from the debate on alethic pluralism (see for instance, besides Wright 1992, also Lynch 2009 and Pedersen and Wright 2013), alethic anti-realism, cashed out in an evidence-dependent way as amounting to truth understood, for instance, as superassertability (see Wright 1992), has traditionally been proposed for those areas of discourse for which it is *a priori* that all true propositions are in principle *knowable* (or at least justifiable). Yet, it is the very gist of the two forms of sceptical paradox we have been considering that "There is an external world" does not fall in that remit.

If these were the only options, we would then face a serious situation. As a result of the sceptical paradoxes, we could not know (or justifiably believe) that there is an external world; nor, it seems, could we make sense of the idea that that is a proposition capable of being true, at least in principle.

I surmise that something like finding himself in this bind led Wittgenstein to propose (if not to fully endorse) the view that hinges are neither known, nor unknown, justified or unjustified, and are actually *neither true nor false* (1969, §94, 205). Connectedly, it was (possibly) finding himself in such a conundrum that led him to hold that their role had to be akin to that of rules. That is, rather than be descriptive, their role would be prescriptive—similar to "*p*!" (where *p*, in this case, would be the proposition that there is an external world and the action commended the one of accepting that *p* and of behaving in accordance with what *p* entails).

4 Anti-realism About Hinges II: A Minimalist Conception of Hinges' Truth

Still, Wittgenstein himself in *On Certainty* toys with a different idea regarding hinges' truth. Namely, that:

(1) Their truth "belongs to our frame of reference." (1969, §83) That is, that we regard them "as certainly true also characterizes [our] interpretation of experience." (1969, §145)

(2) The truth of our statements is the test of our understanding of these statements (1969, §80). "That is to say: if I make certain false statements, it becomes uncertain whether I understand them." (1969, §81)
(3) If one thought that "if everything speaks for an hypothesis and nothing against it" it certainly "agrees with reality [that is, with mind-independent] facts", one would already be "going round in a circle." (1969, §191) Again, "the reason why the use of the expression 'true or false' has something misleading about it is that it is like saying 'it tallies with the facts or it doesn't', and the very thing that is in question is what 'tallying' is here." (1969, §199)
(4) Taken together, hinges form our "picture of the world" (*Weltbild*). (1969, §§93–94) Yet, it is not because we are, or have "satisfied ourselves of its correctness" that we hold on to it (1969, §94); rather, this picture of the world is a "background" (1969, §94) we have inherited and against which we distinguish between true and false. Hinges thus form a "substratum" (*Substrat*) for all our enquiries and assertions (1969, §162). And "the propositions describing [our picture of the world] are not all equally subject to testing." (1969, §162)
(5) "It is the truth only inasmuch as it is an unmoving foundation of [the] language-games." (1969, §403)
(6) Their truth is kept fixed by what rotates around them (1969, §152), like our methods of empirical investigation (1969, §151, §318), and "the rest of our procedure of asserting." (1969, §153)

Thus, as a methodological point, when we are dealing with hinges, we have to focus on their role, rather than on their content (and possibly on the context of their use). As is apparent, they make it possible for us to engage in our empirical inquiries and investigations (1); for Wittgenstein, they also play a meaning-constitutive role, therefore they are conditions of possibility of meaningful discourse (2). Given their role, however, nothing can ultimately speak against them. For local doubts would have to presuppose them and global ones would be meaningless and not rational, for Wittgenstein. However, the fact that nothing can speak against them and that everything speaks in favour of them does not mean that they are unconditionally true—that is, that they are true because they correspond to a mind-independent reality that makes them true (3). Nor are they true because all of our evidence speaks in favour of them and

therefore makes them true (3). For they make it possible for that evidence to be evidence in the first place. Thus, neither the correspondentist nor the evidence-dependent conception of truth applies to hinges. Rather, hinges determine the conditions of possibility for saying that a given proposition *p* tallies with the facts, and is therefore true; as well as for saying that *p* is true because all of our evidence speaks in favour of it. Moreover, in a Kantian flavour, being conditions of possibility for any kind of truth, let it be correspondentist or evidentialist, and of empirical rationality, for they make it possible to have evidence for or against ordinary empirical propositions, they do so because they are *a parte subiecti* (4, 5, 6). Clearly there is also a flavour of historicism in Wittgenstein's thought, for, given his wide conception of hinges, he did think that at least some of them were subject to change at least over time (1969, §§96–99). Indeed, "nobody has ever been on the Moon" (cf. 1969, 106), which is a hinge for him, is no longer such for us. Moreover, in the passages on the riverbed metaphor, he concedes that the riverbed, that is our hinges, may change because of empirical discoveries and progress. Furthermore, there is also a flavour of relativism in *On Certainty*, as he sometimes envisions the possibility of communities that do not abide by our hinges and hold on to altogether different ones (1969, §92, §132).[11]

Now, the hinges I have been focusing on in my systematic work are more general and much more limited in number than Wittgenstein's. They are propositions like "There is an external world," "Our sense organs are broadly functioning correctly," or "Our cognitive faculties are broadly reliable." Still, they certainly are conditions of possibility of any empirical inquiry, whether or not they can also play a meaning-constitutive role. Moreover, they are presupposed by any ordinary empirical doubt, while also making it possible for any empirical evidence to play that very role. We can thus apply to them the idea that although nothing can speak against them and everything speaks in favour of them, that does not mean that they are unconditionally true. Nor are they true because they correspond to a mind-independent reality that makes them true. For it is only by accepting a hinge like "There is an external world" that we do have at least one of the terms in the equation for starting thinking that representations may or may not correspond, or tally with a mind-independent reality. The other term, in a Kantian flavour, can be entered

in the equation only by holding on to the idea that there are minds, whose faculties—perceptual and cognitive—are broadly conducive to finding out about such a mind-independent reality. Indeed, this is what the other hinges I have been discussing in my more systematic work would give us. If so, this kind of hinges would be not just a condition of possibility of epistemic rationality, as I have maintained in previous work, but also a condition of possibility of there being representations at all—let them be linguistic, mental, pictorial, et cetera. Yet, it is equally mistaken to think that these general hinges are true because all of our evidence speaks in favor of them and therefore makes them true. For they are the conditions of possibility for that evidence to be evidence at all. As we saw, we need to take for granted that there are physical objects to let our sensory evidence speak in favour of, for example, "Here is a hand." Thus, neither the correspondentist nor the evidence-dependent conception of truth would apply to them. Rather, hinges would be the conditions of possibility for both saying that "*p* tallies with the facts, and is therefore true"; as well as for saying that "*p* is true because all of our evidence speaks in favor of it."

That being said, in a pluralistic spirit, we might want to recognise the need for both kinds of truth, depending on the area of discourse at hand. I will allow myself to speak loosely of different kinds of truth. Really, however, what that means, in a developed pluralistic framework, is just that truth is differently realized (or determined) according to the kind of discourse at hand. More specifically, there would be a truth-predicate—"true"—a truth-concept—TRUE—a generic truth-property—truth-as-such—and its multiple realizations (e.g. truth as correspondence, or as superassertibility, or as coherence etc.). That is to say, when we talk about Mount Everest and its height, we may have good reasons to think that whatever we think or say about that is true iff Mount Everest has such-and-such a property and it is because that is the case that what we think or say is true. Whereas, when we think or talk about a joke being funny, ultimately what we think or say is true iff the joke is funny, where the kind of truth predicated of "This joke is funny" is one that makes non-eliminable reference to human reactions. Thus, it is because we find that joke funny and are prepared to judge and assert that it is, that it is indeed funny. Now consider a sentence like "This sushi is tasty" and notice that it actually rests on "This is sushi" and "This is tasty." On this pluralist approach, the first sentence would express a proposition whose truth is mind-independent,

while the second will express a proposition whose truth is grounded in our reactions. Mixed cases like this one are widespread and their significance is that, in many cases, our truth-determining reactions concern items belonging to a mind-independent reality. Thus, even if one did not endorse an evidence-dependent conception of truth for propositions about specific physical objects, it is of the essence of those areas of discourse for which an evidence-dependent account of truth has been invoked by alethic pluralists—that is, comic and taste discourse, but also aesthetic and moral discourse—that the relevant human reactions that are deemed truth-determining targeted physical objects (including people), to which we have access and about which we can form beliefs, based on the operations of our sensory and cognitive faculties. In that mediated sense, then, even for a pluralist of the kind I am envisaging, the truth of "There is an external world," et cetera, would have to be taken for granted.[12]

