# *Reasons for* BELIEF

EDITED BY

Andrew Reisner and
Asbjørn Steglich-Petersen

CAMBRIDGE

CAMBRIDGE www.cambridge.org/9781107006874

This page intentionally left blank

# REASONS FOR BELIEF

Philosophers have long been concerned about what we know and how we know it. Increasingly, however, a related question has gained prominence in philosophical discussion: what should we believe and why? This volume brings together twelve new essays that address different aspects of this question. The essays examine foundational questions about reasons for belief and use new research on reasons for belief to address traditional epistemological concerns such as knowledge, justification, and perceptually acquired beliefs. This book will be of interest to philosophers working on epistemology, theoretical reason, rationality, perception, and ethics. It will also be of interest to cognitive scientists and psychologists who wish to gain deeper insight into normative questions about belief and knowledge.

ANDREW REISNER is Assistant Professor of Philosophy at McGill University. He has published a number of articles on theoretical reason and normativity.

ASBJØRN STEGLICH-PETERSEN is Associate Professor of Philosophy at Aarhus University. He has published widely in epistemology and metaphysics, and is editor of *Metaphysics: 5 Questions* (2010).

# REASONS FOR BELIEF

EDITED BY

ANDREW REISNER AND
ASBJØRN STEGLICH-PETERSEN

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town,
Singapore, São Paulo, Delhi, Tokyo, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9781107006874



First published 2011

Printed in the United Kingdom at the University Press, Cambridge

*A catalogue record for this publication is available from the British Library*

*Library of Congress Cataloguing in Publication data*
Reasons for belief / edited by Andrew Reisner and Asbjørn Steglich-Petersen.
p. cm.
Includes bibliographical references and index.
ISBN 978-1-107-00687-4
1. Belief and doubt. I. Reisner, Andrew Evan, 1975– II. Steglich-Petersen, Asbjørn. III. Title.
BD215.R39 2011
121′.6 – dc22 2011006849

ISBN 978-1-107-00687-4 Hardback

# *Contents*

*Contributors* *page* vii
*Preface* viii

Introduction 1
*Andrew Reisner and Asbjørn Steglich-Petersen*

PART I NORMATIVE REASONS FOR BELIEF 11

1 How to be a teleologist about epistemic reasons 13
*Asbjørn Steglich-Petersen*

2 Is there reason to be theoretically rational? 34
*Andrew Reisner*

3 Epistemic motivation: towards a metaethics of belief 54
*Veli Mitova*

4 Error theory and reasons for belief 75
*Jonas Olson*

5 Can reasons for belief be debunked? 94
*Nishi Shah*

PART II REASONS AND EPISTEMIC JUSTIFICATION 109

6 Reasons and belief's justification 111
*Clayton Littlejohn*

7 Perception, generality, and reasons 131
*Hannah Ginsborg*

8 Immediate warrant, epistemic responsibility, and Moorean dogmatism 158
*Adam Leite*

9 Primitively rational belief-forming processes 180
*Ralph Wedgwood*

10 What does it take to "have" a reason? 201
*Mark Schroeder*

11 Knowledge and reasons for belief 223
*Alan Millar*

12 What is the swamping problem? 244
*Duncan Pritchard*

*References* 260
*Index* 271

# *Contributors*

HANNAH GINSBORG, Professor of Philosophy, University of California, Berkeley

ADAM LEITE, Associate Professor of Philosophy, University of Indiana, Bloomington

CLAYTON LITTLEJOHN, Assistant Professor of Philosophy, University of Texas, San Antonio

ALAN MILLAR, Professor of Philosophy, University of Stirling

VELI MITOVA, Postdoctoral Research Fellow, Department of Philosophy, University of Vienna

JONAS OLSON, Docent and Research Fellow, Department of Philosophy, University of Stockholm

DUNCAN PRITCHARD, Professor of Philosophy, University of Edinburgh

ANDREW REISNER, Assistant Professor of Philosophy, McGill University

MARK SCHROEDER, Associate Professor of Philosophy, University of Southern California

NISHI SHAH, Associate Professor of Philosophy, Amherst College

ASBJØRN STEGLICH-PETERSEN, Associate Professor of Philosophy, University of Aarhus

RALPH WEDGWOOD, Professor of Philosophy, University of Oxford

# *Preface*

This book contains a collection of new papers on the topic of *reasons for belief*, sometimes referred to in the literature as 'theoretical reasons'. The papers in this volume address two broad themes: the nature of reasons for belief and the application of reasons for belief to other traditional problems in epistemology. The choice of these two themes reflects the rationale for putting together this volume.

For a period of roughly 35 years, practical philosophers have been appealing to *reasons* to do ever more work in their theorizing. Early debates in practical philosophy posed the question of whether one could have normative reasons for action – considerations that count in favour of an agent's performing a particular action or actions – that were disconnected from an agent's own motivations. Soon the value of thinking of broader problems in ethics in terms of reasons became apparent, and now appeals to reasons populate debates across normative ethics and metaethics alike. Reasons became a kind of common currency for consideration of the normative issues in practical philosophy.

Like those in ethics and other areas of practical philosophy, many of the problems studied in epistemology are also normative. Epistemologists have traditionally examined these problems through discussions of *justification* and *warrant*. Increasingly, however, philosophers interested in the problems of normative epistemology have appealed to reasons both to help explicate justification, warrant, and related concepts, and to address independently other concerns in epistemology. That this has occurred is not surprising; concentrating on reasons has proved valuable to moral philosophers in no small part because reasons provide a common normative currency, and the potential benefits of doing so for epistemology appear no smaller.

Because of the growth in work on reasons for belief, there are now many strands of research in epistemology that rely on using reasons. This volume brings together papers covering a broad swathe of these strands. Part of the aim in doing so is to allow someone new to the subject, or someone whose

research has focused closely on just one or two of those strands, to take advantage of the common currency aspect of reasons. Developments in foundations of reasons for belief, for example, may be helpful to someone who is interested in how to use the concept of reasons for belief to better understand perceptual warrant. Bringing together papers on a wide variety of aspects of reasons for belief fulfils another aim of the volume, which is to provide a single source for scholars interested in reasons for belief to see the state of the art.

A collected volume is only as good as its papers, and in this respect we are happy to have an excellent volume. As editors, we are especially grateful to our contributors both for their contributions and for their patience with the volume's coming to fruition. In addition to the contributors, we owe special thanks to our editor at Cambridge University Press, Hilary Gaskin, and our production editor, Jodie Barnes. They have made the publication process as easy as it can be. The manuscript was substantially improved by a copy-editor, Jo North. Robert Stephens and Oran Magal have also been invaluable throughout the entire process of assembling this volume.

ANDREW REISNER
ASBJØRN STEGLICH-PETERSEN

# *Introduction*

*Andrew Reisner and Asbjørn Steglich-Petersen*

Over the last three decades, practical philosophy has increasingly looked at, and become dependent upon, the concept of *normative reasons* for actions, and action-related propositional attitudes. The concept gradually came into prominence in a series of classic treatments in the late seventies and early eighties and has since then become the focal point and organizing concept for a vast array of work in both ethics and the philosophy of mind and action. The core of the concept is a simple one: normative reasons are facts that *count in favor* of some action or attitude; they are the facts that determine whether or not an agent *ought* to do something, or adopt some attitude. As such, normative reasons are often thought to be the most fundamental concept relevant to understanding *rationality*, which, on this view, is the capacity to recognize and respond to reasons in appropriate ways. The link to rationality means that normative reasons not only determine what ought to be done; at least sometimes, they also play the role of explaining why an agent in fact acted or thought as she did.

Even if there is broad agreement among philosophers on these fundamental features of reasons, a detailed understanding of reasons is still subject to controversy. What kinds of facts can act as reasons? Are they restricted to representational states such as beliefs or desires, or can nonmental states of affairs act as reasons as well? Do reasons for action somehow depend on the value of taking that action, or can value itself be explained in terms of reasons? Are there different kinds of reasons, and how are these reasons then related? Can different reasons be weighed against each other? How do reasons enter into reasoning about what one ought to do? What is the logical form of reasons statements? What is the relationship between normative and explanatory reasons? These are all issues on which philosophers disagree. Nevertheless, there is some general agreement that many problems in practical philosophy are fruitfully addressed by asking what

there is reason to do, desire, or intend, and the enormous interest that the above issues have generated is testament to this fact.[1]

Traditionally, in theoretical philosophy, epistemologists have looked to the concept of *epistemic justification* as their central normative notion, and to *warrant* as a crucial subsidiary one. Important debates on whether justification depends on properties that are internal or external to the epistemic agent, what role justified belief plays in knowledge, and whether justification is context-sensitive, have, amongst many others, taken centre stage in epistemology. Although the concerns of epistemologists are somewhat different from those of practical philosophers, at least superficially, the notion of epistemic justification seems to be closely related to that of normative reasons. Epistemic justification is supposed to speak in favor of adopting the relevant belief, and, at least sometimes, the epistemic justification a person has for a given belief explains why that belief was adopted. The similarity is also reflected by the fact that epistemologists sometimes speak of the *justification* and the *reasons* a person has for her belief interchangeably. Nevertheless, debates over epistemic justification and normative reasons have largely been conducted in isolation from each other.

This raises at least two important questions. First, it raises the question of how we are to understand reasons for beliefs, where "reasons" is understood in the normative sense known from practical philosophy. Questions of this sort already have some pedigree in the philosophical debate. For example, many philosophers have been interested in whether there could be "practical" reasons for belief, in the sense of reasons for belief that are not related to the truth of the belief, but instead to some practical advantage, which the belief would accrue to its holder. However, theorizing over normative reasons for belief hasn't yet reached the sophistication of the parallel debate over normative reasons for action, and many important issues remain largely untouched. For example, in what sense does evidence "speak in favor" of adopting a given belief? Are reasons for belief somehow value-based, even if they are truth-related or "epistemic"? Is there some general theory of reasons, which is common to both reasons for action and reasons for belief? If not, what does this show about the unity of normativity and rationality? How do reasons for belief motivate the relevant beliefs? Would it be coherent to deny the existence of normative reasons for belief? These are some of the questions taken up by the chapters in the first part of the volume.

[1] For a recent volume of papers devoted to the subject of normative reasons for *action*, see Sobel and Wall (2009).

The second important question raised by the above considerations is what lessons there are to be learned in regard to the traditional issues in epistemology concerning the epistemic justification of beliefs, by looking to the (at least superficially) related concept of normative reasons. Might there be something important to be learned about the conditions under which one is justified in holding a belief, by looking to theories of what it means to have a normative reason for some action or attitude? It remains a very real possibility, of course, that there is a deep reason why debates over epistemic justification and normative reasons so far have been conducted independently of each other. Perhaps the two notions really *are* too different to be usefully compared; perhaps the superficial similarity between the two concepts is nothing more than exactly that: superficial. The best way to test this is to see whether considerations concerning the one concept can in fact help elucidate the other. The chapters in the second part of the volume do exactly that, and the result, we think, is very helpful indeed.

In the remaining part of the introduction, we provide summaries of the individual chapters of the volume, highlighting along the way how the chapters speak to the two organizing questions of the volume.

## PART I: NORMATIVE REASONS FOR BELIEF

As mentioned, the chapters in the first part of the volume are all devoted to understanding reasons for beliefs, where "reasons" is understood in the normative sense known from practical philosophy. In the first chapter, "How to be a teleologist about epistemic reasons," Asbjørn Steglich-Petersen considers the possibility of explaining the normative source of reasons for belief in terms of the value of the beliefs that they support. According to the popular teleological conception of normative reasons, reasons are value-based in the sense that whether someone has reason to φ depends on the value of the result of φ-ing, or the intrinsic value of φ-ing itself. Many have been attracted to similar accounts of *epistemic* reasons, according to which epistemic reasons depend on the value of the beliefs that they support, but it has proven difficult to make such an account plausible. The central problem is that the epistemic properties in virtue of which epistemic reasons obtain aren't always valuable. Hitherto, most defenses of the teleological account of epistemic reasons have focused on ways in which epistemic properties, despite appearances, might be seen as valuable in a sufficiently general sense to ground the teleological account. Instead of this, Steglich-Petersen pursues the alternative strategy of developing a teleological account of epistemic reasons, which is compatible with the

seeming fact that epistemic properties aren't always valuable. To this end, Steglich-Petersen distinguishes between reasons to believe that particular propositions are true, i.e. epistemic reasons, and reasons to pursue the aim of forming beliefs about those propositions in the first place. Once we realize the difference between these two kinds of reasons, epistemic reasons can be understood as instrumental reasons to pursue the aims, which the second kind of reasons support. A result of this is that epistemic reasons can be understood as value-based, without epistemic properties being valuable in all contexts.

Andrew Reisner's chapter, "Is there reason to be theoretically rational?," explores the relation between reasons for belief and rationality. During the last decade, it has become increasingly fashionable to believe that rationality strongly supervenes on the mental, whereas normative reasons depend on non-mental facts (for the most part). Reisner explores this claim for the particular case of the rationality of belief. The chapter first gives a brief history of how the current view about the distinction between rationality and reasons came to be held. Reisner argues that the distinction is essentially well motivated and that there is a conceptual distinction between a belief's, or collection of beliefs', being rational and there being reason to have that belief or collection of beliefs. Nonetheless, there remains the question of whether its being rational to hold a collection of beliefs in some way entails that there is a reason to hold that collection of beliefs. Reisner argues that its being rational to hold a collection of beliefs does provide a very strong reason to hold that collection of beliefs, at least for rational requirements of a certain kind (those that are wide-scope consistency requirements). Ordinary evidential reasons for belief, it is argued, give us wide-scope reasons not to hold collections of inconsistent beliefs, and these beliefs are forbidden by the requirements of theoretical rationality.

In the section's third chapter, "Epistemic motivation: towards a metaethics of belief," Veli Mitova takes the first steps towards a more systematic extension of traditional metaethical concerns to the ethics of belief. As pointed out by Mitova, three fundamental debates largely define contemporary metaethics. The first is between internalists and externalists about moral motivation. While internalists think that there is a necessary connection between making a moral reasons judgment and being motivated to act in accordance with it, and that such a connection is necessary if we are to explain the manner in which moral reasons can guide and explain actions, externalists deny any such necessary connection. The second main debate is that between Humeans and anti-Humeans about motivation. The issue here is whether desires, or some desire-like states, as the Humeans

hold, are necessary in order to explain the motivation of actions, or, as the anti-Humeans hold, a cognitive state such as a belief can suffice for motivating action on its own. The last main debate concerns the truth-aptness of moral judgments. Cognitivists claim that moral judgments can be true or false, and that to make a moral judgment entails having a belief with that content. Non-cognitivists deny this, and hold instead that such judgments should be understood as the expression of some non-cognitive attitude. The aim of Mitova's chapter is to transpose, in a systematic way, these debates to the ethics of belief. Her guiding observation is that the three debates are interdependent. In particular, commitment to anti-Humeanism and internalism about moral motivation seems to entail a commitment to moral cognitivism. This forms the starting point for Mitova's argument. She first provides a novel argument for a form of anti-Humeanism and internalism about the way in which normative reasons judgments motivate beliefs, and then relies on these arguments to establish a form of cognitivism about such judgments.

In his chapter, "Error theory and reasons for belief," Jonas Olson considers the radical possibility that there simply *are no* reasons for belief and the prospects of an accompanying *error theory* about such reasons. This option has been explored in some detail in the practical domain, motivated mainly by a desire to avoid commitment to "queer" and non-natural normative entities suggested by moral thought. But a parallel position in regard to reasons for belief has often been assumed to be a non-starter. In fact, the perceived implausibility of an error theory about epistemic reasons has been regarded by many as an embarrassment for error theorists in the practical domain. Using a "companion in guilt" strategy, these philosophers have argued that since error theory about epistemic reasons is so implausible, and arguments for error theory in the practical domain seem to entail a commitment to an analogous theory about reasons for belief, we should reject error theory in the practical domain too. In his chapter, Olson considers whether an error theory about reasons for belief would really be so implausible as is often supposed. In particular, Olson discusses what Terence Cuneo (2007) has recently described as three unpalatable results for such a theory, *viz.* that it would be self-defeating or polemically impotent; that it would imply that there can be no arguments for anything; and that it rules out the possibility of epistemic merits and demerits. In his careful discussion, Olson argues that these results either do not follow, or are less unpalatable than Cuneo supposes.

Nishi Shah is more skeptical concerning the possibility of an error theory about reasons for belief. In his chapter, "Can reasons for belief be

debunked?," Shah develops a novel argument against this possibility. Like Cuneo, Shah thinks that an epistemic error theory would be inconsistent or self-undermining. But as opposed to Cuneo's argument, Shah's depends only upon a number of widely accepted assumptions about the nature of belief itself. Shah argues that even the error theorist is committed to the existence of beliefs – after all, the core claim of the epistemic error theorist is that our *judgments* or *beliefs* about epistemic reasons are systematically false, which presupposes that these judgments or beliefs are there to be false in the first place. If statements about epistemic reasons were to be interpreted as mere expressions of conative states, which are neither true nor false, the error theorist's claim would make no sense. But according to the account of belief previously defended by Shah, ascriptions of belief require making normative judgments. In order to ascribe a belief to someone one must judge, at least implicitly, that the mental state one has just classified as a belief is *correct* if and only if the proposition believed is true. If Shah is right, it is simply part of the nature of belief, that beliefs are governed by this norm of correctness. But if the error theorist accepts this, it seems, as Shah argues, that she will also be committed to the truth of at least some true normative judgments about reasons for belief. Hence, the error theory about such reasons fails.