Now, at this point, we might quite naturally want to stop thinking of hinges as being true (or false) altogether. I think that is what Wittgenstein ultimately ended up doing.[13] Yet, we need not have to do that, as long as we are hospitable to the idea that truth may be sensibly predicated of hinges such that they can figure in disquotational schemas; that they can be embedded in conditionals, and thus be semantically assessable, or be embedded in meaningful negations, at least in further conditional statements. Yet, such predication would not be metaphysically robust, clearly. For the reasons just rehearsed, it could neither be taken as an ascription of a correspondence between these hinges and a mind-independent reality that makes them true, nor could it be taken as expressing the idea that since all of our evidence speaks in favour of them, that very evidence makes them true. It is thus in connection with hinges that I find a minimalist account of truth eminently plausible. However, "minimalism" is a name used for two different positions in the literature on truth and objectivity. One is Horwich's (1998) minimalism, which is a form of deflationary theory of truth, which is non-redundantist and propositionalist as opposed to sententialist. The other one is Wright's (1992) minimalism, according to which a truth-predicate has to satisfy a number of platitudes, like the ones just mentioned, while the truth-property it expresses may vary as a function of the area of discourse at hand. By "minimalism" I will refer to the—no doubt unusual—synthesis of these views. For, like Wright and unlike Horwich, I take minimalism to be compatible with

alethic pluralism, while, *contra* Wright and in the vein of Horwich, I am proposing that in certain areas of discourse the kind of truth predicated is exhausted by the T-schema and the platitudes traditionally licensed by deflationism. Discourse about hinges would be such an area. Thus,

(T_{EXT}) The proposition expressed by "There is an external world" is true iff there is an external world

together with platitudes about negation, conditionals (and possibly some more see Lynch 2009) would be all there is to predicating truth of that proposition. Borrowing Lynch's terminology, we may call hinges "plain truths." That is, propositions whose truth just consists in the satisfaction of the relevant platitudes and is not grounded in anything else. Thus, they are not true in virtue of their correspondence to a mind-independent reality, or in virtue of being supported by all available evidence.

The virtue of such an approach would be to maintain the very possibility of predicating truth for these kinds of problematic propositions, while eschewing the problems that beset any of its possible stronger understandings. It would also tally quite well with the underlying understanding of hinges as conditions of possibility for representation, rather than as representations of a mind-independent reality themselves. Truth, in their regard, would thus be a "logical," or "grammatical" predication, which would not carry any stronger ontological commitment *per se*. Rather, it would simply entail that discourse about mind-independent objects is possible and is actually admitted among the kinds of discourse we engage in. Once that discourse is in place, specific moves within it become possible. Thus, a correspondentist account of truth could still find purchase for *specific* empirical claims such as

(T_{EVE}) The proposition expressed by "Mount Everest is more than 8,000 meters high" is true iff Mount Everest in more than 8,000 meters high.

should one feel so inclined. Of course, the resulting realism would be weaker than what traditional supporters of that view might have thought (or wished). For it would be a kind of "internal realism," dependent on a previous commitment to the legitimacy of an entire area of discourse, which

could not itself be redeemed in a correspondentist fashion.[14] Yet, so much seems to me to be entailed by the endorsement of a hinge epistemology of Wittgensteinian descent. And if one did feel a hint of discomfort, one should always remind oneself of the alternatives, which I take to be worse by the lights of a self-aware, rather than naïve realist perspective. Namely, either to embrace deflationism about truth all the way down, including (T_{EVE}) and its kin; or else to embrace correspondentism all the way down, but with dim prospects of being able to respond to sceptical paradoxes.

Notice, however, that there would be no hint of historicism or even of relativism in the kind of minimalism about hinges I have been proposing. This is because the hinges I have been considering and for which I have been proposing a minimalist treatment are very general ones, and can legitimately be seen as conditions of possibility of representational discourse *tout court*. The same does not seem true of many of Wittgenstein's hinges like "Nobody has ever been on the moon," or "My name is AC," "I have never been to Australia," "Nobody can make it rain," "Nobody can bring the world into existence with their own birth," et cetera, which may be subject to historical, cultural, or even personal variations. To repeat, I have been concerned with those hinges that allow us to think and talk about a mind-independent reality, and with those that allow us to take ourselves to be capable of engaging with such a world through our senses and our cognitive faculties.

Another issue which needs to be addressed, albeit briefly, is whether there is anything optional in these hinges. That is, could we opt for different ones or even opt out of them altogether, without endorsing any different ones? Idealists or phenomenalists may be taken as opting for different ones, such that physical objects would then coincide with clusters of sense experiences, or the like. Now, to the extent that such a position is ultimately coherent, for these sensory impressions would have to be produced by the interaction between our senses, which are physical organs, and something out there, which would be mind-independent, it would actually deprive us of the very possibility of engaging in fully representational discourse and would ultimately deprive us of the very possibility of engaging in epistemic rationality. That is to say, they would deprive us of the very possibility of engaging in those practices by means of which we hold, revise or withdraw from beliefs about mind-independent objects on the

basis of evidence bearing on the mind-independent truth of those beliefs. These views, if ultimately coherent at all, would, at best, confine ourselves within a realm of internally coherent beliefs, which would not allow us to consider the latter as reaching anything else outside of them. Equally, opting out of these very general hinges, without endorsing different ones, would ultimately result in the impossibility of engaging in representational discourse and in epistemic practices constitutive of epistemic rationality, with not even a *prima facie* substitute for them. Most likely, that would result in cognitive paralysis. Thus, to the extent that representational discourse and epistemic rationality are valuable to us, we do have a strongly pragmatic, and as a universal as that might sensibly be reason to hold on to these very general hinges that make all of that possible.[15]

One may then wonder in what sense, if any, minimalism about hinges, as presented here, would count as an anti-realist position. It would clearly count as such if by "realism" one meant the view traditionally associated with correspondentism about truth. For minimalism denies that hinges are true because they correspond to a mind-independent reality. Still, minimalism would not be anti-realism traditionally understood either. For minimalism denies that hinges are true because our evidence speaks in favor of them. Eventually, I think that minimalism would pair very well with a form of quietism about the ultimate metaphysical status of hinges, understood as being the very conditions of possibility of representational talk and of epistemic rationality.

Thus, when it comes to hinges, if one thought that one would already be "going round in a circle" (1969, §191), if one tried to understand hinges' truth along evidence-dependent lines, and if one thought that "the … use of the expression 'true or false' has something misleading about it," for "it is … like saying 'it tallies with the facts or it doesn't,'" when "the very thing that is in question is what 'tallying' is here" (1969, §199), then one should realise that the moment has come to cease to engage in robustly metaphysical speculations. That is where the spade reaches the bedrock and it would be turned if we did not stop and acknowledged that this is simply what we do.

Finally, one might wonder in what sense, if any, turning minimalist with respect to hinges' truth and quietist about the existence of an external world would help one respond to sceptical paradoxes. The answer to that worry is this: by rejecting the supposed unbridgeable distance

between mind and world on which both Cartesian and Humean scepticism trade (as we saw in Sect. 2). For, on the picture advocated here, we are well within our rights in holding on to the existence of an external world broadly manifested to us through our sensory organs. Yet, we are within our rights not because we have a proof of that, but because it is part and parcel of engaging in representational discourse and in epistemic rationality as we—and, crucially, sceptics—do, that we take for granted the existence of physical objects and that our sensory faculties reliably allow us to engage with it.

Notes

1. Clearly, Wittgenstein objected to the very possibility of global Cartesian scepticism for reasons briefly alluded to in the main text. Against a wholesale scepticism of Humean descent, he did appeal to the idea that rational—that is, justified—doubts could only be local. If so, hinges would have to be taken for granted to be able to formulate them. For a discussion of Wittgenstein's various anti-sceptical strategies, see Coliva (2010, chap. 3).
2. For other general presuppositions of epistemic rationality, see Coliva (2015, chap. 5).
3. Moyal-Sharrock (2005) has defended the anti-propositionalist reading of Wittgenstein's hinges. For a critical discussion, see Coliva (2010, chap. 2, 5).
4. On the notion of acceptance, see Coliva (2015, 33–39).
5. The charge is due to DeRose (1995). For a response, see Sherman and Harman (2011) and Coliva (2015, 136–39).
6. See Coliva (2015, 2017a, b, chap. 4) for replies to possible criticism.
7. I prefer to use the label "minimalist," rather than "deflationary," for reasons which will become apparent in the following.
8. This is the well-known problem of the criterion.
9. A Moorean argument would have roughly the following form:

   (I) Here is a hand.
   (II) If there is a hand here, there is an external world.
   (III) There is an external world.

   Now, (I) would be justified on the basis of one's on-going sensory experience. Yet, for that experience to be justifiably brought to bear on a belief about the existence of a specific physical object like a hand,

(III) should already be taken for granted. Hence, whatever perceptual justification one could have for (I) could not be transmitted to (III), since prior justification for (III) would be needed in order to justifiably take (III) for granted and therefore for having a justification for (I) in the first place (Wright 1985). In Coliva (2015, chap. 3), I have proposed a different notion of transmission failure, which would depend merely on taking (III) for granted—without the need to have a justification for it—for (I) to be justified on the basis of one's on-going experience. For present purposes, it is irrelevant which kind of transmission failure is taken to apply to a Moorean argument.