## PART II: REASONS AND EPISTEMIC JUSTIFICATION

As mentioned above, the chapters in Part II all explore different ways in which elements from the debate over normative reasons might help elucidate some more traditional epistemological concerns over the justification of beliefs. In the first chapter of Part II, "Reasons and belief's justification," Clayton Littlejohn begins by observing that there is little to say about the justification of beliefs that cannot be said in terms of reasons. Hence, we must be able to work our way from an account of the demands of reasons to an account of epistemic justification. A prominent view on the former issue is that reasons demand *conformance*: if there is reason for an agent to $\varphi$, its demands are met if the agent in fact $\varphi$s, and otherwise not. A common objection is that conforming is not always rational. For this to be the case, the agent must also *comply* with the reason, i.e. $\varphi$ for that very reason. But according to Littlejohn, this additional demand should be replaced by a demand that we exercise due care to avoid acting against a reason and failing to conform to its demands. If this is true, important consequences follow for how we should think of epistemic justification. In particular, the traditional view that a person's justification for a belief

supervenes on the evidence available to the agent, and the account of justification which relies on the assumption that knowledge is the norm for belief, both seem to be in trouble. If Littlejohn's argument is successful, it would thus constitute an important example of the importance of the notion of normative reasons to traditional epistemological concerns.

In her chapter, "Perception, generality, and reasons," Hannah Ginsborg relies in part on the notion of reasons in evaluating a recent trend in the theory of perception. For a long time, philosophers have debated whether the representational content of perceptual states is conceptual or non-conceptual, but recently a new trend has emerged challenging the assumption that perceptual states are representational *at all*. This idea must be rejected, they say, if we are to adequately characterize what is distinctive about perception in contrast to thought and belief. Ginsborg considers the implication of this view for the idea that perceptual experiences can stand in rational or reason-giving relations to belief, and argues that the two ideas are irreconcilable. Denying that perceptual states have representational content implies denying that perception can play the proper reason-giving role in regard to belief, at least in the sense traditionally invoked in epistemological debates over the justification of empirical belief. In making this point, she relies on the distinction familiar from practical philosophy between a reason understood as a consideration counting in favor of an action or attitude (Scanlon 1998), and a reason understood as a belief that stands in a rationalizing relation to actions or attitudes. Ginsborg argues that while the non-representational view may be able to explain perceptual reasons in the first sense, it cannot explain perceptual reasons in the second sense, which is the more interesting one from an epistemological point of view. Instead, Ginsborg outlines a version of the representational theory of perception, which both addresses the worries motivating the non-representational view, and explains the ways in which perception can rationalize thought and belief.

Adam Leite's contribution to the volume, "Immediate warrant, epistemic responsibility, and Moorean dogmatism," also takes up the problem of accounting for the reasons for belief provided us by our sensory experiences. According to the hotly debated "Moorean dogmatist" response to external world skepticism, our sensory experience provides us with *prima facie* immediate justification, warrant, or reason to believe certain propositions about the world. This position has a great deal of intuitive appeal and can easily seem to be exemplified in our ordinary epistemic practice. At the same time, the strategy endorses forms of reasoning or argumentation that many people find objectionable. In particular, the dogmatists appear to

hold that an ideal rational agent who considers the question of whether she is being deceived by an evil demon could start from a position which presupposes no beliefs at all about the world, consciously take her current experience as a reason for believing that she has hands and so believe that she has hands on that basis, reason from that belief to the conclusion that she is not being deceived by an evil demon, and thereby form the latter belief in a fully satisfactory way. To many people this line of reasoning seems objectionable, but it has proven difficult to locate the source of dissatisfaction. In his chapter, Leite seeks to locate this dissatisfaction in considerations about epistemic responsibility. To this end, Leite develops a theory of immediate warrant and shows how it can be combined with plausible "inferential internalist" demands arising from considerations of epistemic responsibility. The resulting view endorses immediate perceptual warrant but forbids the sort of reasoning that Moorean dogmatism would allow. A surprising result of this discussion will be that dogmatism alone isn't enough to avoid standard arguments for skepticism about the external world.

Ralph Wedgwood's chapter, "Primitively rational belief-forming practices," seeks to explore the common view that one can only have a reason to do something if there is a sound process of reasoning that takes one from one's current mental states to the performance of the relevant action. Wedgwood suggests that this same view may be applied to belief, and in particular to belief formation. He explores the topic by asking where there are primitively rational belief-forming practices – practices that, if carried out, make it rational for an agent to form the belief resulting from those practices. Wedgwood puts three constraints on these practices: that they are rational practices that result in rational beliefs; that they are not infallible practices (even if they are reliable); and that they are primitive or basic, i.e. they do not require some practice-independent justification for their rationality. The particular putative primitively rational belief-forming practice that he considers is "taking experience at face value," which is the practice of coming to believe *p* in response to *p*'s being part of one's conscious experience. After developing a more careful account of what it is to take experience at face value, Wedgwood argues that the practice may well be a primitively rational belief-forming practice and that understanding why will help us to understand in general what is required for something to be such a practice. He suggests that the central criterion is that such practices must be, and indeed are, built into our very capacity to possess certain kinds of concepts and to have certain kinds of propositional attitudes, and that they are therefore *a priori* practices.

In Mark Schroeder's chapter, "What does it take to 'have' a reason?," questions about beliefs derived from perception also take a central role, but in this case to raise broader questions about the role evidence plays in inferential justification: in our having reasons for beliefs that we have arrived at by inference. The aim of the chapter is to consider, and then reject, a series of arguments that suggest that the bar for having reasons for beliefs arrived at inferentially is high. Schroeder argues that beliefs inferred from perceptual experiences would be ruled out by accounts that require a high bar for inferentially arrived at beliefs. Accepting such accounts would either rule out the possibility that perception can play an appropriate role in giving us reasons for our beliefs (serving as evidence for those beliefs), or that we need two accounts of justification for inferential beliefs: a low bar account for perceptually grounded beliefs and a higher bar account for other cases. Schroeder argues that we should retain a unified account with a lower bar. If correct, his view appears to have far-reaching consequences, as it offers up resources for resisting both coherentist and radical externalist views about justification within epistemology.

In "Knowledge and reasons for belief," Alan Millar poses a puzzle of a general form for epistemologists: how can we get so much from so little? The exemplar of this puzzle is testimony: it is commonly thought that we can gain knowledge through testimony, even when we have not reasoned carefully about the testifier's epistemological *bona fides*. A central feature of this puzzle is that we are inclined to believe that this ability to gain knowledge suggests an ability to gain justified belief, even without particular commitments concerning whether knowledge can be analyzed in terms of justified belief. Famous challenges to gaining justification on such a minimal basis abound, and it is Millar's project to develop a sketch of how one might show that under suitable circumstances, one might indeed be able to be justified in believing something on apparently less than ideally robust bases. In building the account, the notion of a recognitional ability is developed. That we have the capacity to recognize the way things are under suitable circumstances allows us to accord experience an explanatory role in our acquisition of knowledge and the possession of reasons for belief. Millar reverses the traditional order of explanation by suggesting that experiences are not themselves reasons for believing that things are as the experiences take them to be. Instead, it is what we know about our environment and our perceptual ability to access it that provides the justification for perceptual knowledge. The order of explanation can be reversed by understanding what kind of reasons we get from certain kinds of thin epistemic bases, such as testimony or indicator phenomena in

general: that is, they are reasons that are given in terms of understanding how they derive from our knowledge of the environment and our capacities for accessing that environment. Millar's chapter concludes that we gain important understanding about our knowledge of the world in traditionally problematic cases like those that arise for perceptual knowledge, if we take a knowledge-first approach. This leads to the development of a particular category of knowledge: detached standing knowledge. In such cases, justification does not serve as a basis for knowledge, but rather we can explain why we are justified in believing something because of what we know. Millar links the availability of reasons for belief to an understanding of how we have come to know what we know.

In the final chapter of the volume, "What is the swamping problem?," Duncan Pritchard returns us to another problem in traditional epistemology. If the only valuable feature of a belief is that it is true, then what non-instrumental value do other normative epistemological concepts such as *justified belief*, *rational belief*, and *knowledge* have? Put another way, we can understand the swamping problem as being, or at least implying, a thesis about reasons for belief: our central way of evaluating belief is in terms of whether there are reasons pointing toward a belief's being true. The swamping problem is normally thought to apply only to certain kinds of epistemological views. Pritchard argues that the problem is not constrained to a narrow range of positions, and he gives a more precise account of the swamping problem in terms of three claims: a general thesis about value, a more specific thesis about epistemic value, and the (putative) correctness of a popular thesis in epistemology, called "T-monism" by Pritchard. It is argued that each of these claims is independently plausible, but that they are collectively inconsistent. Having set things out in this way, the options available to epistemologists for solving the swamping problem are discussed. It is conjectured that, from amongst the available options, the most promising way of resolving the swamping problem will require us to adopt a type of value pluralism which casts into doubt the central role that knowledge has played in traditional epistemology.

PART I

# *Normative reasons for belief*

CHAPTER 1

# *How to be a teleologist about epistemic reasons*

*Asbjørn Steglich-Petersen*

## 1.1 INTRODUCTION

According to the teleological conception, reasons are value-based. Whether someone has reason to ϕ depends on the value of the result of ϕ-ing, or the intrinsic value of ϕ-ing itself (Scanlon 1998: 84). Accordingly, on the teleological conception of reasons for belief, whether someone has reason to believe some particular proposition *p* on some particular occasion depends on the value of the result of believing *p*, or the intrinsic value of believing *p*, on that occasion. Although this basic tenet of the teleological account of reasons for belief has seemed plausible to many, and a number of attempts have been made at fleshing it out in detail, it remains controversial whether it can be made to work.[1]

We can distinguish between two distinct challenges, or broad categories of challenges, for the teleological conception of reasons for belief. On the one hand, there is the problem of explaining why, on the teleological account, one couldn't have reason to believe a proposition that is not supported by evidence. If reasons for belief are reasons that obtain in virtue of the value of the relevant belief, why does the pleasantness (or some other non-epistemic property of value) of believing a particular proposition not provide a reason for holding it? There are those, of course, who accept the possibility of such non-epistemic reasons for belief.[2] But many teleologists about reasons for belief wish to maintain that one can have reason to believe a proposition only insofar as the belief is evidentially supported, or has some other epistemic property, depending on the specific account. We can call this 'the exclusivity problem' for teleologists. The problem is that of explaining why only certain properties of value, namely the epistemic

[1] For prominent examples, see Alston (2005), BonJour (1985), Foley (1987), Goldman (1999b), Lehrer (1990), Lynch (2004).

[2] See Reisner (2008 and 2009a) for a recent discussion and defence of the idea of non-epistemic reasons for belief.

ones, can count in favour of beliefs, if reasons are generally a matter of the relevant actions or attitudes resulting in or promoting something of value. I have argued in previous papers that a teleological conception of reasons for belief is compatible with there only being epistemic reasons for belief, and I shall simply assume in the following that this is the case.[3]

The second major challenge for the teleological account concerns the nature of *epistemic* reasons specifically, regardless of whether such reasons are somehow privileged in supporting beliefs. The problem is that of explaining how epistemic reasons could be value-based in the first place. If epistemic reasons for belief are reasons that obtain in virtue of the epistemic properties of the beliefs that they support, a teleological conception must explain epistemic reasons in terms of the value, whether intrinsic or derived, of those epistemic properties. But as witnessed by the latest decades of theorizing about this question, it is all but clear how or why epistemic properties are valuable or conducive to something of value, in a sense that is sufficiently general to form the basis of a teleological account of epistemic reasons.[4] This is the challenge I address in this chapter. In the following, I shall call it 'the value problem' for teleological accounts of epistemic reasons.[5]

The most common approach to the value problem accepts the basic premise of the challenge, namely that if the teleological account is to succeed, epistemic properties must somehow be of value, or be conducive to something of value, in some *general* sense. Accordingly, the common approach attempts to solve the problem by locating the epistemic properties of value, and explaining the sense in which they are valuable. A candidate account might hold, for example, that true beliefs are of intrinsic value, and that any property of beliefs that ensures or makes it likely that a belief is true (such as the property of being evidentially supported or being the result of a reliable mechanism) therefore is valuable as well.[6] However, in spite of intensive efforts, this strategy has not yet reached a widely accepted solution.

In this chapter, I explore the prospects of a slightly different strategy in solving the value problem for the teleological account of epistemic reasons.

[3] See Steglich-Petersen (2006a, 2006b, and 2009).

[4] For a recent discussion of representative proponents of this line of explanation, see Grimm (2009).

[5] The value problem for a teleological account of epistemic reasons is distinct from several other so-called 'value problems' in epistemology, most notably the problem of explaining the value of knowledge as opposed to mere true belief. For a survey of recent work on this problem, see Pritchard (2007b).

[6] For this particular account, see Lynch (2004).

The starting point of the strategy is to concede the main criticism of teleological accounts of epistemic reasons, namely that epistemic properties (whatever the relevant epistemic properties might be) are *not* valuable generally and in all contexts, and that a teleological account relying on such a claim therefore fails. So the question becomes whether a teleological explanation of epistemic reasons which doesn't rely on that claim is available. I will introduce and motivate such an account.

At the core of the account lies a distinction between two different kinds of reasons concerning beliefs, namely reasons to form beliefs *that* particular propositions are true, i.e. epistemic reasons, and reasons to form beliefs *about* certain propositions or subject matters. I shall argue that the latter kind of reasons is connected to value in a straightforward way: whether one has reason to form beliefs *about* some proposition or subject matter is determined by the value of doing so, in the same way that reasons for actions seem determined by the intrinsic or derived value of those actions. I shall argue further that the two kinds of reasons stand in a certain systematic relationship to each other that any theory of those reasons must explain. To anticipate, epistemic reasons to believe *that p* entail that one ought to believe that *p* only in the context of an all-things-considered reason to form a belief *about p*. I then go on to provide a teleological account of this relationship. I shall argue that the relationship can be explained by supposing that having an epistemic reason to believe that *p* is equivalent to having what I term a 'hypothetical instrumental reason'. The main virtue of the account is that it does not require the epistemic properties underlying epistemic reasons to be of value in all contexts. Only if one has reason to form beliefs about the relevant subject matter do epistemic properties of some potential belief about that subject matter become valuable, and since one can have epistemic reason to believe some proposition without having reason to form beliefs about that proposition, epistemic reasons can obtain without the epistemic properties in question being of value.

## 1.2 DEFINITIONS

To fix matters, it will be useful to begin by clarifying what I shall mean by a number of central notions. First of all, we need to clarify the notion of a 'reason'. By a 'reason', I shall mean a fact or consideration that stands in a reason-giving relation to some particular action or attitude of some particular agent (Scanlon 1998: 17). The reason-giving relation can take different forms, depending on the kind of reason. Generally, the relation is that of *favouring* the relevant action or attitude, so a reason to $\phi$ is

something that somehow favours ϕ-ing. But there are different kinds or 'grades' of favouring. Some reasons may by themselves *require* ϕ-ing, i.e. make it the case that one ought to ϕ. Such reasons are sometimes called 'perfect reasons' (Broome 2004), but I shall instead call them 'all-things-considered reasons'. Other reasons may favour ϕ-ing in a mere 'pro tanto' way. Such reasons can usefully be thought of as considerations that can play a role in a weighing-explanation of what one ought to do, where the weights of several pro tanto reasons are compared to reach a judgement about what one overall ought to do in a given situation (Broome 2004). So having a pro tanto reason to ϕ does not necessarily entail that one ought to ϕ. There might be other pro tanto reasons *not* to ϕ, or to do something else that excludes ϕ-ing, which weigh heavier than one's reason to ϕ. This being said, the distinction between all-things-considered reasons and pro tanto reasons should not be understood too sharply. Sometimes a pro tanto reason to ϕ by itself makes it the case that one ought to ϕ – for example in case there are no opposing reasons not to ϕ. In that case, the pro tanto reason to ϕ is an all-things-considered reason as well. Moreover, what we treat as all-things-considered reasons are often in fact the result of weighing several pro tanto reasons for and against the relevant action. The distinction is thus somewhat blurry, and I agree with those who regard the notion of all-things-considered reasons as derivative of the notion of pro tanto reasons (Broome 2004). Nevertheless, the distinction is useful.[7]

For reasons that will emerge later in the chapter, it will be useful to represent these relationships between reasons and corresponding ought-statements as conditionals, having reason-statements as antecedents and statements about what an agent ought to do in the appropriate sense as consequents:

Necessarily, if S has all-things-considered reason to ϕ, then S ought to ϕ.[8]

Necessarily, if S has pro tanto reason to ϕ, then, in the absence of opposing reasons, S ought to ϕ.

Different kinds of reasons can thus be distinguished, at least in part, by the kinds of ought-statements they entail. In investigating a class of reasons,

[7] Some have suggested that there are reasons that merely 'entice', i.e. make it the case that ϕ-ing would be *attractive* for the agent, without thereby implying that the agent would commit a mistake or otherwise be in the wrong if he fails to respond to it (Dancy 2004). For a discussion relating these different kinds of reasons to evidence, see Steglich-Petersen (2008b).

[8] 'Necessarily' indicates that the relevant conditional is that of entailment rather than a contingent relation. Whenever not explicitly mentioned otherwise, I shall take 'if' to stand for the material conditional.

part of what we are interested in is the kind of ought-statement entailed by that class of reasons, i.e. in what sense the reason in question favours the relevant action or attitude. As the case of pro tanto reasons shows, reasons do not always entail unqualified oughts, but only more complex or conditional statements including oughts. This will become important later on when investigating the nature of epistemic reasons, where it is all but clear in what sense they favour or make it the case that one ought to form the relevant belief.