10. For a philosophically informed discussion of work in contemporary cognitive psychology, see Burge (2010).
11. I do not think that ultimately Wittgenstein was a relativist, but just an anti-realist. I have discussed this issue at length in Coliva (2010, 188–207).
12. For different treatments of mixed compounds and inferences, in the vein of alethic pluralism, see Lynch (2004), Edwards (2011) and Pedersen and Lynch (2018).
13. I thus concur with Moyal-Sharrock's (2005) interpretation. Cf. Coliva (2010).
14. This would tally with Wittgenstein's own understanding of "There are physical objects" as licensing talk about such a category of entities and of "A is a physical object" as being merely "a piece of instruction, which we give only to someone who doesn't yet understand either what 'A' means, or what 'physical object' mean." (1969, 36, cf. 35–37)
15. It falls beyond the scope of this paper to argue for the stronger claim that we have more than simply a pragmatic reason. A line one might want to explore, in order to argue for that stronger claim, is that being inquirers and knowers is essential to our identity, if not nature. Thus, insofar as we want to maintain that identity, or even more strongly, in virtue of our own nature, we cannot but stick to these hinges.

References

Burge, Tyler. 2010. *Origins of Objectivity*. Oxford: Oxford University Press.

Coliva, Annalisa. 2010. *Moore and Wittgenstein: Scepticism, Certainty and Common Sense*. Basingstoke: Palgrave Macmillan.

———. 2012. *Scetticismo. Dubbio, Paradosso, Conoscenza*. Roma-Bari: Laterza.
———. 2015. *Extended Rationality: A Hinge Epistemology*. Basingstoke: Palgrave Macmillan.
———. 2017a. Précis of Extended Rationality: A Hinge Epistemology. *International Journal for the Study of Skepticism* 7 (4): 217–234.
———. 2017b. Replies to Commentators. *International Journal for the Study of Skepticism* 7 (4): 281–295.
DeRose, Keith. 1995. Solving the Skeptical Problem. *Philosophical Review* 104 (1): 1–52.
Dretske, Fred. 1970. Epistemic Operators. *Journal of Philosophy* 67 (24): 1007–1023.
Edwards, Douglas. 2011. Simplifying Alethic Pluralism. *Southern Journal of Philosophy* 49 (1): 28–48.
Horwich, Paul. 1998. *Truth*. Oxford: Clarendon.
Lynch, Michael. 2004. Truth and Multiple Realizability. *Australasian Journal of Philosophy* 82 (3): 384–408.
———. 2009. *Truth as One and Many*. Oxford: Oxford University Press.
McDowell, John. 1982. Criteria, Defeasibility and Knowledge. *Proceedings of the British Academy* 68: 455–479.
———. 1995. Knowledge and the Internal. *Philosophy and Phenomenological Research* 55 (4): 877–893.
Moyal-Sharrock, Danièle. 2005. *Understanding Wittgenstein's On Certainty*. Basingstoke: Palgrave Macmillan.
Nozick, Robert. 1981. *Philosophical Explanations*. Cambridge, MA: Harvard University Press.
Pedersen, Nikolaj J.L.L., and Michael Lynch. 2018. Truth Pluralism. In *The Oxford Handbook of Truth*, ed. Michael Glanzberg. Oxford: Oxford University Press.
Pedersen, Nikolaj J.L.L., and Cory Wright, eds. 2013. *Truth and Pluralism: Current Debates*. Oxford: Oxford University Press.
Pritchard, Duncan. 2016. *Epistemic Angst: Radical Scepticism and the Groundlessness of Our Believing*. Princeton: Princeton University Press.
Pryor, James. 2004. What's Wrong with Moore's Argument? *Philosophical Issues* 14 (1): 349–378.
Pryor, Jim. 2000. The Sceptic and the Dogmatist. *Noûs* 34 (4): 517–549.
Quine, W.v.O. (1969) 2008. Epistemology Naturalized. In *Epistemology: An Anthology*, ed. Ernest Sosa and Jaegwon Kim, 292–300. Oxford: Blackwell.
Sherman, Brett, and Gilbert Harman. 2011. Knowledge and Assumptions. *Philosophical Studies* 156 (1): 131–140.

White, Roger. 2006. Problems for Dogmatism. *Philosophical Studies* 131 (3): 525–557.
Wittgenstein, Ludwig. 1969. *On Certainty*. Oxford: Oxford University Press.
Wright, Crispin. 1985. Facts and Certainty. *Proceedings of the British Academy* 71: 429–472.
———. 1992. *Truth and Objectivity*. Cambridge, MA: Harvard University Press.
———. 2002. (Anti-)Sceptics Simple and Subtle: G. E. Moore and John McDowell. *Philosophy and Phenomenological Research* 65 (2): 330–348.
———. 2007. The Perils of Dogmatism. In *Themes from G. E. Moore: New Essays in Epistemology*, ed. Susanna Nuccetelli and Gary Seay, 25–48. Oxford: Oxford University Press.



Epistemic Schmagency?

A. K. Flowerree

1 Introduction

Constructivist approaches in epistemology and ethics offer a promising account of normativity. Norms are authoritative—when they are—in virtue of being constitutive of our agency (or they derive authority from norms that are constitutive of our agency). But constructivist accounts face a powerful objection raised by David Enoch, the Schmagency Objection. David Enoch (2006) argues that constructivist accounts of normativity face an insuperable challenge. While his Schmagency Objection has been widely discussed as a challenge to constructivist views of moral and practical norms, no one has yet explored how the Schmagency Objection might undermine popular approaches to epistemic normativity. In this paper, I rectify that gap; I explore whether the Schmagency Objection holds in the epistemic case. First, I develop the objection against a popular and prominent form of epistemic constructivism, Belief

A. K. Flowerree (✉)
Department of Philosophy, Texas Tech University, Lubbock, TX, USA
e-mail: aflowerr@uni-koeln.de

© The Author(s) 2018
C. Kyriacou, R. McKenna (eds.), *Metaepistemology*, Palgrave Innovations in Philosophy, https://doi.org/10.1007/978-3-319-93369-6_13

Constitutivism. Belief Constitutivism holds that epistemic normativity is grounded in the constitutive aim of belief, namely a belief is correct just in case its contents are true. Belief Constitutivism is susceptible to a Schmagency Objection, I argue, because it locates the source of normativity in the *belief* rather than the *agent*. In the final section, I propose a version of epistemic constructivism that locates epistemic normativity as constitutive of agency, rather than constitutive of belief. I argue that this version has the resources to respond to the Schmagency Objection. My view has the added benefit of illuminating the role that belief plays in action.

2 The Schmagency Objection

Many practices have constitutive norms. The norms of gift-giving, as recorded in Amy Vanderbilt (1952), dictate that one ought to mail a note of thanks as soon as possible, "within a week preferably." The norms of being cool, at some time and for some people, involved wearing backward baseball caps. The norms of greetings, also at some time and for some people, involve a firm-but-not-too-firm handshake lasting 1–2 seconds. When we participate in these activities, we ought to follow the norms. But the question arises: *should* we participate in these activities? And if we do, should we participate *well*? From the fact that there is a constitutive norm, it doesn't follow we have reason to comply with it. David Enoch develops this point into the Schmagency Objection against constructivist accounts of normativity. In this section, I will lay out the constructivist strategy, present Enoch's Schmagency Objection, and then develop the constructivist response to the Schmagency Objection. I will sketch a general outline of the strategy that any constructivist view must take to overcome Enoch's Objection.

The constructivist strategy, as it is understood by Enoch, and as I will understand it in this paper, attempts to show that normativity is constructed from the constitutive norms of agency.[1] The constructivist argues that practical reasons (reasons that justify engaging in contingent practices like gift-giving, being cool, and greeting each other) and moral reasons arise from the constitutive norms of agency.

In general, constitutive norms govern various activities and give rise to reasons within those activities. We see this sort of normativity all around us. Take the perennial example of chess. Chess is a game that is governed by rules, and the aim of playing chess is to win. Within the game there are certain norms—norms of the mechanics of the game, norms of strategy—and these norms generate reasons for performing particular movements within the game. But unlike chess, the constructivist argues, agency is "closed under its own operation." (Ferrero 2009, 308) That means we cannot coherently ask whether we have good reason to follow the norms of agency. And so while the norms of chess are contingent requirements that can be abandoned at any moment if the agent quits the game, practical and moral norms are inescapable. On the constructivist view, normative questions arise from within a standpoint. There is no standpoint from which one could reason about whether to play the agency game. Whatever you appeal to (implicitly or explicitly) relies on the fact that you are already engaged in the activity of agency. Moral and practical norms are constitutive of the activity of agency.