Although the teleological account of reasons might be committed to certain characteristics of the favouring relation that reasons can stand in,[9] it is primarily an account of another aspect of reasons, namely of the kinds of considerations or facts that are able to constitute reasons, and thus stand in the favouring relation to actions or attitudes in the first place. According to the teleological account, some fact or consideration can be a reason for an action or attitude only in virtue of being a fact or consideration about the value, intrinsic or derived, of the action or attitude that it favours. For example, the fact that an action promotes well-being is a reason to perform it, since well-being is valuable.

Finally, by 'epistemic reasons', I shall refer to considerations that count in favour of holding a belief in a particular proposition solely on account of that (potential) belief's epistemic properties. In this chapter I wish to remain neutral on the exact nature of the epistemic properties giving rise to epistemic reasons, but by 'epistemic properties' I shall generally mean properties that are in some way relevant to whether or not the proposition believed is true. On this very general definition, being evidentially supported is the most obvious candidate for an epistemic property, but I shall not exclude other candidates in advance.

## 1.3 EPISTEMIC REASONS AND VALUE

With this rough characterization of reasons, and the teleological account of these, we can go on to consider the teleological account of epistemic reasons specifically. According to the common version of this account, epistemic reasons for beliefs obtain in virtue of the value, either intrinsic or instrumental, of the epistemic properties of the beliefs they support. When I have an epistemic reason for a particular belief, the belief must

[9] For example, it is sometimes argued that teleological accounts of reasons are committed to the view that all reasons can be compared and weighed against each other. For this view, see in particular Joseph Raz (1986).

have some epistemic property that somehow constitutes or promotes something of value. On the teleological account, epistemic reasons thus share an important characteristic with practical reasons: they both support actions or attitudes in virtue of those actions or attitudes promoting or constituting something of value. What distinguishes epistemic reasons is that the valuable property of beliefs, which the reason is based on, is an epistemic property, such as being true, supported by evidence, etc., depending on the specific teleological account. This means that the teleologian must explain why those epistemic properties are valuable or promote something of value.

Critics often point out that beliefs supported by epistemic reasons do not thereby become valuable or come to promote something of value. If epistemic reasons for beliefs are reasons that obtain in virtue of the relevant beliefs having some epistemic property that is of value or promotes something of value, then *all* beliefs supported by epistemic reasons must be valuable or value-promoting. But not all beliefs supported by epistemic reasons are valuable or value-promoting. In fact, some beliefs are just the opposite, despite being supported by epistemic reasons. Numerous examples of such beliefs have been given in the literature. Most of them invoke propositional contents that it would be trivial or even harmful to form beliefs about, but are nonetheless supported by epistemic reasons. To give but a few recent examples: Thomas Kelly (2003) asks us to imagine being about to watch a suspense thriller. In that case, most of us would prefer not having epistemically supported beliefs about the ending of the movie, since our enjoyment of it hinges at least in part on not knowing the ending. Alvin Goldman (1999b) and Stephen Grimm (2009) offer several examples of propositions that it wouldn't be outright value-detracting, but nevertheless worthless to form beliefs about: What is the 323rd entry in the Wichita, Kansas phone directory? Is there an even number of dust specks on my desk? Such cases seem to show that epistemic reasons for beliefs do not depend on the value of those beliefs. A belief can be true, justified, warranted, supported by evidence, etc., without thereby becoming either intrinsically valuable, or conducive to something of intrinsic value. A world in which I have a true, warranted belief about the number of dust specks on my desk is not better than a world in which I have no such belief. A world in which I have a true, warranted belief about the ending of a suspense thriller I am about to watch is actively *worse*, certainly from the perspective of my own aims, but arguably also intrinsically, than a world in which I have no such belief. Epistemic evaluation of beliefs thus seems entirely independent from questions of value.

It may well seem fantastic to claim that epistemic reasons do not depend on the value of the beliefs that they support. Surely, a warranted belief is more valuable than an unwarranted one; surely, a true belief is more valuable than a false one. But if the above criticism of the teleological account is sound, that is the conclusion we are driven to accept. Although it *often* is the case that a warranted belief is valuable, no *general* claim of that sort seems available. It is always possible to find counter-examples involving epistemically supported beliefs in very trivial propositions, or propositions that the believer would be better off not having beliefs about.[10]

There are two common strategies for averting this objection against the teleological account.[11] Some defenders of the teleological account have tried to avert it by attempting to show that epistemic reasons, despite appearances, *do* indicate that the relevant beliefs are somehow valuable. Michael Lynch (2004) insists, for example, that epistemically well-founded beliefs (in his case, true beliefs), about even extremely trivial or inconsequential matters *are* valuable, since we value truth in its own right. The main problem with this view is that it seems implausible to claim it worthwhile or valuable to form beliefs about trivial matters such as those mentioned above. Other teleologians recognize that some epistemically well-founded beliefs aren't valuable, but attempt to salvage the teleological account by restricting epistemic reasons to support beliefs that are of interest.[12] But this strategy undermines the basic tenet of the teleological explanation of epistemic reasons. If epistemic reasons obtain in virtue of the relevant beliefs having valuable epistemic properties, then beliefs without valuable epistemic properties cannot be supported by epistemic reasons. But it seems that even extremely trivial and uninteresting beliefs can be supported by epistemic reasons. However uninteresting, it is possible to have an epistemically well-supported belief that there is an even number of dust specks on my desk. Although I cannot argue for this claim in full here, it thus seems that both common strategies in salvaging the teleological account are in trouble.[13]

[10] Something not often noted is that the value-independence of epistemic reasons, if true, would be just as puzzling for non-teleological accounts. How can *any* reason, let alone teleological ones, be entirely independent of value? Some have suggested that reasons for belief are instead generated by a special constitutive non-teleological norm governing belief (Shah 2003; Shah and Velleman 2005). For criticism of this view, see Steglich-Petersen (2006a, 2008a).

[11] For a more thorough recent criticism of these two strategies along the same lines, see Grimm (2009).

[12] This strategy is adopted by, for example, Goldman (1999b) and Alston (2005).

[13] If epistemic and practical values could be weighed against each other, these problems would have an easy solution: in all of the mentioned cases, the epistemic value of the relevant belief is outweighed by some competing practical value. It is doubtful, however, whether such weighing is possible. For further discussion, see Reisner (2008) and Steglich-Petersen (2009).

In the following, I propose to accept that epistemic reasons can obtain independently of the value, whether intrinsic or derived, of the beliefs that they support. My belief that the suspense thriller I am about to watch ends in a particular way can be epistemically unassailable, entirely underwritten by epistemic reason, even if it would be better from a value perspective if I hadn't had that belief. Likewise, my belief that there is an even number of dust specks on my desk can be well supported by epistemic reasons, even if nothing of value would be promoted by forming such a belief. Nevertheless, I will argue that there is a broadly teleological account compatible with these claims.

## 1.4 TWO KINDS OF REASONS FOR BELIEF

Normally, reasons for belief are taken to be reasons for believing *that* some particular proposition, or body of propositions, is true. But this is not the only kind of reasons relating to beliefs. Another important class of reasons for belief are the reasons we might have to form or have beliefs *about* particular propositions, or sets of propositions. When one has reason to form a belief *about* some proposition *p*, one has reason to form a belief as to whether *p* is true or not. These two kinds of reasons are clearly independent of each other in the sense that neither entails the other in regard to some particular proposition. Having reason to form a belief *about* or *as to whether p* does not entail having a reason to believe *that p*, nor vice versa. I can have reason to form a belief as to whether the Americans landed a man on the Moon, without having reason to believe that they did in fact land a man on the Moon. For example, I might expect to be met with that question in a quiz, but not yet have any evidence concerning it. On the other hand, I can have reason to believe that there is an even number of dust specks on my desk (in the sense of having the relevant evidence), without thereby having any reason to form a belief *about* that proposition. I might be completely uninterested in the matter, and with good reason.

The notion of reasons to form beliefs *about* propositions or subject matters is in some respects similar to the notion of having a reason to engage in *enquiry*, i.e. a reason to actively pursue evidence bearing on a certain proposition or subject matter. There are important differences, however. For one thing, reasons to form beliefs about propositions are more general than reasons to engage in enquiry in regard to those propositions. It may be that reasons of the former kind *entail* reasons of the latter kind, i.e. that one has reason to engage in enquiry about some subject matter whenever one has reason to form beliefs about it. But that does not make them identical. Another reason to not simply treat them as identical is that

doing so would exclude in advance the possibility of practical reasons for belief, which one might find an altogether too swift argument against this possibility.

Most importantly for our purposes, reasons to form beliefs about propositions seem value-driven in a relatively straightforward sense. Whether one has reason to form a belief about a particular subject matter depends on the value, whether intrinsic or derived, of having a belief about that subject matter. Clearly, the reasons can be grounded on a multitude of values. There is nothing to suggest that a single value, or category of values, should govern such reasons. They may be based on personal interest or curiosity. Arguably, my interest in road cycling gives me reason to form beliefs about the latest developments in the Tour de France, while those without such interests have little or no reason to form such beliefs. In other cases, the reasons might be independent of personal interests, and even be universal in nature. Even if I have little personal interest in Danish politics, it is arguably my duty as a Danish citizen to exercise my democratic rights by voting, and thus to form beliefs about subject matters that are relevant to making an informed decision. At a more general level, we might all have a duty to form beliefs about the ways in which our actions affect others, the environment, or other common goods. The reasons may also be purely instrumental, as when we have reason to form the beliefs necessary to achieve goods of various sorts.

Another characteristic of reasons to form beliefs about propositions or subject matters is that they come in degrees of weight or importance. Although I have reason to form beliefs now about the latest developments in the Tour de France, there might be other and more urgent matters I have more reason to form beliefs about. As such, reasons to form beliefs about propositions can be compared and weighed against other such reasons in the usual ways. In cases where I have equal reason to form beliefs about two alternative propositions or subject matters, and cannot form beliefs about both (I might have limited time to find out about the subject matters), I am rationally permitted to choose either of the two options. Reasons to form beliefs about propositions may also be weighed against practical reasons. It is not uncommon for reasons to form beliefs about subject matters to compete with practical reasons, since finding out about subject matters takes time and resources that could otherwise be devoted to other practical pursuits. In such cases, the value of forming the relevant beliefs must be weighed against the value of pursuing the competing practical pursuits.[14]

[14] For a detailed discussion of the way reasons to form beliefs may be weighed against each other, and against practical reasons, see Steglich-Petersen (2009).

These are all characteristics that set reasons to form or have beliefs about propositions apart from epistemic reasons to believe *that* some particular proposition is true. Whether I have epistemic reason to believe a particular proposition depends on the epistemic properties of that (potential) belief only, and these properties are entirely unaffected by the interest I may or may not have in that proposition, the moral or instrumental worth of believing the proposition, or any other such value-giving property.

As such, the weighing of epistemic reasons with other epistemic reasons is entirely different in nature from the weighing of reasons to form beliefs about propositions or subject matters. In a paper not otherwise devoted to the nature of epistemic reasons, Jonathan Dancy observes the following:

> One of the most striking differences between theoretical and practical rationality is that if I have sufficient practical reason for more than one option [. . .] I am rationally permitted to choose any of them, so long as they are roughly equally supported. By contrast, if I have sufficient and equally good reason for each of a set of alternative beliefs, I am not rationally permitted to choose any of them in preference to the others. (Dancy 2004: 95)

To illustrate the point I take Dancy to make, consider first a case in which you have sufficient practical reason for two alternative options that cannot both be pursued. For example, the fact that it would be fun to spend the afternoon playing cricket is a reason to do so; on the other hand, the fact that it would be pleasant to go swimming is a reason to pursue that option instead. In the absence of defeating reasons, both considerations provide sufficient reason, and they support the options with roughly equal weight. In that case, one is permitted to choose either. But consider now a case where you have roughly equal epistemic reason to hold each of two alternative beliefs, such that both cannot rationally be held at the same time. For simplicity, we might suppose that you have evidence for both *p* and not-*p*, that the two bodies of evidence support the truth of the relevant propositions to roughly the same degree, and that either body of evidence would have been sufficient to hold the relevant belief in the absence of any other evidence bearing on the matter. In this case, one is *not* permitted to hold either belief. The two epistemic reasons defeat each other, rather than make either option permissible. This difference between practical and epistemic reasons is easy to miss, because reasons of both kinds can outweigh reasons of the same kind, in case they support their relative options to a greater degree than the competing reasons. The difference only shows up in cases where there is (roughly) equal reason to pursue either of two alternative options, or hold either of two alternative beliefs.

With these initial characterizations in place, we can explore, in a loose and tentative way, the *relationship* between reasons for having beliefs about some proposition or subject matter, and epistemic reasons for believing that particular propositions are true. It has already been noted that reasons of neither kind entail reasons of the other. But we now need to explore how the two kinds of reasons rely upon each other in determining what one *ought* to believe.

First consider the situation in which S has an all-things-considered reason to form beliefs about some subject matter. Suppose, for example, that S is a registered voter, and that much is at stake in the election. In that case, it is plausible that S has all-things-considered reason to form beliefs about the policies of the presidential candidates. This in itself is not sufficient to make it the case that S ought to believe particular propositions about the policies of the respective candidates, for example that one candidate is in favour of universal health care. If S has no evidence concerning that candidate's health-care policy, it seems plausible that S shouldn't believe any particular propositions on that topic (or only extremely trivial ones, such as the proposition that it is the case of the candidate's health-care policy that S has no evidence concerning it). So having all-things-considered reason to form a belief *about p* does not in itself entail that one *ought* to believe either that *p* or not-*p*. But suppose that S acquires sufficient evidence, and thus an epistemic reason to believe that the candidate indeed is in favour of universal health care. In that case, it seems that S ought to form the belief that the candidate is in favour of universal health care.

Consider next the case where S has no reason to form beliefs about some proposition or subject matter. Suppose, for example, that the subject matter is whether there is an even number of dust specks on S's desk. Let us also suppose that S has excellent evidence, and thus epistemic reason to believe in the sense defined, that there indeed is an even number of dust specks on his desk. In spite of this epistemic reason, it does not seem to be the case that S ought to form the belief that there is an even number of dust specks on his desk. It may be that S as a matter of fact *cannot avoid* forming that belief, since we are psychologically disposed to form beliefs that are supported by consciously considered evidence. But it is nonetheless not the case that S ought to form that belief. If S failed to form the belief, we wouldn't fault him or regard him as normatively worse off for that reason.[15] We may regret that S's general psychological disposition did not make him form the belief, since more often than not it is desirable

[15] For further arguments to the effect that one cannot derive a normative conclusion from the fact that we are psychologically disposed to form beliefs in accordance with our consciously considered evidence, see Steglich-Petersen (2006a) and Dretske (2000).

to be disposed to form beliefs in accordance with the available evidence. But we would not regard S as having failed to do something he ought to have done. Nevertheless, S has epistemic reason to believe the relevant proposition. So one may have overall epistemic reason to believe that *p* without it being the case that one ought to believe that *p*.[16] If, on the other hand, a belief about the number of dust specks should become valuable for S (an eccentric billionaire might give him a prize for forming a belief about that proposition[17]), thus giving rise to an all-things-considered reason to form a belief as to whether the content is true, this, together with the epistemic reason, would mean that S ought to believe that there is an even number of dust specks on his desk.

So far, then, the following relationships hold. If S has epistemic reason to believe that *p*, S is such that, *if* he has all-things-considered reason to form beliefs about or as to whether *p*, *then* he ought to believe that *p*. If the latter conditional is false (which it is just in case the antecedent is true and the consequent is false), it is also false that S has epistemic reason to believe that *p*. In other words, if it is not the case of S that having a reason to form beliefs about *p* materially implies that he ought to believe that *p*, we may deduce that S does not have epistemic reason to believe that *p*.

The same holds for reasons to form beliefs about or as to whether *p*. If S has all-things-considered reason to form beliefs about *p*, S is such that, *if* he has epistemic reason to believe that *p*, S ought to believe that *p*. If the latter conditional is false (which it is just in case the antecedent is true and the consequent is false), it is also false that S has reason to form beliefs about *p*. In other words, if it is not the case of S that having epistemic reason to believe that *p* materially implies that he ought to believe that *p*, we may deduce that S does not have all-things-considered reason to form beliefs about *p*. These relationships can thus be summed up as follows (where 'T' stands for 'belief that' and 'A' stands for 'belief about'):

| | |
|---|---|
| (T) | Necessarily, if S has epistemic reason to believe that *p*, then [if S has all-things-considered reason to form a belief about *p*, S ought to believe that *p*]. |
| (A) | Necessarily, if S has all-things-considered reason to form a belief about *p*, then [if S has epistemic reason to believe that *p*, S ought to believe that *p*].[18] |

[16] For further defence of this claim, see Steglich-Petersen (2008b).

[17] Since the billionaire does not require you to adopt a belief that a particular proposition is true, but only a belief *as to whether* the proposition is true, this example does not raise problems concerning the possibility of practical reasons for belief.

[18] Of course, in both schemas, the necessity operator attaches to the main connective only, and thus not to the embedded conditional. Part of the upshot of the above discussion is that reasons of neither kind ever on their own entail that one ought to form the relevant belief.

It is important to make clear that these schemas are not intended as analyses or definitions of the respective kinds of reasons. The schemas merely state certain complex relationships between reasons to form beliefs about propositions, epistemic reasons to believe propositions, and whether an agent ought to form a particular belief. As such, the schemas are compatible with the truth of a number of other claims about the relationship between these three conditions, some of which may be stronger than the ones focused on here. As indicated by the above discussion, however, the schemas do describe what appear to be fundamental properties of these two kinds of reasons. Furthermore, the schemas are quite permissive in a number of ways. For one thing, they leave open the details as to what it takes to have a reason to form a belief about some subject matter, and what epistemic properties some potential belief that *p* must have in order to make it the case that one has epistemic reason to believe that *p*. Perhaps more importantly, the schemas are silent as to whether there could be non-epistemic reasons to believe that *p*. I mentioned at the outset that I would simply assume in this chapter that some version of evidentialism is true, but nothing in these schemas excludes that it could be the case that one ought to believe that *p* in the absence of an epistemic reason. As mentioned, I take it to be a separate problem to explain the exclusivity of epistemic reasons, and I have dealt with it elsewhere (see footnote 3).

Since this chapter is primarily about epistemic reasons, I will now leave schema (A) aside, and focus on schema (T), stating a necessary conditional with a statement about epistemic reasons as the antecedent. This schema can be altered and qualified in various ways to take into account a number of subtleties. For example, one might find it plausible that the normative force of the reason one has to form beliefs *about* some subject matter carries over to the ought-statement concerning belief in a particular proposition, on the condition that one has epistemic reason for believing that proposition. For example, consider two mutually independent propositions *p* and *q*, where *p* is of great significance and *q* is of lower significance to the believer, such that one has an all-things-considered reason to form a belief about *p* and a mere weak pro tanto reason to form a belief about *q*. Suppose further that one has evidence of identical strength regarding the truth of the two propositions, and thus equal epistemic reason to believe either. Despite the equally strong epistemic reasons, it seems that the ought-statement concerning belief that *p* should be stronger than that concerning belief that *q*. In particular, it seems that in the case of *p* it is entailed that one ought to believe it, while in the case of *q* it is merely entailed that one

ought to believe it in the absence of opposing reasons (these differences are analogous to the explanation of all-things-considered reasons and pro tanto reasons set out in section 1.2). So the force of the ought-statement concerning belief in particular propositions depends on the force of the reasons to form beliefs *about* those propositions, and not on the force of the epistemic reasons one has for believing those propositions. Indeed, it is doubtful whether epistemic reasons have normative 'force' or 'weight' in that sense at all. When we speak of the 'force' or 'weight' of epistemic reasons, we usually have in mind a different quality, namely the degree to which the evidence, which the epistemic reasons depend upon, confirms the relevant proposition. But this only affects the *degree of belief* one should adopt in that proposition, not the force of the ought-statement concerning the formation of the relevant belief. If this is plausible, the schema might be qualified in the following way, where the inserted parentheses indicate that the degree of belief one ought to adopt is determined by the epistemic reason, and the split of the schema into two, corresponding to pro tanto reasons and all-things-considered reasons respectively, indicates that the normative force of the ought-statement concerning belief in a particular proposition is determined by the reason one has to form beliefs about that proposition:

| |
|---|
| (T – all things considered) |
| Necessarily, if S has epistemic reason to believe that *p* (to degree D), then [if S has all-things-considered reason to form a belief about *p*, S ought to believe that *p* (to degree D)]. |
| (T – pro tanto) |
| Necessarily, if S has epistemic reason to believe that *p* (to degree D), then [if S has a pro tanto reason to form a belief about *p*, then, in the absence of opposing reasons, S ought to believe that *p* (to degree D)]. |

However, in the following I will ignore these qualifications of schema (T), and focus on the initial simple version of (T) only. The important thing for our purposes is that the schema specifies a precise sense in which epistemic reasons for some belief are related to the value of having that belief. If reasons for forming beliefs about propositions are value-driven, then epistemic reasons are value-driven too, in the sense that they entail that one ought to believe particular propositions in the context of a value-driven reason only. Pointing out this dependence of epistemic reasons on values is a significant step towards solving the value problem for the teleological

account of epistemic reasons. It is not in itself, however, sufficient. What is still missing is an *explanation* of why this dependence holds.

## 1.5 A TELEOLOGICAL EXPLANATION OF THE RELATIONSHIP

The teleological explanation to be offered in this section is a version of the instrumentalist account of epistemic reasons. According to one influential version of this account, epistemic reasons are a species of standard instrumental reasons. Roughly, S has instrumental reason to ϕ if and only if there is some aim that S has reason to pursue, and ϕ-ing is the best available way (or a step in the best available way) of realizing that aim.[19] For example, if I have reason to go to London, and going to the airport is a step in the best available way of going to London, I have instrumental reason to go to the airport. The standard analogous instrumentalist explanation of epistemic reasons proceeds roughly as follows. First it is observed that forming a belief involves having a particular aim, for example that the belief is true or amounts to knowledge. So if S has reason to form a belief about *p*, S has reason to pursue the aim involved with forming a belief about *p*, i.e. believing the truth concerning *p*, or coming to know whether *p*, depending on the specific account. This means that one has instrumental reason to believe that *p* just in case one has reason to form a belief about *p*, and believing *p* is the best way to realize the aim involved with forming such a belief. Finally, it is pointed out that the conditions under which believing *p* is instrumental or conducive to the aim involved with forming a belief about *p* coincide with the conditions under which one has epistemic reason to believe that *p*, thus making it plausible to regard epistemic reasons as a special case of instrumental reasons.

As already pointed out, however, this simple instrumentalist explanation is flawed since epistemic reasons do not depend in such a straightforward sense on reasons for forming the relevant beliefs, and thus for adopting the putative aim involved with forming those beliefs. I can have an epistemic reason to believe that there is an even number of dust specks on my desk, regardless of whether I have reason to form a belief about that matter in the first place. So a straight assimilation between epistemic reasons and instrumental reasons is not available.

However, there is a different sense of instrumental reasons to which epistemic reasons might be assimilated instead. This is the purely hypothetical

[19] Although this rough account will be sufficient for our purposes, the question of how to correctly understand instrumental reasons is a subtle and much-debated issue. For representative strands of the recent debate, see Broome (2007c), Schroeder (2004), and Wallace (2001).

sense of instrumental reasons that we have to pursue hypothetical aims, regardless of whether or not we in fact have those aims, or have reason to pursue those aims. For example, regardless of whether I have reason to go to London, I am such that *if* I have reason to go to London, *then* I ought to go to the airport. Countless such conditional statements are true of each of us all the time, corresponding to all of the instrumental actions to hypothetical aims that are available to us. For each of the endless number of hypothetical aims I might adopt, there will be a number of available actions that are instrumental to that aim. Once again, we can characterize this kind of reasons in terms of a conditional statement (where 'HI' stands for 'hypothetical instrumental reason'):

| | |
|---|---|
| (HI) | Necessarily, if S has hypothetical instrumental reason to ϕ in pursuit of aim A, then [if S has all-things-considered reason to pursue aim A, S ought to ϕ]. |

Once again, this conditional does not amount to an analysis or a definition of hypothetical instrumental reasons, but merely states an important relationship between such reasons, all-things-considered reasons, and ought-statements. A number of qualifications and comments are needed. The first concerns the possibility of competing instrumental reasons to pursue the same aim. Often there will be several ways to achieve the same aim, so one cannot infer from the fact that there is instrumental reason for S to ϕ in order to achieve A, to it being the case that S instrumentally ought to ϕ, since there might be a stronger instrumental reason to ψ, if ψ-ing is an overall better way of achieving A. So in the above schema, 'S has hypothetical instrumental reason to ϕ in pursuit of aim A' should be understood as indicating an *overall* hypothetical instrumental reason, i.e. that ϕ-ing is the best available way to pursue A. Under this assumption it might be observed that the normative force attaching to the consequent ought-statement carries over from the normative force of the reason for pursuing the relevant aim: if the reason for pursuing the relevant aim is a mere pro tanto reason, the consequent ought-statement should be qualified to state that S ought to ϕ in the absence of opposing reasons not to pursue aim A. Finally, a comment is needed to dispel any worries about the scope of the consequent ought in the embedded conditional. In discussions about instrumental rationality, i.e. the rationality of taking the means to one's aims, it is often assumed that instrumental rationality gives rise to *wide-scope* oughts, rather than *narrow-scope* ones. On the narrow-scope account, if S has some aim A, and ϕ-ing is necessary in pursuing A, S ought to ϕ. But this has the counterintuitive result that one ought to take the

means to aims that one does not have reason to pursue. For example, if an arsonist has the aim of burning down his neighbour's house, he ought to pour some gasoline on it if that is necessary to burn down the house. Proponents of the wide-scope account suggest instead that the central demand of instrumental rationality should be phrased in terms of an ought that takes the entire conditional as its scope, in the following way: S ought to ensure that *if* S has aim A, and ϕ-ing is necessary in pursuing A, S ϕs. According to wide-scopers, this avoids the counterintuitive result since it is silent as to whether one should comply with it by giving up the aim, or by taking the means to it. Given this, however, it should be clear that none of these worries are relevant to the above schema concerning hypothetical instrumental reasons. In the schema, the embedded conditional moves from S having an all-things-considered reason to pursue A, to it being the case that S ought to take the means. So it never demands that one take the means to an aim one does not have reason to pursue.

With this explication of hypothetical instrumental reasons in place, we can return to the issue of epistemic reasons. If we interpret reasons to form beliefs about propositions as reasons to pursue a certain aim involved with believing (for the present purposes we need only identify this as the aim of forming a belief as to whether some proposition is true), and interpret believing particular propositions as ways of pursuing that aim, such that one forms a belief in a particular proposition in order to form a belief about or as to whether that proposition is true, the above schema for hypothetical instrumental reasons (HI) is structurally identical to schema (T) describing the relationship between epistemic reasons to believe that *p*, reasons to form beliefs about *p*, and whether one ought to believe that *p*. Both schemas state a conditional relationship between the obtainment of a certain kind of reason and the truth of another conditional, moving from the obtainment of a different kind of reason to an ought-statement. Furthermore, in both schemas, the normative force of the consequent ought-statement is carried over from the reason in the embedded conditional, and the content of what one ought to do is carried over from the reason on the left hand side of the overall conditional. This structural identity suggests that one possible explanation of the relationship outlined in section 1.4 is that epistemic reasons are simply special cases of what we have called 'hypothetical instrumental reasons'. To be precise, they are instrumental reasons for believing particular propositions in pursuit of the hypothetical (i.e. present or non-present) aim of forming beliefs about those propositions. It is this fact that explains the special relationship between epistemic reasons, reasons to form beliefs about certain propositions, and ought-statements,

explicated in the previous section. This claim obviously stands in need of further clarification. A helpful way of doing so will be to test it against a number of objections.

## 1.6 OBJECTIONS

Perhaps the most influential objection to standard instrumentalist accounts of epistemic reasons is that it would make epistemic reasons hypothetical in a sense they are clearly not. As mentioned in the previous section, on the standard version of the instrumentalist account, epistemic reasons depend on the contingent aims of the believer. In an influential discussion, Kelly has pointed out that epistemic reasons are not hypothetical in that sense. Most importantly, epistemic reasons are usually considered to be *intersubjective* in a way that they could not be if they depended on the particular aims of agents. When we offer evidence to someone, we take the evidence to provide epistemic reason for that person to believe a particular proposition, regardless of whether she happens to have the relevant cognitive goal (Kelly 2003: 621).

The present account avoids this objection in two distinct ways. First of all, it does not make what one ought to believe depend on the aims or desires of agents, but rather on the reasons agents have to form beliefs about certain propositions or subject matters. As mentioned above, it seems that such reasons are value-based in a relatively straightforward sense: whether one has reason to form beliefs about some subject matter depends on the value of doing so. I have left it open whether such reasons might be categorical in certain instances, but there is no principled reason why they couldn't be on the present account. Secondly, on the proposed account *epistemic* reasons are not hypothetical in the problematic sense that Kelly has in mind. Whether someone has epistemic reason to believe a particular proposition depends on the evidence available concerning the truth of that proposition. As such, epistemic reasons are intersubjective – providing evidence to someone is equivalent to providing that person with an epistemic reason, regardless of the particular aims and desires of that person. Epistemic reasons are hypothetical only in the sense that *whether one ought to form a belief in accordance with them* depends on an additional factor, namely whether one has reason to form beliefs *about* the relevant propositions in the first place. And it is fully compatible with the present account that these additional reasons might be categorical.

Another possible objection is that the present proposal downplays the normative significance of epistemic reasons to an implausible extent. The

proposal would in effect assimilate epistemic reasons to the infinite number of conditional statements about instrumental reasons that are true of us all the time. For example, whether or not I have reason to go to London, it is true of me right now that *if* I have all-things-considered reason to go to London, *then* I ought to go to the airport. This conditional truth follows from the fact that going to the airport is necessary in order for me to go to London. Countless other such conditional statements are true of me at all times. It is true of me that if I have reason to call my mother, then I ought to turn on my phone, and so on *ad infinitum*. Can it really be true that epistemic reasons are as normatively insignificant as that?

But setting it up this way makes the situation seem worse than it really is. Two considerations should make it appear decidedly more palatable. First of all, if I *did* have a reason to go to London, it *would* be normatively significant that a necessary step for me to do so is to head out to the airport. The same goes for epistemic reasons, and we *do* have reason to form beliefs about a rather large number of propositions and subject matters. So very often, it *is* of normative significance to be such that *if* one has reason to form beliefs about some proposition, *then* one ought to believe that proposition. Secondly, it does seem true that we have much more evidence available than we have reason to respond to by forming the appropriate beliefs. If epistemic reasons are generated by evidence available to us, we have a large, perhaps infinite, number of epistemic reasons all the time, and far from all of them are about matters worth forming beliefs about. It would be absurd for an account of epistemic reasons to entail that we are at fault for failing to respond to all those epistemic reasons, so *any* account of epistemic reasons must be able to explain why we are quite justified in not caring about the vast majority of them. The proposed account does just that, by making the normative import of epistemic reasons depend in a systematic way on a different kind of value-based reasons.

In this way, the account may also help to resolve a current dispute over the normative status of evidence. In his original criticism of instrumentalism about epistemic reasons, Kelly (2003) argued that the instrumentalist cannot explain how we can have epistemic reasons for believing propositions that we have neither reason nor desire to form beliefs about, such as propositions about the ending of a movie we are about to watch. Adam Leite (2007) has objected to this that we ought to distinguish between *having epistemic reason* to believe some proposition, and *having evidence* suggesting that proposition to be true. While having an epistemic reason is a normative property of some person, having evidence is a purely descriptive property. So having evidence doesn't necessarily imply having an epistemic

reason. But in his reply to Leite, Kelly (2007) argues that there is no clear non-normative sense of evidence.[20] As the argument goes, evidence cannot be understood save as something that makes a difference to what one is justified in believing. Kelly even suggests that 'evidence' might be more or less synonymous with 'reason to believe', at least as long as this latter term is understood in the epistemic sense (2007). If that is correct, the instrumentalist cannot rely on a sharp distinction between those two terms in answering Kelly's initial challenge, as Leite has suggested. It seems, however, that the account proposed here can accommodate the intuition that evidence must be understood in normative terms, and even that 'evidence' is synonymous with 'epistemic reason', without thereby committing to the view that having evidence requires that one form a belief in accordance with it. For suppose that we accept the equivalence between having evidence and having an epistemic reason. On the current account, none of these terms would then be understandable in non-normative terms, since they (necessarily) entail a conditional statement about what one ought to believe. But since the 'ought' in question is embedded in a conditional, Kelly's initial challenge is avoided. The embedded conditional can be true without the consequent ought-statement of that conditional being true. One can thus have evidence, and therefore epistemic reason to believe some proposition, without it being the case that one ought to form a belief in that proposition, as in Kelly's problematic movie case.

A further possible objection is that the account fails to explain why one can be rationally unassailable in holding an evidentially supported belief in a proposition that one has no reason to have or form beliefs about in the first place. It seems, for example, that even if I have no reason to have or form a belief about the ending of the movie I am about to watch (even, in fact, if I have all-things-considered reason *not* to), if I inadvertently were to acquire evidence about the ending and form a belief accordingly, that belief would be rationally unassailable. The most plausible answer to this objection, I think, is to point out that even if the belief in question is unassailable in the sense that the believer cannot be blamed for having formed it, the belief is nonetheless regrettable. It is regrettable for someone to form a belief about the ending of a movie he or she is about to watch, especially if the enjoyment of the movie depends partly on not knowing how it ends. This point seems to generalize to all beliefs that we have reason not to form. We routinely withhold evidence from people because we judge that they have reason not to form the belief in question, and

[20] Kelly attributes this point to Jaegwon Kim (1988).

this seems a perfectly justifiable practice. So if the beliefs in question are 'rationally unassailable', this must be in a specific sense rather than *tout court.* But if that is the case, the present account can explain the rational unassailability of such beliefs in a perfectly straightforward way. On the present account, if S has epistemic reason to believe that *p*, S is such that, *if* she has all-things-considered reason to form beliefs about *p*, S ought to believe that *p*. A belief can thus be rationally unassailable for S from an epistemic point of view (supported by epistemic reason, that is), *without* S being such that she ought to believe that *p*.

## 1.7 CONCLUSION

I have proposed that epistemic reasons be regarded as a species of what I called *hypothetical instrumental reasons*. The arguments came in two stages. First it was argued that, necessarily, if S has epistemic reason to believe that *p*, then S is such that *if* she has all-things-considered reason to form beliefs about *p*, S ought to believe that *p*. It was then argued that this relationship between epistemic reasons, reasons to form beliefs about certain propositions, and normative statements about what a person ought to believe, could be explained by supposing that epistemic reasons are hypothetical instrumental reasons. The main virtue of this account is that it provides a sense in which epistemic reasons are teleological, or value-based, without commitment to epistemic properties being valuable in a general sense, or to there being a sense in which every epistemically well-founded belief is valuable. In this way, the account avoids the perhaps most pertinent problem faced by teleological accounts in recent years, namely what I termed the 'value problem'. Furthermore, the account avoids problems typically associated with purely instrumentalist variants of the teleological account, since it avoids making epistemic reasons aim- or desire-dependent, while at the same time explaining why epistemic reasons do not in themselves require that a belief be formed. Needless to say, the account must be clarified and developed further before its merits can be assessed fully. But I hope to have shown that the teleological account may be more resilient than is often supposed.[21]

[21] This paper was presented to audiences in Aarhus, Stockholm and Geneva, and at the annual meeting of the Philosophical Association of Southern Africa. For helpful comments and criticism, I wish to thank Carl Erik Kühl, Johanna Seibt, Raffaele Rodogno, Andrew Reisner, Jane Heal, Pascal Engel, Kathrin Glüer-Pagin, Peter Pagin, Jonas Olson, Ward Jones, Veli Mitova, Timothy Chan, Davide Fassio, Philipp Keller, Igor Douven, and two anonymous referees for Cambridge University Press.