A virtue of this account is that can answer certain sceptical challenges. The sceptical challenge that constructivists (particularly Christine Korsgaard) have in mind is the requirement of reason turned inward (Korsgaard 1996). In order for something to be truly normative, Korsgaard argues, it must be something that we can rationally view as authoritative. Constitutive norms are resilient in the face of sceptical challenges because they bind the activity *essentially*. The normative question is answered by reflecting on the type of creatures we are (creatures who are constituted by a set of norms) and the types of norms at play (norms that govern us and our activity essentially).

In general, constructivists who hope to derive normativity from the nature of agency hope to ground all normativity: moral, practical *and* epistemic. In metaethical debates, the focus has primarily been on practical and moral norms. But attentive readers will notice that constructivists and their critics express hope that the epistemic case is no different from the practical.[2]

2.1 Schmagent's Soliloquy

Enoch argues that constitutive norms cannot ground normativity. He illuminates this point by giving us the character of a schmagent. The schmagent rebuffs the constructivist's normative judgements about what he ought to do. Here, the schmagent delivers the Schmagent's Soliloquy:

> Classify my bodily movements and indeed me as you like. Perhaps I cannot be classified as an agent without aiming to constitute myself. But why should I be an agent? Perhaps I can't act without aiming at self-constitution, but why should I act? If your reasoning works, this just shows that I don't care about agency and action. I'm perfectly happy being a "Schmagent"—a nonagent who is very similar to agents but who lacks the aim (constitutive of agency but not "schmagency") of self-constitution. I am perfectly happy performing "schmactions"—nonaction events that are very similar to actions but lack the aim (constitutive of actions but not schmactions) of self-constitution. (Enoch 2006, 179)

The Schmagent's Question, then, is "Why should I be an agent?" or "Why should I play the agency game?" We should understand the schmagent's question as a query for justification (i.e. a reason) for why it is that the schmagent ought to view the norms of agency as authoritative. The Schmagency Objection is that if there is no answer to the Schmagent's Question, then we cannot see the requirements of agency as justified. And if we cannot see those requirements as justified, then we cannot be justified by following them.

This soliloquy is meant to capture the way in which constructivist accounts fail to ground the right kind of normativity (or perhaps any normativity at all) for moral and practical norms. Enoch's central point is that agents are in a position to challenge the normative status of a normative practice (as we did above with gift-giving, being cool, and greetings). The constructivist, Enoch claims, has no satisfactory answer to the Schmagent's Question. From the fact that the agent engages in some normative practice we cannot conclude that the agent *ought* to engage in a normative practice.[3] In order for the practice to normatively bind the agent, the agent must have *reason* to engage in the practice.

This problem is especially pressing for the constructivist because she cannot appeal to any reasons outside of the norms of agency (Enoch 2006). The constructivist approach attempts to explain all normativity in terms of the constitutive norms of agency. As a result, if we challenge whether agency itself is justified, the constructivist cannot appeal to anything outside of agency for her response. Otherwise, the view is no longer constructivist. And if we can demand justification for activity of agency, appeals to what would justify something *within* the practice of agency will not help us justify *the practice itself*. And so, the constructivist has no resources to answer the Schmagency Objection.

Without a reply to the Schmagency Objection, Enoch thinks, the most we can provide in terms of normative assessment is a semantic classification, and this renders all norms toothless. In order for something to be classified as a game of chess, or an agent, certain norms must have force. But what happens if we don't follow the rules of chess, or we rebuff the requirements of agency? It doesn't follow that we've done something *wrong* or *violated a requirement* in a deep sense. Instead, it follows that we weren't really playing chess, or we weren't actually agents. If we were to confront the haphazard chess player or a recalcitrant agent, they could reply to us, "so I'm not following the rules—who cares? Why should the rules motivate me? The only sanction is that I fail to deserve the folk-theoretical term chess player or agent."

Constructivists claim that agency is inescapable, we have no recourse but to be agents. As Korsgaard writes, agency is "our plight" (2008, 2). But this, Enoch argues, is insufficient. You may be doomed to some activity, but that does not mean you have *reason* to engage in the activity. It doesn't mean you have *reason* to follow the norms of the activity. Enoch writes, that it is never a good answer to the question "why should I φ?" to say "because you are a φ-er." (2011, 213) Thus, Enoch concludes, constructivism cannot ground normativity.

2.2 The Constructivist Response to the Schmagency Objection

The constructivist response is to show that either the answer to the Schmagent's Question is trivial, or the question is semantically defective.

The Schmagent's Question is ambiguous between two readings. On one understanding of the question, it can be answered trivially. On the other, it presupposes an incoherent state of affairs. It is helpful to keep in mind Rawls' (1955) distinction between internal justification (justification within a practice), and external justification (justification by something outside the practice). The Schmagent's Question is a request for either internal or external justification. If it is an internal question, it takes place within the activity of agency, and so can be answered according to the internal norms of agency. If it is an external question, then the constructivist argues that the question is semantically defective. The Schmagent's Question presupposes that we can consider reasons *outside* the practice of agency. But, says the constructivist, this is like asking, whether a "telephone is correct rather than a tree." (Velleman 2009, 145) The schmagent demands that we justify the demands of agency *independently of any standard whatsoever*. But the constructivist has argued that reasons can only be in force relative to some standards. And so the constructivist can give no answer to the question, but that is because the question is semantically defective.[4]

While the exact details of the response to the Schmagency Objection can differ, any successful response must involve showing the external version of the Schmagent's Question to be defective. The constructivist cannot provide an answer to the question. She cannot, in principle, appeal to reasons why one should be an agent rather than a schmagent. This would be to give up constructivism. And it seems that if she accepts that there is no reason to be an agent, the fully reflective agent cannot see herself as justified in following the norms of agency. In order to diffuse the Schmagency Objection, the constructivist must reject the Schmagent's Question.

On the internal disambiguation of the question, the constructivist argues that the Schmagent's Question is answered trivially. This is because constitutive norms bind the activity *essentially*. And so from within the practice, the question "Why are these the norms of the practice?" is answered by appealing to the nature of the activity itself. When we are playing chess, the rules of chess bind us. The nature of the activity and its constitutive rules answer the question, "because those are the norms of this activity." And this is a sufficient answer, within a practice, for the normative status of some rule. Within a practice, it is sufficient to appeal

to the nature of the practice itself. To ask for additional justification for the rules is not intelligible. In chess, the question, "why is it permissible for you to castle?" can only be answered by "because that's the rules of chess, the game we are playing!" Contra Enoch, if the schmagent's question is raised within the practice, it is appropriate to appeal to the practice itself to diffuse the question. Similarly, if the question is, within the practice, "why should I follow the norms of agency?" it is appropriate to answer, "Because you are an agent!" Perhaps there is no deep reason for why we are playing chess rather than schmess. Perhaps there is no deep reason for how I ended up an agent rather than a schmagent. But given that we are playing chess, the norms of chess apply to our game. And given that I am an agent, the norms of agency apply to me in virtue of my very nature. So from within the practice, there can be no Schmagency Objection to constructivism.

And the external version of the question, the constructivist argues, is incoherent. In order for the objection to get going, we must be able to raise the question from outside the practice of agency. In general, we can raise external challenges for justification from outside a practice. We can ask *which* practice to engage in.[5] We can decide, for example, whether to play chess or schmess. Enoch argues that in order to have reasons to make moves within the game of chess, I must have reason to play chess. This is why the schmagent retreats from the constructivist's criticism by saying, "Fine! I'll be a schmagent, then!" The schmagent, standing between two practices, picks one. The schmagent is asking for a *reason* to engage in one activity over another (or no activity).[6]

In order to diffuse the Schmagency Objection, the constructivist must show that the Schmagent's Question is ill-formed. The structure of the response goes something like this. The Schmagent's Question presupposes a substantive claim about normativity, one that the constructivist has explicitly rejected. The substantive claim is this: that we can settle what we have reason to do *independently* of the practice of agency. The constructivist argues that in order to have authority over us, a norm must relate to our agency, and it must be constitutive. The Schmagent's Question presupposes that something could be authoritative for us independently of agency. Because the question involves a presupposition failure, the constructivist rejects the external version of the question.[7]

Constructivists give different accounts of the semantic defect contained within the Schmagent's Question. For Korsgaard, the question is defective because normative concepts embody solutions to practical problems, problems that arise within the domain of agency (1996, 47). As agents, our normative concepts embody norms of agency. They can only correctly be applied *within* the practice of agency. Perhaps if there are schmagent normative concepts, they can answer what to do within the practice of schmagency.[8] But for creatures like us, we can correctly apply normative concepts *within* the practice of practical reasoning. The concepts cannot be correctly deployed *outside* the practice. By contrast, Velleman (2009) argues that something can be normative only relative to some standard. So it is nonsense to ask whether something is normative independently of any standard whatsoever. To ask what we have reason to do independent of any standard whatsoever is semantically ill-formed. As Velleman puts it, you have not even asked a question yet, since there is nothing that could count as a correct answer to the question.