CHAPTER 2

# *Is there reason to be theoretically rational?*

*Andrew Reisner*

## 2.1 INTRODUCTION

An important advance in normativity research over the last decade is an increased understanding of the distinction, and difference, between *normativity* and *rationality*. *Normativity* concerns or picks out a broad set of concepts that have in common that they are, put loosely, guiding. For example, consider two commonly used normative concepts: that of a *normative reason* and that of an *ought*. To have a normative reason to perform some action is for there to be something that counts in favor of performing that action, perhaps that doing so will result in a good outcome. Likewise with ought, when there is sufficient evidence for something, one ought to believe it (at least under normal circumstances). Not all guidance need be directed toward a specific mental state or a specific action. Subject to the requirements of normativity, too, are relations. It is commonly believed, for example, that we ought not to hold contradictory beliefs.[1]

At least some of the requirements that concern relations amongst an agent's mental states are, or seem, distinctive. Agents who fail to satisfy these requirements are considered irrational to some degree. On many current views, being irrational is distinct in some way from not being how one ought to be; *rationality* is a concept commonly thought distinct from *normativity* by philosophers working on reasons and oughts. Much of the literature on this topic over the last decade stems from attempts to capture the characteristic features of the requirements of rationality. Two influential views in particular did much to set the agenda. The first of these was put forward by John Broome.[2] His view, the particulars of which I shall discuss in more detail below, is that the requirements of rationality

[1] That is, one ought not to believe $p$ while believing $\sim p$. This is an ought that governs a relation. Contrast the relational case with the ought that governs a single state with complex content: one ought not to believe $p$ & $\sim p$.

[2] Broome (2000).

could be expressed using a normative relation, which he calls a "normative requirement." Normative requirements are conditionals governed by an all-things-considered *ought*. In the case of rationality, the conditional is made up entirely of mental states. As an example, you ought not to (believe *p* and also believe not *p*). For the moment, two points should be stressed. The first is that Broome believed that the norms of rationality (or at least a great many of them) are distinguished by their logical form[3] and that the normativity of rationality is that of the final, all-things-considered ought. To put it another way, the requirements of rationality have no distinct normative status, but they have a distinct shape. The second thing that should be stressed is that the requirements of rationality have their normativity because of the logical or conceptual relations amongst the contents of the relevant mental states. Rationality requires of us that we do not believe both a conditional and its antecedent, while believing the negation of the conditional's consequent. The explanation for this, according to Broome, was that the contents of those beliefs must accord with the logic of *modus ponens*.[4]

Derek Parfit develops the second influential line of reasoning in a paper from 2001.[5] Here, Parfit makes the now widely accepted claim that the requirements of rationality are not themselves either *oughts* or *normative reasons*. According to Parfit, what we ought (or in his terms, "have most reason") to do is given to us by the facts. He adopts a strong form of externalism, in which these facts are for the most part not facts about our mental states. What is most rational, however, is what we would have most reason to do, were our non-normative beliefs true. His view is strikingly different from that set out in Broome's earlier paper for a number of reasons, but there is just one aspect that requires attention here. Parfit's view draws on a contrast between *reasons* or *oughts*, on the one hand, and *rationality* on the other. Rationality places requirements on the relations amongst our mental states, but those requirements are not, and do not have the force of, normative reasons or oughts. Broome's view was that, in fact, the requirements of rationality are normative requirements, i.e. oughts that govern conditionals.

Broome was later to adopt Parfit's point of view about the distinction between normativity and rationality, although Broome retains his very

[3] In fact, this statement requires some care. Not all normative requirements were rational requirements. And, in principle, there might be normative requirements that governed only mental states that were not rational requirements, but rational requirements were normative requirements and were formally indistinct from other normative requirements.

[4] Broome (2000). [5] Parfit (2001).

different ideas about what rationality requires of us.[6] Indeed, there was broad consensus within just a few years of the publication of Parfit's work that there is an important distinction at hand.[7]

This chapter aims to settle a lingering question raised by having made a distinction between rationality and normativity: is there reason (a normative reason or an ought) to be rational, even though we have accepted that rationality and normativity are conceptually distinct? In particular, this chapter will look at the more specific matter of whether there is reason to be theoretically rational. I shall argue that there is very strong reason to be theoretically rational. Unfortunately, answering the question about theoretical rationality does not settle questions about the normativity of practical rationality, but I shall outline in a speculative way how the considerations raised here in favor of the normativity of theoretical rationality might be used to show that the requirements of practical rationality are normative.

## 2.2 THE STRUCTURE OF NORMATIVE REQUIREMENTS

We often say that you ought to do this or that. Quite often what you ought to do is a simple thing, like going to the shop. In such cases we say, "You ought to go to the shop." Sometimes we want to express a more complex thought about what you ought to do. A more complex thought is that there is some sort of conditional thing that you ought to do. In English, when we want to say that there is a conditional thing that you ought to do, or make the case, we express it by the locution, "If $p$, then you ought to $q$." While this locution seems straightforward enough at first blush, it is capable of expressing normative conditionals of two different forms. The distinction between these forms is clear when they are formalized. The first logical form follows the English word order. Here, the ought has a *narrow scope*:

$$\text{F1. } p \rightarrow \mathrm{O}q$$

F1 expresses a material conditional with the normativity attached to the consequent. To use Broome's terminology, we can say that the normativity in this case is *detaching*. We say it is detaching, because when the antecedent of this conditional is true, then it is the case that you ought to satisfy the

[6] Broome (2005).

[7] See Broome (2005), Kolodny (2005, 2007), Parfit (2001), Reisner (2009b), Wedgwood (2003). For those who remain uncertain about the distinction, see Raz (2005).

consequent. Thus, when $p$ is the case, we can detach, by *modus ponens*, the proposition that you ought to $q$. Let us call this reading of "If $p$ then you ought to $q$" the "narrow-scope" reading. We can say that the conditional in F1 has a *narrow-scope* ought. We say that the ought has a narrow scope because it governs only the consequent of the conditional.

The second logical form does not follow the English word order. In order to express this alternative logical form unambiguously in English, we have to bend normal English syntax slightly, so that "ought" takes a proposition rather than an infinitive. So, we would say "You ought that if $p$ then $q$":[8]

$$\text{F2. } O(p \rightarrow q)$$

This reading is the *wide-scope* reading of the conditional. F2 is a *wide-scope* ought. We call this ought a "wide-scope" ought, because the normativity does not govern only the consequent, but rather governs the entire conditional. To use Broome's terminology again, we can say that the normativity here is *non-detaching*. It is non-detaching because from $O(p \rightarrow q)$ and $p$, one cannot detach the conclusion $Oq$. We should keep two things in mind about the distinction between wide-scope oughts and narrow-scope oughts. The first is that narrow-scope oughts are detaching: one can detach a normative consequent when the antecedent is true, whereas one cannot do so with wide-scope oughts. The second is that in the case of wide-scope oughts, the agent to whom the ought applies is responsible for the truth of the conditional. Thus, one has satisfied a wide-scope ought as long as any of the truth conditions for the conditional that it governs have been met.

In Broome's terminology, a normative requirement is a wide-scope ought.[9] That $p$ normatively requires $q$ just means that you ought that if $p$ then $q$, to use the awkward English phrasing. Broome had, in the past, argued that many requirements of rationality are normative requirements.[10]

Before looking at why one might think that rational requirements, or at least some rational requirements, might be normative requirements, it

[8] This change in syntax does less violence to the language than might be feared. Infinitives have implied subjects when they complement a modal auxiliary verb. When you say "You ought to go to the shop," the infinitive "to go to the shop" has "you" as its implied subject. Furthermore, other ways of saying you ought to do something – "It is fitting that . . ." or "It is meet that . . ." – take propositions rather than infinitives.

[9] In the original paper, "Normative Requirements," Broome conceives of normative requirements as being more complex than a wide-scope ought governing a material conditional. However, he has since given up the additional logical features and has identified normative requirements as wide-scope oughts. See Broome (2004).

[10] See Broome (2000).

will be helpful to look at why one might think that rationality is a type of normativity at all.[11] One reason for taking the requirements of rationality to be requirements of normativity, i.e. ought or reasons claims, is that doing so explains the pressure to be rational. If rational requirements are something other than oughts or reasons, then there is a separate question of why, or if, one ought to be rational.[12] If the requirements of rationality are normative requirements that concern relations amongst our mental states, then there is no question as to whether one ought or has reason to be rational; to ask the further question would be to ask if one ought to do what one ought to do, or if one has reason to do what one has reason to do.[13]

The use of wide-scope oughts, or perhaps wide-scope reasons, in giving an account of (at least some) rational requirements is important for any view that holds that rational requirements are normative. This is because of what Michael Bratman calls the "bootstrapping problem."[14] The bootstrapping problem arises when we try to explain the role of belief in theoretical rationality and the role of intention in practical rationality.

Consider my belief that it is Tuesday and also my belief that if it is Tuesday, then I am in Belgium. These two beliefs rationally require me not to believe that I am not in Belgium. If we are giving an account of rational requirements in terms of normativity, there are three plausible analyses of this particular example. The first analysis is that there is a narrow-scope ought attached to the consequent. If I believe that it is Tuesday and I believe that if it is Tuesday, then I am in Belgium, then I ought not to believe that I am not in Belgium:

$$\text{F3. } [Bt \;\&\; B(t \rightarrow b)] \rightarrow O \sim B \sim b$$

The second analysis is similar to the first, but with a reason rather than an ought attached to the consequent. If I believe that it is Tuesday and I believe that if it is Tuesday, then I am in Belgium, then I have a reason not to believe that I am not in Belgium:

$$\text{F4. } [Bt \;\&\; B(t \rightarrow b)] \rightarrow R \sim B \sim b$$

F3 and F4 are both subject to the bootstrapping objection. I may have no reason whatsoever to believe that it is Tuesday or that if it is Tuesday then I am in Belgium. Indeed, I may have very good reason to believe that I am not in Belgium. If both my belief about the conditional and my belief

[11] Niko Kolodny (2005) discusses this issue in a great deal more detail.
[12] Kolodny (2005) thinks that there is no reason to be rational.
[13] I take the liberty of treating "do" here as a universal verb.
[14] Bratman (1999). See especially pp. 24–27.

about its antecedent are unjustified or there is no reason to believe them (or perhaps there are strong reasons not to hold those beliefs), then it is hard to see what reason there is to deny a belief that is inconsistent with them. Just the having of some beliefs does not in normal circumstances stand as evidence for the truth of the logical consequences of those beliefs.[15]

The thrust of the objection to using detaching normative relations when characterizing rationality is that no matter how little reason there is for an agent to have certain beliefs or intentions, or indeed no matter how much reason there is for an agent not to have those beliefs or intentions, having those beliefs or intentions would make it the case that there is some reason for the agent to have the mental states that follow rationally from the initial ones. In other words, if rationality is expressed by detaching normative relations, there would be a reason for an agent to believe what follows immediately from two irrationally held beliefs, and there would be a reason for an agent to intend what follows immediately from an irrational intention and an irrational belief about the necessary means to carrying out that intention.

As a result of the difficulties involved in expressing the requirements of rationality with detaching normative relations, there is some appeal to a third analysis: the idea that rationality is really a system of non-detaching normative relations. I shall ultimately argue that this analysis is very much on the right track, although spelling out just how and why requires some care. In particular, we must be careful to distinguish between the view that rational requirements are themselves special instances of oughts or reasons and the view that they give rise to oughts or reasons. It is the latter view that has the most plausibility.

It will be helpful to say just a little more about what notion of rationality it is that we are trying to capture. There are two distinct aspects to rationality, only one of which we need be concerned with. The first one, the one that is not (or is not very) relevant here, is the descriptive aspect. Descriptive theories of rationality aim at describing how it is that we reason or try to give an accurate picture of the relations amongst certain people's mental states. The second aspect is the norm or requirement expressing part, and it is this one in which we are interested. This part of rationality is that part which tells us what are the correct relations amongst an agent's mental states, the lack of which are constitutive of her being (at least partially) irrational.

[15] I am assuming that there are no non-evidential reasons for the beliefs in question. It is worth noting that one can set up parallel bootstrapping objections for principles of practical rationality. Indeed, this was Bratman's original use of the objection.

It is worth noting that not all normative requirements that govern mental states are rational requirements. But it may be difficult to develop a precise set of criteria for distinguishing between normative requirements on mental states that capture a requirement of rationality and those that do not. Whatever the correct account is, it will have something to do with the nature of the relations amongst the contents of the various mental states involved. One might say that these requirements will hold in virtue of certain logical and conceptual relationships amongst the contents of an agent's mental states. Perhaps the requirements are understood as expressing a certain category of ideals concerning human mental life.

We need not take up the question of which normative requirements governing mental states are supposed to be rational requirements, because the aim here is to consider whether rational requirements are normative requirements, not to settle which normative requirements might also be rational requirements. All we need ask is whether clear examples of rational requirements are normative, and if so, why and in what way.

## 2.3 WHY RATIONAL REQUIREMENTS ARE NOT NORMATIVE REQUIREMENTS

The argument here requires the making of a distinction between two types of normative reasons: object-given reasons and state-given reasons. What we might think of as ordinary or typical reasons, those arising out of some conceptual "fit" between the attitude and its object, are object-given reasons. State-given reasons arise out of the benefits or harms that follow from having a particular attitude. An example is helpful to understanding the distinction. Normally we admire someone because she is admirable – perhaps she has through hard work and personal integrity achieved a position of importance. Or perhaps she has endured some suffering with nobility and dignity and still finds time to selflessly help others. These are the sorts of characteristics that would, under normal circumstances, provide or be reasons to admire someone. They are object-given reasons.[16] By way of contrast, there might be a very different sort of reason to admire someone. A friendly billionaire might offer you half of her fortune if you will admire Alex. Alex is a lazy sadist with a poor sense of humor. He lacks all the qualities that we might normally consider admirable. Still, the prospect of a large reward for so doing is a reason for you to admire him. This sort of reason is a state-given reason, a reason that depends not on

[16] See Parfit (2001).

how well the object of one's attitude fits the attitude, but rather on the incentives for holding that attitude.[17]

Just as there can be state-given reasons for certain pro-attitudes, like admiring, and for neutral propositional attitudes, like intending and believing,[18] there can also be the equivalent of a state-given reason for having or not having certain combinations of mental states. An example will make this point clearer. This example makes use of a paradigm principle of rationality, that believing a conditional and its antecedent rationally requires you not to believe the negation of its consequent. This may be expressed as the following normative requirement:

$$\text{NR1. } O\{[Bp \ \& \ B(p \rightarrow q)] \rightarrow \sim B \sim q\}$$

Note that the ought governs the entire conditional rather than the consequent, so we cannot detach $O \sim B \sim q$ when the antecedent is true.

Some eccentric billionaire might offer you a prize for believing $p$, believing if $p$ then $q$, and yet believing not $q$. According to the normative requirement, NR1, above, you would not be as you ought to be, should you have the first two beliefs while believing not $q$. NR1 says that it is normatively required that the following conditional not be true of you: if you believe $p$ and believe if $p$ then $q$, you believe not $q$. Imagine now that with the money you will get from believing $p$, believing if $p$ then $q$ and believing not $q$, you could and would feed all the hungry people in the world. We might very well think that you ought to, assuming that you can, believe $p$, if $p$ then $q$, and not $q$. You are, in this case, as you ought to be. We can express this as a new normative requirement:

$$\text{NR2. } O\{[Bp \ \& \ B(p \rightarrow q)] \rightarrow B \sim q\}$$

On a fairly standard account of the semantics of ought, *ought not* implies *not ought*. This means that NR1 and NR2 jointly entail a contradiction, so one or the other must be discarded.[19]

Intuitions about which one of the two normative requirements to reject might go either way, but it strikes me, at least, that it would be quite hard to explain how it is that saving all the starving people in the world does not have deontic or normative priority over violating a principle of rationality. If this is the case, then at least sometimes NR1 is false. That we sometimes ought not to have our collection of beliefs be such that when we believe

[17] This is only a rough account of the distinction. The notion of *fit* is problematic. See Danielsson and Olson (2007) for one way of working out this notion.

[18] For a defense of the view that there are state-given reasons for belief, see Reisner (2009a).

[19] This is more obvious if we look at a logical transform of NR1: $O \sim \{[Bp \ \& \ B(p \rightarrow q)] \ \& \ B \sim q\}$.

$p$ and believe if $p$ then $q$, then we do not believe $\sim q$ has a great deal of significance for how we interpret the claim that we are rationally required to be in a state in which we do not, when we believe $p$ and believe if $p$ then $q$, then believe $\sim q$.[20]

One way of interpreting this situation is that sometimes we redesignate as rational certain relations amongst our beliefs that we normally designate as irrational. This interpretation carries with it a rather substantive view that there is more to rationality than the normal constraints of logic on what it is rational for us to believe or intend. I mean to use "logic" here in a very general sense. We ought not to be happy with this interpretation of the situation for two reasons. The first is that we risk losing the notion of *rationality* when we allow conditions unconstrained by the logical relations amongst the contents of our mental states to determine which relations amongst mental states we judge to be rational.[21] The second is that this interpretation seems to conflate two distinct notions: what is rationally required or permitted, with what is normatively required or permitted. Drawing on Parfit again, we may describe the error as conflating what is rational with what there is reason for. Of course, the two may overlap or even be coextensive, but it seems to be a mistake to think that "ought" and "rationally required" are just interchangeable.

So, we can try another way of interpreting the situation, and I think that this one is the most sensible. On this view, the initial error is in thinking that *rational requirements* are conceptually identical to *normative requirements* (or form a proper subset thereof). The truth of any particular normative requirement is, like any other ought, determined by the various features of the world on which normativity is dependent. Eccentric billionaire examples can be generated for any requirement of rationality that could be given as a normative requirement; sometimes it is the case that it would be so bad to be rational that one ought not to be.

Rationality, on the other hand, is dependent on some complex of the logical features of the relations amongst an agent's mental states and their contents; the rational supervenes strictly on the mental, and this is not the case for the normative. That it would be extremely bad for an agent to have mental states that would qualify her as (locally) rational does not affect the truth of the matter about whether or not an agent's mental states are such that she is (locally) rational.

[20] This assumes that there are no genuine normative dilemmas.
[21] For a fuller explanation of this point, see Reisner (2009b).

However, one must be careful not to draw too strong a conclusion from this argument. We can see that it is possible to be rationally required to be a certain way when we are not normatively required to be this way. This argument in the past had led me, at least, to press for a strong conclusion: that rationality is not normative.[22] As it turns out, or so I shall argue, we should draw a much weaker conclusion: that rational requirements are not normative requirements where the ought involved is an all-things-considered ought. Rational requirements, or at least theoretical rational requirements, entail strong reasons (at least under certain reasonable assumptions). Indeed, they may be identical to normative requirements governed by a defeasible ought. Or, as is more likely, the presence of a rational requirement may make it the case that we defeasibly ought to conform to the requirement.

## 2.4 THREE OBJECTIONS

In this section, I wish briefly to consider three objections to the argument for distinguishing between rational requirements and normative requirements (of the sort governed by an all-things-considered ought). The first objection is the not uncommonly held view that there are no genuine state-given reasons and that putative state-given reasons are in fact object-given reasons for having higher-order attitudes. The second objection is closely related to the first: that apparent state-given reasons for propositional attitudes are really ordinary reasons for bringing it about that you have the attitude in question. The third objection is that a dispositional account of rationality will not be subject to arguments that I have given.

The arguments here have depended on the existence of state-given reasons. Some people believe that there are no state-given reasons.[23] They take it that state-given reasons are really second-order object-given reasons. On this view, that some eccentric billionaire offers you one billion dollars to admire lazy Alex is not a reason to admire Alex, but rather it is an object-given reason to desire to admire Alex.

There are a variety of reasons for rejecting this view, and even for doubting the viability of the object-given/state-given distinction.[24] I shall just

[22] I make this claim in my doctoral thesis: see Reisner (2004). Broome arrives at a similar conclusion: see Broome (2005).

[23] See Parfit (2001).

[24] For a comprehensive discussion, see Rabinowicz and Rønnow-Rasmussen (2004).

offer one argument of my own as to why one might want to accept that there are genuine state-given reasons. Call this the *blocked ascent* argument.[25]

Suppose you are offered a large prize for admiring Alex, the lazy and humorless sadist. On the view that there are no state-given reasons, there will be only an object-given reason to desire that you admire Alex. To put this in a general way, to find the reason that you actually have, you must ascend to find a higher-order object-given reason. The problem with this view is that it is possible to block the ascent by changing the incentives in the example slightly. Instead of being offered the prize for admiring Alex, you are offered the prize for admiring Alex and for having *no higher-order attitudes* that have your admiring Alex as part of their contents. This last clause includes cases where your admiring Alex is deeply embedded, so that you lose the prize by desiring that you desire that you desire that (and so on) you admire Alex. Any ascent leads to your losing the prize.

In the case of blocked ascent, it is not at all clear why one would say that there is any reason given by the billionaire's offer at all, if one denies that there are state-given reasons. Because one has no object-given reasons for higher-order attitudes and because one, *ex hypothesi*, cannot have a state-given reason for the first-order attitude, there is apparently no reason generated at all. This may not seem like an absolutely objectionable conclusion in an eccentric billionaire example. However, it becomes much more objectionable in mad scientist cases, in which failure to admire Alex leads not to the loss of a mere monetary prize, but rather leads to a mad scientist committing an awful atrocity. In blocked ascent cases, those who think that there are no state-given reasons must accept that there is no reason to admire someone awful, when doing so is the only way to save the world. While someone who denies that there are any state-given reasons could bite the bullet, they do so at great cost.

The second objection is closely related to the first. Rather than taking state-given reasons to be object-given reasons for having higher-order attitudes, one could instead take state-given reasons to be object-given reasons for bringing it about that one satisfies what there is a putative state-given reason to intend or believe.[26] In the case of Alex the lazy sadist, one does not have reason to admire Alex. Rather, one has reason to bring it about that one admires Alex. Here again, there is a straightforward reply. One can rewrite the example such that the billionaire gives you the prize only if you admire Alex and do not bring it about that you admire Alex. What is

[25] A fuller version of the blocked ascent argument can be found in Reisner (2009a).
[26] Pamela Hieronymi (2005) suggests this account.

slightly less satisfactory with using this style of reply to answer the second objection is that it is not implausible that doing something ("do" here is just used as the universal verb and is not meant to imply an action verb *per se*) is a special case of bringing it about that one does something. One way to bring it about that you do something is to do it.

It is easy enough to come up with conditions for winning the prize that will block causal ascent; they will be of a piece with those that block attitudinal ascent. Still, there are two much more fundamental reasons to doubt the effectiveness of the "bringing it about" objection. The first reason is that if "doing" is a special case of "bringing it about that you do," then it appears that all saying that you have reason to bring it about that you admire Alex does is to add reasons, not to subtract them. One way to bring it about that you admire Alex is just to admire him. Of course, you could also bring it about by taking a special pill that will cause you to admire him. You have a reason to do either on the bringing-it-about view, since either is an instance of bringing it about.

The second reason is this: suppose, despite what I have just argued, you think that admiring Alex is not just a special case of bringing it about that you admire Alex. There will still be a worry deriving from the fact that "brings about" is a success verb. If you bring it about that you admire Alex, then it is an analytic necessity that you admire Alex. Those who support the bringing-it-about account of state-given reasons must also be committed to the controversial view that having a reason for $x$ does not give you a reason for what necessarily follows from $x$. If one does not deny this inference, then having a reason to bring about that one believes $b$ will entail that one has a reason to believe $b$.

The third objection I shall address is that one could interpret normative requirements dispositionally. On this view, the ought that governs a normative requirement says something about a way that people are set up, not about what the relations amongst their mental states are in individual instances. Here, the claim that *I ought not to x* is consistent with *I ought to have a disposition to x*. So, if I ought to have the disposition to have my mental states be thus and so, it is not inconsistent to say that in a particular instance, my mental states ought not to be thus and so.

This third objection is successful against the particular examples that I have given above, but eccentric billionaires can do what they like, and there is no reason why they cannot offer incentives not to have particular dispositions. If you get the prize by being disposed not to have the rationally required set of mental states, then the same type of conflict arises that rules out the view that putative state-given reasons are really higher-order

object-given reasons and the view that putative state-given reasons are really reasons to bring it about that one has a certain attitude.

## 2.5 THE NORMATIVITY OF THEORETICAL RATIONALITY

If the objections to the claim that rational requirements are normative requirements have been dealt with satisfactorily, we can draw two conclusions. The first is that rational requirements are not conceptually identical to normative requirements in which the ought is an all-things-considered ought. The second is that rational requirements do not entail normative requirements governed by an all-things-considered ought. However, this leaves room for the view that either a wide-scope reason or a wide-scope *defeasible* ought is identical to, or is entailed by, a rational requirement.

The argument for there being reason to be theoretically rational in one of these two senses is quite straightforward. What is difficult to determine is just what its precise upshot is. Let us consider a belief that one clearly has a normative reason not to believe, assuming that there are evidential reasons for belief. In the formalization, I shall use "N" as the normative reason operator:

E5. You have a reason not to believe $p$ and not $p$

F5. $\mathrm{N} \sim [\mathrm{B}\,(p \;\&\; \sim p)]$

Assuming classical logic, the probability that $p \;\&\; \sim p$ is false is 1. So, there is a perfectly straightforward epistemic reason not to believe a contradiction: it is certainly false.

Rational requirements govern relations amongst an agent's beliefs. There is, for example, a rational requirement not to have contradictory beliefs. In the formalization, I shall use "RR" as the rational requirement operator:

E6. You are rationally required not (to believe $p$ and to believe $\sim p$)

F6. $\mathrm{RR} \sim (\mathrm{B}p \;\&\; \mathrm{B} \sim p)$

We can write up F6 as a rational requirement of the traditional form, using a material conditional:

F6a. $\mathrm{RR}(\mathrm{B}p \rightarrow \sim \mathrm{B} \sim p)$

Is there a good epistemic reason to comply with the rational requirement in F6? The answer looks to be that there is, on just the same grounds that there is reason not to believe a contradiction. The probability that both

of your contradictory beliefs are true is zero, and this is a perfectly good epistemic reason not to have both beliefs simultaneously.[27]

This line of argument can be extended quite straightforwardly to any requirement of theoretical rationality, as long as it is a consistency requirement. Consider just one more rational requirement as an illustration:

E7. You are rationally required not to believe $\sim q$ if you believe $p$ and if $p$ then $q$.

F7. $\mathrm{RR}\{[\mathrm{B}p \;\&\; \mathrm{B}(p \rightarrow q)] \rightarrow \sim \mathrm{B} \sim q\}$

We can see again that the probability that $\{p \;\&\; (p \rightarrow q) \;\&\; \sim q\}$ is false is 1.[28] This provides a perfectly good epistemic reason to comply with F5. Of course, there is no reason to comply with F7 in any one particular way rather than another, just so long as one complies with it in one of those ways.[29] To put things more formally:

E8. When there is a wide-scope rational (consistency) requirement, there is a wide-scope normative epistemic reason.

F8. $\mathrm{RR}\{[\mathrm{B}p \;\&\; \mathrm{B}(p \rightarrow q)] \rightarrow \sim \mathrm{B} \sim q\} \rightarrow \mathrm{N}\{[\mathrm{B}p \;\&\; \mathrm{B}(p \rightarrow q)] \rightarrow \sim \mathrm{B} \sim q\}$

The reasons to conform to a requirement of rationality are ordinary, truth-directed, epistemic reasons. If they have any distinctive feature, it is that they are wide-scope. This feature of the kind of reasons there are to be theoretically rational is very similar to a feature of Broome's original normative requirement account of rational requirements: there is something distinctive about the form of the normativity (it is wide-scope), but there is nothing distinctive about the kind of normativity.

The argument that there is reason to be theoretically rational relies on two tacit premises, either of which might be false. The first is that there are epistemic or evidential reasons of the kind that I am relying on. I have been assuming that if $p$ is evidence for $q$, then at least under suitable

[27] One might want to factorize the explanation in this way. The probability of $p$ given $\sim p$ is 0, and the probability of $\sim p$ given $p$ is 0. This is perfect evidence that both beliefs are not true at the same time.

[28] One could factorize this argument the same way. The conditional probability that any one of the three beliefs is true, given that the other two are, is 0. So, that provides clear evidence that the conjunction of all three beliefs is false.

[29] Note that the claim here is consistent with Gilbert Harman's point that knowing that one should reject a proof does not entail that one ought to disbelieve any of the particular premises. See Harman (1999).

circumstances, $p$ is a reason for you to believe $q$. This is a claim that one might deny. The second assumption is that we can straightforwardly derive wide-scope epistemic reasons in the way required by the arguments above. It behooves me to say something about both of these assumptions.

The first assumption is accepted (with varying specific constraints) by most philosophers who work on normative reasons for belief, or at least by most of those who are not radical pragmatists. Of course, one need not accept this claim. I do not know how to argue that evidence, under suitable circumstances, gives us reasons in a way that would be more persuasive than the basic intuition that it does so.[30] Rather than try to argue for this assumption, let me instead note that there is another assumption that will do the relevant work. This is the assumption that if one knows that $p$ is false, then one has a reason not to believe $p$. Granting that logically competent individuals know that contradictions are necessarily false, then one has a reason not to believe a contradiction. Accepting this assumption does not directly commit one to views about evidence giving us reasons.

As to the second assumption, the worry is that our first assumption alone is not enough to generate the wide-scope reason:

$$\text{F6b. } \mathrm{N} \sim (\mathrm{B}p \ \& \ \mathrm{B} \sim p)$$

Rather, it is only strong enough to get us:

$$\text{F5. } \mathrm{N} \sim [\mathrm{B}(p \ \& \sim p)]$$

Why this worry? We know that the complex proposition $p$ and not $p$ is false. That is good enough to give us F5. But, we do not know that $p$ is false, nor do we know that not $p$ is false. So, no single proposition in F6b is known to be false. What we need is an argument that allows us to derive F6b from F5.

It turns out that one cannot directly derive F6b from F5. One may still accept that the probabilistic argument offered above implies F6b; I believe that to be so. However, one can derive F6b from what I take to be a less controversial principle: evidence requires you not to believe $p$ and $\sim p$. Here is the derivation:

| | |
|---|---|
| 1. $\mathrm{ER} \sim \mathrm{B}(p \ \& \sim p)$ | (assumption) |
| 2. $\sim \mathrm{ER} \sim (\mathrm{B}a \ \& \ \mathrm{B}b) \rightarrow \sim \mathrm{ER} \sim [\mathrm{B}(a \ \& \ b)]$ | (assumption) |
| 3. $\mathrm{ER} \sim [\mathrm{B}(a \ \& \ b)] \rightarrow \mathrm{ER} \sim (\mathrm{B}a \ \& \ \mathrm{B}b)$ | (2, contraposition) |
| 4. $\mathrm{ER} \sim (\mathrm{B}p \ \& \ \mathrm{B} \sim p)$ | (2,3 *modus ponens*) |

[30] John Broome has expressed some concern about this assumption to me in correspondence.

Premise (2) of this argument relies for its plausibility on the view that we can agglomerate under evidential requirement. If you doubt this, then you will not be persuaded by this argument. Given what it is to be an evidential reason for belief, it should be uncontroversial to infer F6b from (4).[31]

## 2.6 The relationship between normativity and theoretical rationality

We now know that the presence of theoretical rational consistency requirements implies that there is reason to conform to those requirements, but we are left with several outstanding questions. In this section, I want to explore whether there is any conceptual relationship between theoretical normativity and theoretical rationality. I shall make the hedged claim that there may be. In the next section, I shall look at whether this argument has any implications for the normativity of practical rationality, and I shall argue that it may, but that much more work is required to spell out how.

To begin with, it makes sense to consider whether the reverse of $F_4$ is true. That is, whether the presence of a wide-scope reason governing a group of an agent's beliefs implies, materially or otherwise, that there is a theoretical rational requirement on that agent. There are some cases that suggest that there can be wide-scope reasons governing an agent's beliefs that do not imply a rational requirement. These reasons are of two kinds, although a further distinction will be required between different categories of the second kind.

The first sort of reason is a pragmatic reason – a reason to believe something because it would be good for you to believe it. I invoked such reasons earlier in arguing that rational requirements are not normative requirements (in the sense of being wide-scope oughts). One might have pragmatic incentives, for example, to hold contradictory beliefs. That does not make holding contradictory beliefs rational.

We might want to make a distinction between pragmatic and epistemic reasons.[32] Epistemic reasons are in some sense "truthy" reasons – reasons that have something to do with having or getting true beliefs. Construed broadly enough, we might include amongst the epistemic reasons those reasons that come from the advantages they confer to us in increasing the number of interesting true beliefs that we have, even if they do not point to the truth of the specific beliefs that they are for.

[31] Earlier versions of this argument had serious flaws. John Broome suggested the argument used here. Johannes Stern pointed out the hopelessness of my earlier approach.

[32] See Harman (1999).

Peter Railton[33] suggests as an example considerations of research productivity. It may make sense to continue believing a theory in the face of evidence to the contrary, because the research and thought that flow from one's belief are producing large numbers of interesting true beliefs (or facilitating their production). Whether this is a properly epistemic reason may be controversial. We need not settle the matter here. If such reasons are epistemic reasons, then something's being a wide-scope epistemic reason governing a collection of beliefs is insufficient for its being a rational requirement.[34] This is because there could be a wide-scope epistemic reason to believe the contents of a theory that is logically inconsistent.

What we are left with is a narrower possibility: that a wide-scope evidential reason gives us a rational requirement. This possibility looks more promising, but I suspect that it is wrong. I had not been able to think of a counter-example, but one was suggested to me. If there is a normative reason not to have inconsistent beliefs, and if rational requirements just are reasons of this kind, then one would be irrational any time one had inconsistent beliefs. This is far too demanding, and we may want to limit the scope of rational requirements to occurrent beliefs, or relevant beliefs, or some other restricted class of an agent's beliefs. This suggests that rational requirements are not identical to wide-scope normative reasons for belief.[35]

If the identity claim does not hold between wide-scope evidential reasons and theoretical rational requirements, there is a strong conceptual link. What qualifies a requirement governing relations amongst an agent's beliefs as a requirement of rationality (as opposed to some other kind of requirement) is that it is a requirement concerning the consistency of the contents of the agent's beliefs. Because pairs (or larger groups) of inconsistent beliefs cannot all be true, there will also be wide-scope reasons in favor of not having inconsistent beliefs.