Regardless of the exact nature of the conceptual confusion, both agree that the concepts cannot be correctly applied outside the domain of agency. The Schmagent's Question is like asking, "How do I castle outside the game of chess? I know what it is to castle *in* chess, but how would I castle *independently* of the practice of chess? Independently of *any* practice whatsoever?" But the concept of castling is essentially constructed from *within* the practice of chess. There is no answer to this question, because the concept can only have correct application within the practice. Similarly, normative concepts are only correctly applied *within* the domain of agency. And so trying to ask normative questions outside the domain of agency is like asking how to castle outside any practice whatsoever. The question cannot be answered because it involves a conceptual confusion.

The success of the Schmagency Objection, then, rests on a substantive assumption about the nature of normative concepts and normativity in general. The realist disputes the constructivist's conception of what it is for a reason to be authoritative. The objection is persuasive to the realist, and not to the constructivist. In order to settle whether the objection succeeds, we would need to make progress on the disagreement between the realist and the constructivist. Here, I do not mean to definitively settle

the debate. My main aim is to sketch the Schmagency Objection, and show the response that the constructivist must give in order to diffuse the objection. She must deny the question. In the next section, I will explore whether a popular constructivist account of epistemic normativity can make the same move.

3 Epistemic Schmagents

I mentioned at the outset that constructivists and their critics have expressed hope that the same moves can be made for epistemic normativity. However, when constructivist views are applied to the epistemic domain, they take a very different form than from the practical and moral case. In the practical case, constructivists argue that practical and moral norms arise out of the nature of *agency*. By contrast, the most prominent epistemic constructivist views hold that epistemic norms arise out the nature of *belief*, not the nature of agency.[9] And the activity of agency, I will argue, is structurally quite different from the nature of the propositional attitude of belief. The result of this asymmetry is that a pressing Schmagency Objection emerges for constructivists who locate epistemic normativity in belief.

According to one prominent constructivist strategy, which I will call Belief Constitutivism, epistemic normativity is generated by belief's internal standard of correctness.[10] This standard of correctness is expressed (with or without caveat) as:

(SOC) S's belief that p is correct iff p is true.

The proponent of Belief Constitutivism argues that epistemic normativity is best understood as constitutive of our concept (Shah and Velleman 2005). This means that our concept of belief contains within it certain normative standards. We evaluate beliefs positively if they meet these standards, and negatively if they do not. For example, part of what it is to be a knife is to be an object that is good at cutting things. To be a knife is to be sharp. A knife is a good knife just in case it is sharp. And it is a bad knife if it is not sharp. This sort of normativity is often present in arte-

facts. Artefacts are objects that are created to serve a function, and they can serve that function well or poorly. It is part of the nature of the thing to perform some function, and so we can say that it is *constitutive* or essential to the thing that it is normatively governed by that standard (Korsgaard 2008; Wright 1992; Velleman 2000).

What makes Belief Constitutivism constructivist? According to the proponent of Belief Constitutivism, epistemic normativity arises from belief's internal standard of correctness. When we deploy the concept of belief, the standard of correctness determines the conditions under which the belief is correct. Belief Constitutivism does not ground epistemic normativity in anything beyond the attitude of belief. For the attitude to count as a belief, it must have an internal standard of correctness. SOC captures a good-making feature of belief, *qua* belief. But this standard of correctness is a normativity of kind, not a normativity of activity. Epistemic normativity, on Belief Constitutivism, is like knife normativity.

(SOC-Knives) A knife is correct iff it is sharp.

It is part of the nature of knives that they ought to be sharp. This sort of normativity is very different from the normativity of activities. The normativity of games gives us reasons to perform certain actions within the game. But SOC-Knives don't give us reason to perform any action, or do anything. It is a good-making feature of a knife that it is sharp. Knives can be evaluated as correct, as exemplars of the kind knife, just in case they are sharp. But I may not need a sharp knife (suppose I am using it to tap a rhythm). From the fact that it is correct for knives to be sharp it does not follow that I have any reason to sharpen a knife, or even any reason to care if my knife is sharp. I have a reason to care if my knife is sharp only if I need a sharp knife. But if I'm tapping a rhythm, then the sharpness is irrelevant to me.

A constitutive standard of correctness can ground normative evaluations. It can ground the judgement that, for example, the knife is not a very good knife because it is dull. It can ground the judgement that it is an excellent knife because it is sharp. If we were to apply this to belief, we could ground the claim that a true belief is a good belief and a false belief

is a bad belief. But from these evaluations, what else follows? A standard of correctness does not, in itself, provide an agent with any *reason* to do anything. It is not authoritative for the agent. It only grounds judgements for evaluating the object. Sometimes, butter knives are useful; sometimes, false beliefs are useful. It would make no sense to criticise an agent for using a dull knife to cut butter. The knife is perfectly suited for the agent's purpose. If the dullness of the agent's knife gives the agent no reason to sharpen it, then why would the falseness of a belief give the agent a reason to give up the belief? In order to see the judgement of a false belief as a criticism of the agent, we must tell some further story about why the agent should have correct beliefs. An internal standard of correctness is ill-suited to ground epistemic normativity because it cannot show the normative standards to be authoritative for the agent. In what follows I will develop this central point.

A standard of correctness like SOC indicates a goodness of kind. However, goodness of kind does not ground any sense in which the agent should be concerned with bringing about correct instances of that kind. Above, we discussed SOC-Knife. It is correct for a knife to be sharp. Sharpness is a good-making feature of a knife. But if I only need to cut butter, I may not care if my knife is sharp. It is no criticism of me that I keep dull butter knives in my kitchen drawer.

Goodness of kind does not mean a good instance of the kind is good. Goodness of kind is not good, for example, if the thing is bad. A perfect exemplar of a guillotine is a heinous object. A good virus adapts quickly to ensure continued existence, but we still ought to eradicate such viruses if we can. And many good instances of kind are neither good nor bad, but their goodness depends on feature of the situation. Knives illustrate this. So what about beliefs? Appealing to internal correctness doesn't tell us whether good instances of the kind are good. In order to understand why the agent should accept goodness-*qua*-belief as relevant, we must have some further account about why that kind (or good instances of that kind) generates reasons for us to regulate an item according to the good-making features of its kind. But this is just to admit that the constructivist project has failed. If we must appeal to additional normative facts in order to ground epistemic normativity, we have not given a constructivist account.

In the previous section, I outlined that the constructivist must diffuse the external question in order to avoid the Schmagency Objection. But this version of Belief Constitutivism does not even have the resources to answer the *internal* constitutive question. Because there is no practice, no activity, there are no norms internal to the practice. A standard of correctness doesn't give the agent any reason to do anything. I am not a belief, nor am I constituted by any particular belief that satisfies the norm of correctness. It is always open to me to step back from my beliefs and question whether I should continue in them.[11] Judith Jarvis Thomson, in her discussion of Belief Constitutivism, reaches the same conclusion.[12] She writes, "There is no enterprise of believing, and *a forteriori*, there is no such thing as carrying out well the enterprise of believing." (2008, 112) And so Belief Constitutivism, as formulated, cannot answer *any* normative challenge.

3.1 From *Belief* to *Believer*

Perhaps we should shift from talking about *belief* to talking about *believers*. The attitude of belief is integrated into our psychology. Beliefs are not like knives, ordinary objects that we can use for whatever purpose we deign. Beliefs are a part of us. We cannot help but have beliefs; we cannot help but be believers. And there is an activity that we recognise as believing. It involves regulating one's beliefs for truth. A charitable reading of SOC suggests that we interpret it as applying to the agent's relationship to her beliefs, not only the belief itself.

The Schmagency Objection arises here as well. The agent is a believer, engaged in the practice of belief. The agent can still question the significance of truth for *her*. "Fine!" she might say, "Truth is normative for belief! But why should I believe? And if I do believe, why should I believe well? There are plenty of things that have an internal standard of correctness that need not be normative *for me*. So why should believing well matter for me? Rather than be a believer, I will be a schmeliever!"