Further, the reason not to have inconsistent beliefs is of a particularly strong sort of evidential reason, because the conjunction of one's inconsistent beliefs is certain to be false. This may explain why requirements of theoretical rationality may appear, at first blush, to be normative requirements. Because under normal circumstances evidential reasons dominate non-evidential or pragmatic reasons for belief, there is overriding reason to be theoretically rational. Indeed, in general, one ought to be theoretically

[33] See Railton (1997).

[34] I am not sure that such reasons are normally wide in scope, but it remains a possibility that at least some are.

[35] This suggestion was made to me by John Broome.

rational. The defeasibility of this ought is explained by the way in which evidential and pragmatic reasons for belief weigh against each other.[36]

We are left with the interesting result that the requirements of theoretical rationality, at least insofar as they are consistency requirements, are inherently normative, although this is not necessarily because the concept *rational requirement* entails the concept *reason*. We can also now see why it was initially tempting to analyze rational requirements as being normative requirements, and that that analysis was not very far off the mark for theoretical rationality; in general, the requirements of theoretical rationality do provide us with normative requirements. At minimum, we are left with an explanation of why the requirements of theoretical rationality have the ability to exert normative pressure on us; there is a reason to be theoretically rational, at least for wide-scope consistency requirements.

## 2.7 LESSONS FOR PRACTICAL RATIONALITY?

Nothing so neat as the story we can tell for theoretical rationality can be told for practical rationality. Important practical rational requirements have nothing to do with consistency, and even those that do may not lay claim to the normativity that came with theoretical rational requirements. If there is a story to be told about the normativity of practical rationality, it will be a quite different one.

We can turn first to a requirement that might be called "*enkrasia*." *Enkrasia* tells us that:

> E7. You are required to intend to $\varphi$ whenever you believe that you ought to $\varphi$.
>
> F7. $RR(BO\varphi \rightarrow I\varphi)$[37]

There is no straightforward way of reading *enkrasia* as a consistency requirement.[38] If there is some normativity to *enkrasia*, it does not obviously derive from consistency.

[36] See Reisner (2008) for a more detailed account of how pragmatic and evidential reasons for belief can be weighed. Danielsson and Olson (2007) make a similar point about weighing reasons for other propositional attitudes.

[37] This is a view with a long history. A detailed account of its importance can be found in Wedgwood (2007a). He calls the view "normative judgement internalism." Another important recent discussion is in Broome (2007a).

[38] Wedgwood has argued that it is implied by a particular version of conceptual role semantics about normative terms. See Wedgwood (2007a). Its generation by Wedgwood's conceptual role semantics might count as a kind of consistency, but Wedgwood's view is controversial.

Some practical rational requirements are clearly underpinned by considerations of consistency. A weakened version of the instrumental principle is one such case:

> E8. You are required not to intend not to take what you believe to be the necessary means to the ends that you intend.
>
> F8. RR{[I$\varphi$ & B ($\varphi \rightarrow \psi$)] $\rightarrow \sim$I $\sim \psi$}.

The weakened instrumental principle requires the contents of one's relevant beliefs and intentions to be consistent. So, in principle, one could say that one has overwhelming evidential reason to follow the weakened instrumental principle. It is difficult to understand what evidential reasons for intentions or actions are, however, and evidence is not likely to be a basic reason-giving norm of practical reason. So, evidence does not obviously make the instrumental principle normative. There may be other ways to argue for the normativity of consistency-based principles of practical reason. Perhaps the near certainty of failure to carry out one's intentions, unless one conforms to the instrumental principle, gives the appropriate sort of practical reason. But this sounds pragmatic in a way that is not the case for theoretical rationality, and perhaps we should be skeptical of this explanation.

However, even if consistency does not look poised to play the same role in giving normativity to practical rational requirements as it does to theoretical rational requirements, consistency requirements and even *enkrasia* may be normative due to related considerations. We might think that there is a broader notion of consistency than the strictly logical one that we hold as an ideal for our mental lives.[39]

On this view, we might explain object-given reasons for propositional attitudes in terms of ways our mental states ideally should relate to their content. For example, object-given reasons to believe something, on this view, are evidential reasons because it is an ideal of our beliefs that they are regulated by truth. Object-given reasons to intend something are identical to the reasons there are to perform the intended action, because intentions should be regulated by consideration of the goodness of the intended actions. Whether this story about object-given reasons is correct, and whether it is sufficient to show that the requirements of practical rationality are normative, remains to be seen. However, it provides at least one possible path for developing not only an account of normativity of practical

[39] Wlodek Rabinowicz and Philip Pettit both suggested this to me.

rationality, but a unified account of the normativity of all rationality. I am skeptical about the prospects for developing an account of the normativity of practical rationality in this way, but it is at least one avenue that might be explored.

## 2.8 conclusion

The question about theoretical rational requirements left unsettled in this chapter is whether they might be identified with defeasible normative requirements, or whether they only imply defeasible normative requirements. This question could be settled decisively in the negative by finding wide-scope reasons governing beliefs that are not rational requirements. Given that wide-scope evidential reasons "overshoot" what we intuitively take to be rational requirements, there is good reason to think that there are wide-scope evidential reasons governing beliefs that are not rational requirements. To say something decisive about this, we need a fuller account of the demandingness of rational requirements.

However this question is settled in the end, we are left with the interesting, and perhaps surprising, conclusion that theoretical rationality, or at least that part of it which concerns consistency, is in fact normative. And this is so even if we make a conceptual distinction between rationality and normativity. Reasons for belief and the rationality of belief both share a deep underlying regulatory conceptual framework.

In the past it has been considered desirable to try to build a theory of practical rationality from our understanding of theoretical rationality, the latter providing a model, at least structurally, for the former. The prospects for developing a strategy to show that practical rational requirements are normative using a parallel model to that employed for theoretical rationality are mixed at best. The foundational role that evidence plays both in giving us reasons for belief and in giving us theoretical rational requirements is not obviously matched by anything in practical rationality. Whether an appeal to ideals about our mental life unites practical and theoretical rationality remains to be seen.[40]

[40] This chapter in its final form has come into being due to comments from a number of philosophers. Particularly helpful were John Broome, Pascal Engel, Daniel Laurier, Kevin Mulligan, Jessica Pepp, Philip Pettit, Wlodek Rabinowicz, Toni Rønnow-Rasmussen, Robert Stephens, Asbjørn Steglich-Petersen, Johannes Stern, Sarah Stroud, and an anonymous referee for this volume. I received valuable feedback in a group setting at the Higher Philosophy Seminar at the University of Lund and from those in attendance at the Danish Philosophical Society Workshop in Epistemology.

CHAPTER 3

# *Epistemic motivation*
## *Towards a metaethics of belief*

*Veli Mitova*

### 3.1 INTRODUCTION

Present-day epistemology abounds in normative talk. The talk is distinctive insofar as it concerns beliefs rather than actions. But on the face of it, it sounds just like metaethical talk: we think about epistemic *reasons*, we quarrel about epistemic *norms* and *virtues*; we celebrate the *Value* Turn; we sport an *ethics* of belief. Not only do epistemologists extensively use metaethical vocabulary, but they have become increasingly concerned with analysing epistemic normative talk in light of some standard metaethical questions. They have been forging, in other words, what we might call 'a metaethics of belief'.

The ongoing debate about the justification of epistemic normativity, and the derivative debate about whether there are any categorical epistemic norms and reasons, are clear examples of epistemologists engaged in a metaethics of belief.[1] Another is the fledgling literature on the nature of epistemic normative judgements. Terence Cuneo (2007), for example, has recently argued for cognitivism and realism about these judgements. And two of the contributors to this volume – Olson and Shah – are concerned with error theories about epistemic judgements, and so are also at work on a metaethics of belief.

Despite helping themselves to the metaethicist's vocabulary and sharing some of his concerns, however, epistemologists have been curiously uninterested in applying particular well-rehearsed metaethical questions to epistemology. Two such questions concern motivation: how should we

[1] Two broad kinds of justification are floated in the literature. Pragmatists justify epistemic norms on instrumental grounds: observing these norms is the best means we have for getting at truth; and true beliefs are necessary for attaining our goals (e.g. Kornblith 1993). On this view, all epistemic norms and reasons turn out to be hypothetical. By contrast, friends of the normativity of the mental argue that the concept of belief itself provides a justification for epistemic norms: since the concept itself is normative, applying it to one's attitude involves endorsement of epistemic norms (e.g. Wedgwood 2002a). On this view, some epistemic norms and reasons are categorical.

understand epistemic motivation? and what is the relationship between making an epistemic normative judgement and being motivated by it?[2] Needless to say, the lack of interest in these questions has left our metaethics of belief somewhat potted, to put it mildly. The aim of this chapter is to set the programme for building up a metaethics of belief from the unexplored angle of motivation. The advantages of this approach will prove to be significant: firstly, we will end up with a more comprehensive metaethics of belief for developing an account of epistemic motivation; and secondly, we will have a fresh approach to settling loftier metaethical issues such as the nature of epistemic normative judgements.

The chapter proceeds as follows. First, it introduces and transposes to epistemology three standard metaethical debates – internalism versus externalism about the connection between moral judgements and motivation; Humeanism versus anti-Humeanism about the nature of motivation; and cognitivism versus non-cognitivism about moral judgements (section 3.2). I then defend internalism (sections 3.3–3.5) and anti-Humeanism (sections 3.6–3.7). I finally show how this position on motivation affects our commitments in the higher, more metaphysical, realms of our metaethics of belief: in particular, it makes cognitivism about epistemic judgements *prima facie* plausible (section 3.8). The arguments I develop are programmatic. They neither are, nor are meant to be, decisive. I am assuming – and the vast metaethical literature supports me – that these issues are too big to be conclusively settled in anything short of a book.

## 3.2 THREE DEBATES IN METAETHICS TRANSPOSED TO EPISTEMOLOGY

### *3.2.1 The debates*

Three debates largely define contemporary metaethics.[3] The first is between internalists and externalists. The question which drives this debate is whether there is a necessary conceptual connection between making a moral judgement and being motivated to act in accordance with it. Internalists think that there is such a connection. They defend some version of the claim that making a normative judgement, such as 'It is wrong to

[2] Sometimes justifications of epistemic normativity care about motivation, but only normatively – they care about why we *should* be motivated by epistemic norms. In this sense, they remain in the upper reaches of a metaethics of belief.

[3] This description of the debates is loosely based on Smith (1994). But this way of describing the debates has by now become common currency.

burn the cat for fun', would necessarily motivate a rational person. Now, many axes have been ground on this claim, so it has sprouted many and subtle versions, but I will keep it simple here. The internalist's idea is that if there were no necessary connection between a moral judgement and motivation, moral judgements could neither guide nor explain our actions.[4] The externalist rejects this necessary connection. He thinks that I can be fully rational, make the judgement that it is wrong to burn the cat, and be unmoved by this judgement.

The second debate is between Humeans and anti-Humeans about motivation. The question here is whether being motivated is necessarily a matter of having some conative state such as desire, or whether a cognitive state like belief, on its own, can motivate a person to act. (From now on I will generally say 'belief' and 'desire', for short.) Humeans say that a desire is a necessary component of motivation; anti-Humeans deny this. So, for a Humean, unless I had some appropriate desire – say to burn the cat, or a desire for something that would be achieved by burning the cat – I could not be motivated to burn it. The anti-Humean, by contrast, thinks that a certain belief about burning the cat – perhaps that doing so will save lives – is sufficient to motivate me to burn it in the absence of a desire to burn it or to save lives.

The third debate is between cognitivists and non-cognitivists about moral judgements. The question at issue here is whether moral judgements express cognitive states such as beliefs or conative states such as desires. To oversimplify again, the cognitivist thinks that moral judgements express cognitive attitudes: the judgement that it is wrong to burn the cat for fun expresses a belief. The non-cognitivist denies this. According to him, moral judgements are expressions of conative attitudes towards the object of the judgement. So my judgement about the cat amounts to disapproval, or dislike, of cat-burning for fun.

The three debates are interdependent. Suppose we started from the issue of motivation and worked our way up to the nature of normative judgements. This is the standard non-cognitivist move: if we are (i) Humeans about motivation and (ii) internalists about the connection between moral judgements and motivation, we then think that (i) only desires can motivate, and (ii) moral judgements necessarily motivate. Non-cognitivism follows: moral judgements must be desire-like rather than belief-like. In general, settling any two of the debates suggests a position on the third.

[4] See, for example, Williams (1980).

### *3.2.2 The debates transposed to epistemology*

These are the three debates I aim to transpose to epistemology. But before I do so, I would like to put to rest a possible doubt concerning the very plausibility of the transposition. The doubt starts with noting a disanalogy between belief and action: action is voluntary; belief is not. From here the doubt can grow in either of two directions. One is that belief's involuntariness renders nonsensical *any* normative talk about beliefs, and *a fortiori*, renders (even more) nonsensical metaethical analyses of such talk. I have no response to this gloss of the doubt. The present chapter is addressed to those who think that we rightly have an ethics of belief. It is not meant to convince sceptics about epistemic normativity.

The other gloss of the doubt is to concede that normative talk about beliefs makes sense, but to urge that, given the disanalogy between belief and action in this basic respect of voluntariness, the *particular* questions we ask in metaethics proper are irrelevant to a metaethics of belief. For lack of space, I can only offer here three brief suggestions as to why this doubt is misplaced. First, action is no more voluntary than belief in a totally brute unconstrained sense of voluntariness. In order to act rather than behave, I need to do so for reasons. So the objection exaggerates the disanalogy between belief and action. Second, paradigmatic moral judgements are not only about actions; they are also about emotions, moral sensibilities, and other attitudes, which are no more under our *direct* control than are our beliefs. So, even if there were a drastic disanalogy between belief and action, it would not open a wide enough gap between moral and epistemic *judgements*. Finally, even the staunchest doxastic involuntarists grant that we have *indirect* control over our beliefs – over how attentive we are to the evidence, to inconsistencies, etc. And this is all we need for normative talk about our beliefs to gain purchase. For to concede any control is to concede control over whether our beliefs go well or badly. And as soon as normative talk gains purchase, questions about the nature of judgements featuring this talk, about its connection to motivation, and about the nature of motivation itself become important.

So, assuming that it is not implausible to ask the three standard metaethical questions in the epistemic context, what will the questions look like? To begin with, what are epistemic normative judgements? Epistemic normative judgements, or 'epistemic judgements' for short, form as large and diverse a category as do moral judgements. They can be about what one ought to believe (for example, 'If I am to have a belief about $p$, I ought to believe in accordance with the evidence concerning $p$'). They can be

about what one has reason to believe (for example, 'That *p* is inconsistent with *q* gives me reason to abandon either my belief that *p* or my belief that *q*'). They can concern epistemic values and virtues ('When I am trying to form a belief about *p*, ignoring a *p*-expert's testimony is epistemically irresponsible'), they can be evaluations of beliefs one already holds ('This belief is unjustified'), and so on.

What is common to all of these judgements is, firstly, that they feature some normative term; and secondly, that this term concerns the *procedures* by which one comes to have or to revise one's beliefs, rather than the beliefs' *contents*. This is not to deny that we can evaluate our beliefs from any point of human concern. The idea just is that when we evaluate them from a point of view other than the epistemic, we are generally concerned with their contents: regardless of how the belief was formed, we can evaluate it according to whether having a belief with this content is a morally good thing (say, the belief that women are intellectually inferior); or a prudentially good thing (say, the belief that my partner is unfaithful), and so on. *Epistemic* normative judgements, by contrast, are blind to the content of the belief. They concern, instead, the belief's formation and revision, or the evaluation of already existing beliefs. These are the judgements to which I apply the three standard metaethical questions.

Internalism in the epistemic context,[5] then, would be the view that there is a necessary connection between making an epistemic judgement and being moved to believe in accordance with it. The externalist would deny this necessary connection. Anti-Humeanism about epistemic motivation would be the view that no desire is necessary for a belief to occur. The Humean would think a desire necessary. Cognitivism would be the view that epistemic judgements are cognitive in nature. The non-cognitivist would think they are conative.

## 3.3 INTERNALISM ABOUT EPISTEMIC JUDGEMENTS: PRELIMINARIES

In this section I set the scene for defending internalism about how epistemic normative judgements motivate us. In the next two sections I defend it. I should note, though, that I am not *dialectically* required to provide an argument for internalism. One of the central aims of this chapter is to show that loftier metaethical questions can be answered from a stance on

[5] This position should not be confused with classical *epistemic* internalism and externalism. The latter is a position about the kinds of things which *justify* a belief.

motivation. One lofty result which my stance on motivation yields is (some version of) cognitivism about epistemic judgements. Since internalism is the classical motivation for non-cognitivism, I could just help myself to internalism without begging any relevant question. All the same, having an independent argument for internalism would make for a richer metaethics of belief. And developing such a metaethics at the level of motivation is the other aim of the chapter. But those who are unimpressed by my argument for internalism should still accept both that we can harvest important upper-level metaethical results from a stance on epistemic motivation and the particular result harvested here, some form of cognitivism.