In Sect. 2, I argued that the constructivist can only diffuse the Schmagency Objection by showing the external version of the question to be ill-formed. But in the case of the Epistemic Schmagent, the question

is quite well formed. In fact, it's one that we regularly face. Consider one case. Psychologists have documented the phenomenon of optimism bias.[13] Optimism bias involves a systematic skewing of beliefs centred on assessing oneself and one's opportunities as much better than they are. People who have optimism bias, psychologists tell us, are more successful in their endeavours. They handle disappointments better. They are happier. Suppose Alex is confronted with this literature, and she decides to cultivate an optimism bias. We can imagine her saying much the same words as the schmagent. "Why should believing well matter for me? I'd far rather be an schmeliever (someone who governs her beliefs according to the norm of optimism bias), than a believer! My life will be so much happier!"

Alex's response is intelligible. More than that, it even sounds plausibly rational. And remember, the structure of the Schmagency Objection is that if we cannot give an answer to the Schmagent's Question, then the norms of believing could not be authoritative for us. And so Belief Constitutivism has a schmagency problem. Not only does the Schmagent's Question sound intelligible, it even sounds like the right way to go, in some instances. Belief Constitutivism can only show the importance of truth for belief, not the importance of truth *for the agent*. It can ground normative assessments of a qualified nature. *Qua* believer, Alex ought to believe her chances of success are quite low. But *qua* schmeliever, Alex ought to assess her chances of success as much higher. And Alex can intelligibly ask, "Why be a believer rather than a schmeliever?" Belief Constitutivism can answer what we should believe *qua* believers, but it leaves open what role belief should play in our psychology. Perhaps the upshot is that for creatures like us, we shouldn't be believers, or we shouldn't be good believers.[14]

Where do we go from here? There are a few options. One is to ground epistemic normativity in some further normativity. The result would be a kind of conditional normativity. Another strategy would be to engage the schmagent and try to justify the authority of epistemic norms on practical grounds. This strategy is consistent with a global constructivism. Epistemic norms are justified on practical terms, and practical norms are justified by the constitutive norms of agency. In the final section, I will develop this option. Another possibility would be to say that there is

some non-constructed reason for the authority of epistemic norms (But this is for the constructivist to admit defeat).

The last possibility that I would like to explore is the possibility of just not answering the Schmagency Objection. What if we tried to live with it? Here, I will argue that the problem with this approach is that it gives *no normativity whatsoever*. Even if I have an attitude, and I cannot help but follow the norms of the attitude, I can still question whether I *ought* to regulate my attitude according to those norms.

Consider the case of Sam's Paranoia.[15] Suppose Sam is clinically diagnosed with paranoid thinking. Suppose his particular ideation is that the postal service is spying on him. Paranoid thoughts are classified by an internal standard of correctness. In order to count as a paranoid thought, the thought involves a certain psychological and representational response to external stimuli. The more the thought embodies this response, the better it is *qua* paranoid thought. And we can characterise not only what it is to be a paranoid thought, but also a paranoid thinker. The paranoid thinker engages in an activity of responding to inputs and maintaining the paranoid ideation. The result is that Sam is engaged in the activity of paranoid thinking, and he regulates his thoughts according to the standard of correctness for paranoid thoughts. Sam receives his diagnosis of paranoia in stride, though, and decides to go to a therapist to try to eradicate his paranoid thinking. But suppose a philosopher meets him along the way and they have this exchange:

> *Philosopher*: Wait, Sam! The standard of correctness for paranoid thoughts is that the more paranoid they are the better they are as instances of paranoid thoughts! You ought to be a good paranoid thinker!
>
> *Sam*: Well, if I ought to be a good paranoid thinker, then I just won't be a paranoid thinker!
>
> *Philosopher*: You seem stuck in your paranoia.
>
> *Sam*: Perhaps true, but I want as little paranoia as possible! I can't work, I can't sleep, I can't get anything done. My life is falling apart. I'm going to be a bad paranoid thinker.

Sam is right. Just because Sam has an attitude that is governed by a constitutive norm does not mean that Sam must see himself as bound to

that norm. And this is true, even if Sam cannot eradicate the attitude. Suppose Sam is unable to entirely opt out of paranoid thought. He still ought to be bad *qua* paranoid thinker, rather than try to have paranoid beliefs that fully satisfy their correctness conditions. Even if the agent inescapably possesses the attitude, this does not give her reason to conform to the standard of correctness.

It might be objected that paranoid thinking is not normative for Sam because it does not have a *rational force* behind it. Since paranoia is by nature irrational, it could have no rational force. By contrast, belief has rational force for us. We are required to be believers. In the final section, I will develop this possibility in greater detail when I propose my positive view. But here, I only want to note how such a strategy undermines Belief Constitutivism. Suppose the authority of epistemic normativity comes from the rational necessity imposed by the importance of belief for our rationality. If this is true, we no longer need to talk about Belief Constitutivism. Instead, we could just talk about how belief is constitutive of rationality, and rationality is (of course!) rationally required of us.

The external question must be answered. As agents, we find ourselves engaged in a multitude of activities. Some activities are ones we should engage in, some are not. The constructivist project is to show that all normative questions are settled within the practice of agency. Normative questions arise *for agents*. Normative questions are answered *by agents*. And so it makes sense that if we are to be truly constructivist about epistemic normativity, we should not locate epistemic normativity in belief, or in believing, but in agency itself. It is this version of epistemic constructivism that I will defend in the final section. This version of epistemic constructivism has the same resources to answer the Schmagency Objection as its moral/practical counterparts.

4 Epistemic Constructivism Defended

In this last section, I will explore the possibility that epistemic norms, like practical and moral norms, are constitutive of agency. If this is true, then the constructivist can argue that the Epistemic Schmagency question is

unintelligible. Thus, my Epistemic Constructivism is on at least as firm setting *vis-à-vis* the Schmagency Objection as the constructivism considered in Sect. 2. This is a reason to prefer my Epistemic Constructivism to Belief Constitutivism.

Before presenting my view, I will contrast it to a popular alternative. Some have thought that the relation between believing and agency is an instrumental one (Kornblith 1993). The practice of agency generates a reason for me to be a believer, and this is because believing truly is necessary to successful action. On this view, being a believer is *necessary* but not *constitutive* of agency. Epistemic norms are distinct from the requirements of agency. We can formulate the norms of agency without making any reference to the epistemic norms. The epistemic norms are necessary like eating a hearty breakfast is necessary for doing well on an exam. The hearty breakfast makes possible doing well on the exam, but eating a hearty breakfast isn't part of taking an exam.

This response could defuse the Schmagency Question. The external question is intelligible, and it is answerable. We can tell Alex that she should be a believer because it is instrumental to bringing about her ends. We have given a conditional constructivism: we have reason to engage in the practice of being a believer. And if the constructivist project is successful, the reason to engage in being a believer can be generated by yet another constructivist norm. This possibility is attractive.

But unfortunately, an instrumentalist defense of epistemic normativity has a fatal flaw. On this conception of agency, successful action is interpreted as "bringing about the intended state of affairs." And beliefs that conform to epistemic norms are often *not* necessary to successful action. Think back to the schmeliever who indulged in optimism bias. The reason to indulge in optimism bias was that it made the agent happier and more successful. So if the question is to be a believer rather than a schmeliever, we can't defend being a believer on instrumentalist grounds. If our beliefs are drastically and pervasively false, we can see that successful action is unlikely. But the converse is not true. Having beliefs that systematically follow the epistemic norms need not make successful action more likely. There may be other norms that are better suited for bringing about successful action. And so while we may have instrumental reason to sometimes and in some ways engage in the practice of being a

believer, this approach can only ground a contingent and limited reason to be a believer.[16]

In contrast to an instrumentalist defense, I propose that we should view epistemic norms as among the norms that are constitutive of action. On this conception of agency, action is not understood as merely producing a state of affairs. Following Korsgaard, I understand an action as "[a]n essentially intelligible object that *embodies* its reason."[17] (2008, 228) The complete description of an action includes three things. It includes the thing to be brought about (the *act*), the *end* which the act serves, and the *reasons* the agent takes to show that the act serves the end. Korsgaard understands these reasons to be captured by the agent's relevant beliefs. So a complete description of the agent's action implicates not only the act, and the end, but crucially, the contents of the beliefs that are relevant to making intelligible what she is doing.[18] Korsgaard writes that, "[g]iving a description or explication of the action, and giving a description or explication of the reason, are the same thing… The reason for an action is not something that stands behind it and makes you want to do it: it is the action itself, described in a way that makes it intelligible." (2008, 227) On this conception of action, we cannot characterise the action without making reference to the agent's beliefs. Contra the instrumentalist picture, beliefs do not *bring about* actions. They are partially constitutive of the action.

On this version of epistemic constructivism, the relationship of belief to action is like paint to a painting. The painting is not understood as *merely* globs of paint—the creative imagination of the artist gives form to the painting—but it *is* globs of paint. We could not display the painting without also displaying the paint. When an agent acts, her movement is attributable to her as an action; it is purposive and guided by her representation of the world. Our understanding of an action is intelligible only if we see it as the agent engaging in movement that is *guided* by her representation of the world. So any description of what the agent is doing will have to include what the agent took to be true when she acted. It is part of what makes the action the action it is.

Epistemic Constructivism grounds the normative status of being a believer within the activity of agency. Believing—the activity that is governed by certain norms—is constitutive of agency. And believing well is

constitutive of acting well. Here, I cannot give a full defence of this view, but I will sketch the outline. An agent is a creature who sets and pursues ends. In order to set and pursue ends, the agent requires a representational capacity to represent the world as it is. An agent cannot set an end of changing the world without having some representation of how the world already is. And this representation will have built into it certain correctness conditions, namely that it accurately represents the world. In order to reason about how to bring about her ends, the agent will need to have a set of background beliefs about how the world works, what means are effective for accomplishing ends, and so forth. And finally, because the agent exists through time, the agent will need to dynamically regulate her representational capacity, gathering information, updating beliefs, dissolving beliefs that do not meet the epistemic norms.

I want to contrast my Epistemic Constructivism with what we discussed in Sect. 3. On my version, the source of epistemic normativity is the agent herself. It is not primarily the belief, or even being a believer. The agent requires that her capacity for representing the world be regulated according to certain norms, epistemic norms. So her agency requires her to be a believer. Her agency also generates the epistemic norms that are authoritative for her beliefs. And so the question, "why should I be a believer?" can be met with the answer, "because you are an agent, and this is what agency requires!" The norms of agency settle that one should be a believer, and manage one's beliefs in a particular way.

On my version of Epistemic Constructivism, the Epistemic Schmagency Objection does not arise. If we understand belief as something that *produces* successful action, we can always ask whether another set of norms will do better. This is why the optimism bias of the schmeliever presented such a compelling challenge. But if the agent's beliefs are playing a different role, a role of making intelligible what the agent is doing by illuminating what the agent takes to be true of the world, what the agent seeks to accomplish and for what end, then nothing else could fulfill that role besides the agent's beliefs. While we can have debates *within* agency over how and when epistemic requirements on belief may be overridden in a particular case, we cannot formulate a distinct Epistemic Schmagency Objection for my Epistemic Constructivism. The question of "why be a

believer?" is answered, "because you are an agent, and this is the role that belief plays for agents."

There is still the regular old Schmagency Objection for constructivism. As I argued in Sect. 2, I think the constructivist has a response to this (though of course more can be said). But, at least for my Epistemic Constructivism, there is no unique Epistemic Schmagency Objection.

5 Conclusion

In this paper, I examined the Schmagency Objection and sketched the general strategy that the constructivist must take to defuse the objection. I then developed a version of the Schmagency Objection for a popular version of epistemic constructivism, Belief Constitutivism. Finally, I motivated a different version of epistemic constructivism, one that locates epistemic norms as being (partially) constitutive of agency. I have not offered a full defence of this account, and more needs to be said in order to establish that epistemic norms *are* constitutive of agency. But I have raised a plausible version of epistemic constructivism, one that avoids the pitfalls of Belief Constitutivism. A fuller defence will have to wait for another time.[19]

Notes

1. In this paper, following Enoch, I am calling constitutivist-constructivist views constructivist. Whether non-constitutivist-constructivist views (alá Sharon Street) face a similar question is outside the focus of this paper. See Street (2010).
2. Enoch writes, "I conclude… that normativity will not come from what is constitutive of action. (Nor, I think, will epistemic normativity come from what is constitutive of beliefs, for reasons similar to the ones I put forward in what follows)" (2006, 171).
3. I am glossing over some of the details of Enoch's Schmagency Objection. There are two complexities that are worth noting. First, Enoch worries about indeterminacy over *which* normative practice we are discussing.

Second, Enoch phrases the question in terms of whether the agent/schmagent should *care*. But the constructivist is interested in whether the norms are *authoritative*, not whether the agent *cares*. While these nuances are interesting, they are not important for this paper.

4. Velleman writes that until we specify the constitutive standards to which we are appealing, we haven't asked a determinate question. He writes, "What the Kantian argues is that the criterion in relation to which guidance is possible must lie in the very nature of that which is to be guided. A question must establish criteria for what can count as a correct answer; if it fails to establish criteria for an answer, then it is not a fully constituted question. If 'Why be an agent?' isn't about a choice or a shmoice or any third thing for which there is a criterion of correctness, then you aren't owed an answer, because you haven't yet asked a question" (2009, 144).
5. This point becomes especially clear in Enoch's follow-up paper Enoch (2011).
6. Enoch overstates the grip of this question, though. In general, we do not need a reason to engage in a practice. We might decide whether to play chess or schmess by flipping a coin. The rules of the road dictate that I should drive on the right side, so I have reason to drive on the right side. But suppose I find myself driving, and I ask myself "why am I driving?" and I can think of no reason why I am. I still ought to conform to the rules of the road, even if I cannot answer the request for justification for why I am driving. Similarly, I could take up the hobby of croquet on a whim. I may not have a reason for doing so (other than my whim), but I still have reason to follow the rules of the game. So, dialectically, if I find myself in a practice, and I cannot justify why I am in the practice, so long as I stay in the activity, I still have reason to follow the rules of the practice.
7. To be clear, the presupposition failure occurs *by the constructivist's lights*. So while the constructivist can reject the objection as question-begging, the result is impasse, not triumph.
8. Unless schmagency is really just a particular conception of agency, as Ferrero (2009) argues.
9. Why *belief* rather than *believing*? Shah and Velleman came to a consensus in their 2005 paper, "Doxastic Deliberation." Velleman (2000) put forward the view that believing has a constitutive aim, namely truth, and this aim structures our believing. Shah (2003) responded with the counterexample of wishful thinking. Wishful thoughts should be classified as

faulty beliefs, but the Velleman (2000) model would classify them as not beliefs at all, since they are not governed by the constitutive aim of believing. Shah and Velleman (2005) propose that we should instead locate the normativity in *belief*. It's important to note that this move wasn't forced by the considerations of Shah (2003), but was rather motivated by considering a number of factors.

10. In addition to Shah and Velleman, other philosophers endorse SOC, though not necessarily as constructivists. See Wedgwood (2002). While Wedgwood's version is consistent with constructivism, he himself is a realist, see Wedgwood (2007).
11. The Belief Constitutivist might respond, but beliefs are not under our direct voluntary control. As a result, we can never get the reflective distance required for this to be possible for us. See my "Evidentialism in Action" (n.d.) for a detailed response to this concern.
12. Thomson distinguishes between the notion of internal-correct and external-correct. Internal-correct refers to an enterprise that can be done well or poorly. External-correct refers to the satisfaction of a standard of correctness. Thomson objects that SOC cannot pick out an enterprise because there is no such thing as internal-correct, only external-correct. There is nothing that counts as well or poorly following the SOC. Instead, SOC is just a standard used to evaluate beliefs.
13. See Hazlett (2013) for an overview of the psychological literature on optimism bias.
14. See Hazlett (2013) for an argument that, in fact, we shouldn't be good believers.
15. This case is structurally similar to Enoch's Reluctant Patriot. See Enoch (2011, 216).
16. A similar point is made by Côté-Bouchard (2016).
17. In Korsgaard (2008), she develops this notion in chap. 7, "Acting for a Reason." In her understanding of action, Korsgaard is drawing on Kant and Aristotle.
18. Crucially, Korsgaard does not think that the beliefs function as reasons that *cause* the action. The beliefs function as reasons-for-performing-this-act-for-this-end, and the agent's reasoning is embodied in the action.
19. I am grateful to Cory Davia, Nick Leonard, Nadja El Kassar, Andrea Robitzsch, Meica Magnani, Luis Rosa, Jakob Ohlhurst, Lisa Benossi, Thomas Grundmann, Sandy Goldberg, Baron Reed, Kyla Ebels-Duggan, and Jennifer Lackey for feedback on various stages of this paper.

References

Côté-Bouchard, Charles. 2016. Can the Aim of Belief Ground Epistemic Normativity? *Philosophical Studies* 173 (12): 3181–3198.

Enoch, David. 2006. Agency, Shmagency: Why Normativity Won't Come from What is Constitutive of Action. *Philosophical Review* 115 (2): 169–198.

———. 2011. Shmagency Revisited. In *New Waves in Metaethics*, ed. Michael Brady, 208–233. Basingstoke: Palgrave Macmillan.

Ferrero, Luca. 2009. Constitutivism and the Inescapability of Agency. *Oxford Studies in Metaethics* 4: 303–333.

Flowerree, A.K. n.d. Evidentialism in Action.

Hazlett, Allan. 2013. *A Luxury of the Understanding: On the Value of True Belief*. Oxford: Oxford University Press.

Kornblith, Hilary. 1993. Epistemic Normativity. *Synthese* 94 (3): 357–376.

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

———. 2008. *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*. Oxford: Oxford University Press.

Rawls, John. 1955. Two Concepts of Rules. *Philosophical Review* 64 (1): 3–32.

Shah, Nishi. 2003. How Truth Governs Belief. *Philosophical Review* 112 (4): 447–482.

Shah, Nishi, and David Velleman. 2005. Doxastic Deliberation. *Philosophical Review* 114 (4): 497–534.

Street, Sharon. 2010. What is Constructivism in Ethics and Metaethics? *Philosophy Compass* 5 (5): 363–384.

Thomson, Judith Jarvis. 2008. *Normativity*. Chicago: Open Court.

Vanderbilt, Amy. 1952. *Amy Vanderbilt's Etiquette*. New York: Doubleday and Co. Inc.

Velleman, David. 2000. *The Possibility of Practical Reason*. Oxford: Oxford University Press.

———. 2009. *How We Get Along*. Cambridge: Cambridge University Press.

Wedgwood, Ralph. 2002. The Aim of Belief. *Philosophical Perspectives* 36 (s16): 267–297.

———. 2007. *The Nature of Normativity*. Oxford: Oxford University Press.

Wright, Crispin. 1992. *Truth and Objectivity*. Cambridge, MA: Harvard University Press.

Index[1]

A

Anti-realism
  epistemic anti-realism, 1, 2, 4, 9, 84, 91n23, 91n25
  moral anti-realism, 4
Assertion, 3, 8, 9, 24, 79–81, 90n18, 187, 188, 197–203, 205–207, 209–211, 213, 213n1, 214n3, 278

B

Belief
  aim of belief, 8, 198, 202, 213, 290
  normativism about, 26, 85, 248, 251
Berker, Selim, 7, 122–126, 128, 129, 132–139, 140n4–7

C

Conee, Earl, 7, 97, 99–101, 103, 106, 107, 117n3, 117n4
Constitutivism, 1–10, 173–193, 290, 297, 298, 300, 301, 303, 304, 307
Contextualism, 226, 229–232, 237
Craig, Edward, 217, 219–223, 225–228, 230, 232, 233
Cuneo, Terence, 4–7, 11n10, 11n11, 40n10, 40n14, 41n25, 46, 59–66, 68n12, 69n17, 71–77, 83, 88n2, 88n3, 88n5, 89n7, 89n8, 89n12, 90n14, 91n21, 91n22, 91n25, 91n26, 149–151, 154, 158, 161, 163, 167n2, 236, 260n4

[1] Note: Page numbers followed by 'n' refer to notes.

© The Author(s) 2018
C. Kyriacou, R. McKenna (eds.), *Metaepistemology*, Palgrave Innovations in Philosophy, https://doi.org/10.1007/978-3-319-93369-6

E
Enoch, David, 10, 90n17, 92n30, 165, 289, 290, 292, 293, 295, 307n1, 307n2, 308n3, 308n5, 308n6, 309n15
Epistemic akrasia, 181–183, 194n4
Epistemic consequentialism, 5, 7, 121–140
Epistemic constructivism, 289, 290, 303–307
Epistemic normativity, 2, 3, 8–10, 56, 62, 84, 91n23, 103, 104, 154–164, 166, 173–175, 194n2, 213, 289, 290, 297, 299, 301, 303, 304, 306, 307n2
  constitutivism about, 8, 173–193, 297, 298
Epistemic reductionism, 5, 6, 45–66
Error theory
  epistemic error theory, 8, 147–150, 154, 157, 163
  moral error theory, 5, 8, 56, 147–150, 157, 163, 166
Evidentialism, 97–101
Expressivism, 2, 4–6, 17, 18, 20, 22–35, 38, 40n10, 40n14, 40n18, 40n19, 41n21, 41n27, 248, 253, 257, 262n18
  epistemic expressivism, 9, 11n5, 17, 22, 36, 37, 39n3, 241–260

F
Feldman, Richard, 7, 60–62, 91n23, 97, 99–101, 103, 106, 107, 117n3, 117n4
Fixed points
  epistemic fixed points, 6, 7, 71–88
  moral fixed points, 6, 7, 71–88

G
Genealogy, 9, 76, 84–86, 91n24, 91n25, 217–237
Gibbard, Allan, 4, 23, 26, 27, 30–34, 36, 37, 40n11, 40n16, 41n21, 41n26, 41n27, 42n28, 249–251, 261n11, 261n12

H
Hinge propositions, 10, 268, 269
Hinges, 10, 25, 89n7, 90n13, 199, 267–285

J
Justification, 4, 6, 7, 67n10, 68n12, 74, 78, 88n3, 97–101, 103–117, 117n2, 118n14, 118n15, 139, 167n3, 218, 244, 261n14, 267, 268, 270, 272–275, 286n9, 294, 308n6

K
Knowledge, 2–5, 8, 33, 38, 76, 78–81, 83–87, 89n8, 90n15, 90n17, 90n18, 109, 147–166, 194n4, 199, 202, 203, 213, 217–223, 229, 231, 232, 234, 237, 261n14, 267, 270, 272–275
  factivity of knowledge, 79, 80, 86, 90n15
Korsgaard, Christine, 11n8, 140n1, 214n8, 291, 293, 296, 298, 305, 309n17, 309n18
Kyriacou, Christos, 5–7, 11n1, 11n5, 11n10–12, 46, 59–66, 68n12, 69n17, 88n1, 88n4,

89n10, 90n14, 91n22, 91n25, 91n26, 92n29, 103, 194n1, 194n6, 260n4, 261n14

M
Mentalism, 7, 97–117
Minimalism, 281, 283, 284
Moore, G.E., 5, 67n8
Moore's paradox, 176, 188, 193
Moral/epistemic parity principle, 149–151
Moral non-naturalism, 89n7, 89n11
Motivational internalism, 98, 183, 185

N
Normativity
  epistemic normativity, 2, 3, 8–10, 56, 62, 84, 91n23, 103, 104, 154–164, 166, 173–193, 213, 289, 290, 297–299, 301, 303, 304, 306, 307n2
  robust normativity, 153, 154, 157, 158, 163, 165, 168n7

P
Pluralism, 10, 218, 226, 228, 277, 282, 286n12

R
Rationality, 4, 62, 84, 105, 107–111, 115, 118n12, 133, 134, 137–139, 152–154, 232, 244, 261n15, 270, 279, 280, 283–285, 285n2, 303
Realism
  epistemic realism, 1, 2, 6, 7, 9, 72, 78, 84, 86, 88, 148, 150, 151, 153, 154
  moral realism, 4, 6, 46, 47, 56, 66, 67n6, 73, 75, 76, 83, 151, 153
  normative realism, 4
Reasons
  epistemic reasons, 33, 53, 56, 57, 82, 90n19, 91n21, 97, 100, 101, 103–107, 109, 115–116, 134, 135, 137, 139, 149, 156, 157, 159, 160, 162, 163, 168n7, 174, 213, 244, 258, 259
  reasons primitivism, 9, 241–260
Relativism, 2, 4, 9, 217–237, 279, 283
  epistemic relativism, 9, 168n6, 217, 218
Representation, 10, 18, 19, 23, 24, 26, 28–30, 32, 38, 40–41n19, 198–203, 206–209, 213, 214n5, 257, 271, 275, 279, 280, 282–285, 302, 305, 306
Reynolds, Steven, 217, 222, 223, 225, 226, 228, 230, 231

S
Schmagency, 10, 289–307
Shafer-Landau, Russ, 6, 7, 41n25, 67n6, 71–77, 83, 88n2, 88n3, 88n5, 89n7, 89n8, 89n12, 92n28
Skepticism, 85

Smith, Michael, 9, 180–183, 242–247, 249, 253, 255, 256, 259

T

Trade-offs, 7, 122–125, 128–134, 137–139, 141n12
Transparency, 8, 177, 178, 182, 183, 185, 186, 191
  of deliberation, 175–186, 190–191, 193, 194n4, 194n5
Truth
  truth norms, 174–193, 194n2, 194n5, 194n6
  value of truth, 122, 164, 197, 198, 204, 213n1

V

Veritism, 122, 123, 125, 126, 128, 132, 138, 140n5, 141n10, 141n11

W

Williams, Bernard, 188, 197, 198, 212, 213, 217, 222–226, 228, 230, 233, 234, 236
Wittgenstein, Ludwig, 199, 219, 267–269, 277–279, 281, 283, 285n1, 285n3, 286n11, 286n14
Wright, Crispin, 271, 273–275, 277, 281, 282, 286n9, 298