In the epistemic context, recall, internalism says that a rational person is necessarily motivated by his epistemic judgements. Externalism says that one can make an epistemic judgement, be unmoved by it, and remain rational. The internalist thesis which I will defend here is that we are necessarily motivated by our epistemic normative judgements *insofar as we are creatures with intentions*. The argument for this claim is this:

| | |
|---|---|
| (P1) | Being curious about some aspect of the world, A, constitutively requires that one is motivated by one's epistemic judgements when forming beliefs about A. |
| (P2) | Having intentions constitutively requires that one is curious about the world at large. |
| (C) | So, having intentions constitutively requires being motivated by one's epistemic judgements. |

I have defended elsewhere (2009) a version of this argument. All I hope to do here is to make it *prima facie* plausible. Before I launch into the details, a word about the argument's central terms and scope. First, 'curiosity' is a term of art: being curious about something means *wanting to know*, or *wanting to find out*, how this thing is. It is a conative attitude with an epistemic object. It is, moreover, a *serious* attitude, which *typically* motivates the agent to find out about the object of curiosity. But of course, it does not result in action infallibly: an agent has many other attitudes, which can override curiosity. So I can want to find out about something and yet not act on this curiosity due to other greater wants, or stronger reasons for doing something incompatible with finding out.

Second, being *motivated*, or *moved*, by a normative judgement means that one either tries to comply with it, or recognizes that one has done something wrong in not trying to comply with it. Third, having an *intention*

to bring about x is having, at a minimum, a pro-attitude to two states of affairs obtaining:

**(a)** that x comes about;

and **(b)** that it is I who brought x about.

The argument is confined to this kind of intention, which involves a *commitment to action* on the part of the agent. (The category includes intentions to maintain or prevent x, but I will talk of intentions to bring about x, for short.) There may be other kinds of intentions, perhaps intentions *that* a certain state of affairs obtain, which do not commit the intender to action. But the present argument has nothing to say about whether such attitudes are legitimately classed as intentions, or whether they involve curiosity. This caveat is crucial: as I show in section 3.5.1, the agent's commitment to action is precisely what gives rise to his curiosity.

Now for the scope of the argument. The argument works from the *concepts* of intention and curiosity.[6] The claim is that it is incorrect to attribute to a person curiosity about A if he is not moved by his epistemic judgements concerning how to form beliefs about A (premise 1), and incorrect to attribute to him an intention if he is not curious about the world (premise 2).

## 3.4 INTERNALISM ABOUT EPISTEMIC JUDGEMENTS: PREMISE 1

This section defends premise 1 of the overarching argument: being curious about aspect A of the world constitutively requires that one is motivated by one's epistemic judgements when forming beliefs about A.

A comment on what the premise is *not* claiming: the claim is not that I will infallibly *form* beliefs in accordance with my epistemic judgements. The claim is, rather, that if I am curious, then I am infallibly *moved* by my epistemic judgements – if I did not try to comply with them, I would recognize at some level that I am in the wrong, as per the definition of being motivated in the previous section.

It seems trivial, given the definition of curiosity, that one cannot be genuinely curious about some aspect of the world A, and not be moved by at least a class of one's epistemic judgements. The relevant class is judgements featuring norms which prescribe particular procedures – like observing the evidence or avoiding inconsistencies – for forming and revising one's beliefs

[6] Note, though, that the argument is neutral on whether one needs to have these concepts in order to be credited with having the corresponding attitudes.

about A. For being curious about A is *wanting to know* or *find out* how A is. And one's epistemic norms-judgements represent precisely one's ideas about how one goes about the business of finding out something one wants to know. Deliberately not being moved by my *own* epistemic judgements about how I should form beliefs about A is being wilfully blind to the world with respect to A. Such wilful blindness just *is* not wanting to know how the world is with respect to A. And this, of course, is the opposite of curiosity about A. Being curious just is being motivated by one's epistemic judgements.

Two points are worth emphasising here.[7] First, I do not mean that being unmoved by a *single* judgement is being wilfully blind to the whole world. Rather, being unmoved by my judgement concerning how I should believe with respect to A is being wilfully blind to the A-portion of the world. But naturally in the limit, if I am unmoved by *all* of my judgements about how I ought to form each of my beliefs, then I am wilfully blind to the whole world.

Second, and to reiterate, the claim about the connection between curiosity and motivation concerns only judgements about how I ought to *form* and *revise* my beliefs. It does not apply to judgements that I ought to believe as I already do. Say that I have believed for two years that Obama won the 2008 election. Suppose that today for some reason I judge that I ought to believe just that – that Obama won the 2008 election. Obviously there is no sense to the idea that this judgement will motivate me to be curious about anything. But this is as it should be: premise 1 does not claim that if I make an epistemic normative judgement, then I am curious. It is the converse claim that if I am curious about Obama, then I am moved by my judgements about how to form beliefs about him. In the case where I make the judgement that I ought to believe as I already do I am not curious about that portion of the world. I've already made up my mind about it. Correspondingly, the judgement that I ought to believe as I already do cannot – and cannot be expected to – motivate me or make me curious.

This establishes premise 1 of the overarching argument for internalism, at least for a limited class of epistemic judgements – judgements about how I ought to form and revise my beliefs. More sophisticated epistemic judgements – concerning one's epistemic responsibility, or epistemic character, say – might prove trickier. If so, I will rest content with a limited form of internalism (but see the end of section 3.5.2).

[7] Thanks to an anonymous CUP referee for pressing me on these points as well as for the Obama example.

## 3.5 INTERNALISM ABOUT EPISTEMIC JUDGEMENTS: PREMISE 2

I now turn to premise 2 of the overarching argument for internalism. The premise states that having an intention constitutively requires that one is curious about the world. I will defend it in two steps:

**3.5.1** Having an intention constitutively requires that one is curious about the intended outcome.

**3.5.2** Being curious about the intended outcome constitutively requires that one is curious about the world.

### *3.5.1 Having an intention constitutively requires that one is curious about the outcome*

Recall (section 3.3) that I understand an intention to bring about x as (at a minimum) a pro-attitude to two states of affairs obtaining:

**(a)** that x comes about;

and **(b)** that it is I who brought x about.

So I cannot have an intention unless *I want these states of affairs to obtain*. The first step in establishing premise 2 is to make the further claim that I cannot have the intention unless I also *want to find out* whether these states of affairs will come to obtain once I have acted on the intention. I become curious, that is, at the time of forming the intention, about two states of affairs:

**(a)** whether x will come about;

and **(b)** whether it will be I who brought x about.

The intuition behind this claim is that x coming about, and coming about due to me, are the only conditions under which the intention is fulfilled, and so the typical conditions under which it is relinquished. Since part of what it is to have an intention is that one relinquish it when one finds out that it is fulfilled, one will want to *find out whether* it is fulfilled. So it is part of having an intention that one is curious, from the moment of forming the intention, about the intended outcome of one's actions.

Consider a thought experiment about someone who fails to have an intention, and fails to have it for no other reason but lack of curiosity about the outcome. I declare that I intend to fix the world's water crisis. A friendly water genie conjures two doors before me and says that going through one of them would fix the world's water crisis, while going through

the other would leave the world as it is. In both cases I will get the belief that I have fixed the crisis. And the false belief, he faithfully promises, would never be controverted by the world. He doesn't tell me which door is which now, but asks me to arbitrarily pick a door, but to choose *now* whether once I have gone through a door I want him to tell me which door was which. Let me time-index the scenario in order to make it clearer: I will arbitrarily pick a door at $t_1$, go through it at $t_2$; at $t_o$ the genie asks me whether I want to [know at $t_4$ which door was which]. He is asking me, that is, whether I am now curious about (want to know) if my intention will be fulfilled when I have acted.[8] Suppose I said that I didn't care to know. Can I be credited with a genuine intention to fix the water crisis?

I think not. Finding out that I have fixed the crisis would make me relinquish the intention to fix it. If I went through a door and the genie told me it was the crisis-fixing door, I would relinquish the intention. If he told me it was not the crisis-fixing door, I would retain the intention. But since I am indifferent to finding out, I am indifferent to what would be the relinquishment-conditions for the intention of fixing the water crisis. So I can't have that intention.[9] If an intention is to explain my action of going through a door, it must be an intention which I think will be fulfilled by the action. So, it must be the intention *to get the belief that I have fixed the water crisis*. Getting this belief is the only common feature between the two doors. If I am indifferent between them, to the extent that I don't care to learn which one was the water-fixing one, I am indifferent to their distinguishing features, such as, significantly, the feature of fixing the water crisis. Of course, getting the belief (whether true or false) that I have fixed the water crisis will make me relinquish the intention to fix it, if I had one. But I know now, at $t_o$, that this would be a spurious relinquishment, since I will not know at $t_4$ whether the intention is fulfilled.

This intuition is corroborated when we recall the kind of intentions that concern me here – intentions that commit the intender to action. When I form an intention *to do* something I commit to *seeing it out*, so to speak. 'Seeing it out', 'seeing to it', and 'seeing it through' are expressions which

[8] We will obviously need plenty of memory cleansing here. If I chose to know afterwards which door was which, and he told me I went through the wrong door, he will have to make me forget that, so that I get the belief that I have fixed the water crisis as he promised.

[9] This conclusion can't be blocked by saying that I have an intention to fix the crisis, but it is an irrational intention due to my not having taken the obvious means to my end – going through the right door. The experiment is designed to rule out this move: since I don't have the option of knowing which door is the right one *before* I go through a door, I have *not* failed to take the means to my end. I take the best means available – go through either door.

capture a constitutive feature of intention: that it involves – as the seeing metaphor suggests – seeing *whether* one has seen it through or out.

It may be objected, however, that there are clear cases of intending in which we believe that it is impossible to find out about the outcome, and in which, therefore, we cannot be curious about the outcome.[10] For example, I buy life insurance with the intention that if I die before I stop working, my family will be provided for. But, assuming that I am an atheist, I believe that it's impossible for me to find out what happens to my family after I die. And so it does not make sense for me to *want* to find out. This, the objection goes, provides a counter-example to the Constitutive Claim, a case in which I have an intention but am not – and cannot be – curious about the outcome.[11]

The reply, which I have developed at greater length elsewhere (see footnote 10) is this. Distinguish three intention-candidates which may be involved in the example:

**(i)** the intention *that* my family are provided for after my death;
**(ii)** the intention to *do my best to ensure* that my family are provided for after my death;
**(iii)** the intention to *absolutely ensure* that my family are provided for after my death.

The three intentions have different fulfilment conditions, and so differ with respect to the possibility of finding out whether their outcomes obtain. Intention **(i)** – if intention it be – is an intention whose outcome I cannot find out about, and so seems eligible for a counter-example. It does not, however, involve a commitment on the part of the agent to act. Yet I restricted my argument when I introduced its general shape (section 3.3) only to those intentions which do involve such a commitment. The Constitutive Claim is not a claim about other pro-attitudes.

I can find out the outcome of intention **(ii)** – I can check on various insurers' records with paying out insured families and choose the best, get the appropriate legal information, and so on. If I do these things, then I will have succeeded in doing my best to ensure that my family are provided

[10] Thanks to an anonymous CUP referee for this objection and for the counter-example. The example involves my belief that it is *physically* impossible for me to find out. For putative counter-examples involving logical impossibility, as well as a longer treatment of physical impossibility, see Mitova (2009: 124–127).

[11] The original objection by the CUP referee included the claim that it is a problem for my account that I also don't want to die before I stop working. But this desire does not affect the Constitutive Claim, which only says that I want to *find out* whether if I die my family gets the money. It says neither that I want to die nor that I want my family to get the money (the latter desire will be an uncontroversial part of my intention).

for. But these are things that I can, and can want to, find out. So intention **(ii)** is not a counter-example to the Constitutive Claim either.

Intention **(iii)** looks like another potential counter-example to the Constitutive Claim. I believe that after I die, I cannot find out whether I have succeeded in guaranteeing that my family are provided for, and so I cannot be curious about finding out. The bad news for the objector, though, is that the attitude to this guarantee cannot be one of intention, not because of any reasons to do with curiosity, but because the attitude breaches a broadly Davidsonian requirement on intention. The requirement is that the agent not *believe* the intended outcome to be impossible. If he does believe the outcome impossible, then he can wish for it, but not intend it. Yet in the example I believe (and must believe if I am rational) the outcome impossible: There is no action that I can perform in order to ensure in the absolute sense that my family are provided for. After I die, I cannot act. So I cannot intend to guarantee that my family are provided for.

Whatever we might think in general about this Davidsonian requirement, it is at any rate accepted by the objector, indeed it drives his objection: It was precisely because I believed it impossible to find out about the intended outcome after I die, that the objector could say that it is impossible for me to *want* to find out. And if he is happy with this requirement for the *weaker* pro-attitude of desiring, then he must accept the requirement for the stronger pro-attitude of intending.

So the reply to the objection is that intentions involved in insuring my life do not constitute a counter-example to the Constitutive Claim. Either (i) they are intentions which the Constitutive Claim does not purport to cover; or (ii) they are intentions whose outcome we can, and do, want to find out about; or (iii) the attitudes are not intentions because they breach a Davidsonian requirement on intention.

### *3.5.2 Being curious about the intended outcome constitutively requires that one is curious about the world*

So, having an intention constitutively requires that I am curious about whether the outcome comes to obtain. Some brief considerations from the holism of the mental should suffice to extend this claim to curiosity about the world at large, and so to get us premise 2.

Beliefs are not isolated mental items but come in networks, the members of which partake in complex inferential patterns. There will be a two-way inferential traffic between one's beliefs about intended outcomes and one's other kinds of beliefs. So beliefs about outcomes will be *inferentially*

*dependent* on other kinds of beliefs one has. And given the complexity of belief systems, one will often not know in advance which of one's beliefs about the world will turn out to be relevant to one's beliefs about outcomes. So, one cannot want to know how things stand with intended outcomes without also wanting to know about *whatever* bits of the world are relevant to one's belief about the outcome. So, one cannot be curious about outcomes without being curious about the world.

This thought is confirmed when we consider what it would mean to be moved by one's norms-judgements for beliefs about outcomes alone. It is impossible for my belief that I have rid the world of the water crisis to be justified, without its inferential predecessors (which are not about outcomes) being likewise justified *by the same standards*. One's epistemic judgements feature this common standard. So, either one is moved by one's epistemic judgements for all of one's beliefs or for none. Since curiosity about outcomes explains motivation in the case of beliefs about outcomes, we have explained how epistemic judgements motivate us for all beliefs.[12]

### *3.5.3 Internalism*

If the thoughts in the last three sections are on the right track, internalism for epistemic judgements has been made plausible: it is impossible for us not to be motivated by our epistemic judgements as long as we have intentions. Of course, this does not mean that we will form our beliefs in accordance with these judgements in every single case. But it does mean that our judgements have a necessary motivational grip on us – when we don't try to form a belief in accordance with our judgements, we recognize that we are in the wrong.

The internalism we secure in this way is quite modest in at least two respects: it concerns only judgements about belief *formation* and *revision*;

[12] An important qualification is in order here. (Thanks to Asbjørn Steglich-Petersen for this point.) I have talked of curiosity about the world as if every proposition about the world will be relevant to our beliefs about outcomes. This keeps the argument fairly straightforward but is inaccurate. Of course, for each of us there is a (dynamic) class of propositions, in which we don't feel the slightest curiosity and yet we remain creatures with intentions. Think of the proposition 'The third entry on page 6 of the London Telephone Directory is X'. Conceding the existence of this class of propositions is not a problem for the present argument for two reasons. Firstly, they are also propositions on which no other beliefs hang inferentially. And so according to the argument above, it is natural that we are not curious about them, since curiosity in propositions comes from their inferential relevance to beliefs about outcomes. Secondly, the important claim for internalism is not that we are inevitably moved by *epistemic norms* for every proposition, but that when we make a *judgement* that such norms are applicable, we are then moved by this *judgement*. And clearly, our epistemic judgements do not typically concern propositions of absolutely no interest to us. So, the above argument should strictly feature curiosity about 'most of' the world and 'all of our inferentially relevant beliefs', but this would make it too cumbersome.

and only judgements which feature *norms* for formation and revision. The first qualification does not restrict the account unduly. After all, judgements about how to form and revise beliefs – unlike judgements that I ought to have the belief I already do – are the ones for which we want an account of motivation. Similarly, in metaethics, internalism is not a position about judgements that I ought to act as I am in fact acting: it is not committed to the (implausible) claim that while I am saving a drowning child, my judgement that I ought to be saving him somehow motivates me further than the original judgement that I ought to save him. So there is no reason to expect internalism to be any more ambitious in the epistemic case.

Second, and more seriously, the above argument gets us to norms-judgements but does not automatically secure internalism for reasons-judgements, and for judgements featuring epistemic values and virtues. Now, moving from norms-internalism to other types of internalism is notoriously challenging, to put it mildly (see, for example, Audi 1997). But if we allow ourselves to tread somewhat roughshod and circumvent the sophisticated literature on this issue, we would think that establishing norms-internalism is at least a good foundation for reasons-internalism. Very baldly put, on the assumption that oughts imply reasons, it would be somewhat odd to be a norms-internalist and a reasons-externalist.[13] The issue deserves much more detailed attention, of course, but I will not attempt to do it justice here. The limited internalism established by my argument is enough to get a metaethics of belief going.

## 3.6 epistemic anti-humeanism

I now turn to the nature of motivation itself. In this and the next section, I argue that it is most plausibly understood on the anti-Humean model. The question at issue between Humeans and anti-Humeans, in the epistemic context, will be this: can a belief that $q$ motivate another belief that $p$ in the absence of any relevant desire? Humeanism says that a desire must be present for motivation to occur. The anti-Humean denies the need for desire. So, to establish anti-Humeanism we must show that cognitive states like beliefs, on their own, at least sometimes motivate other beliefs. And so that the argument avoids begging the question against non-cognitivism, we should show how ordinary descriptive beliefs – rather than normative-judgement beliefs – can motivate other beliefs.

[13] In fact, it is usually assumed that something stronger is the case – that oughts imply reasons and reasons imply oughts. (But see Dancy 2004.)

We start with the innocuous claim that belief is the sort of attitude which is generally responsive to truth-considerations. This is the unproblematic *descriptive* version of the well-rehearsed thesis that belief aims at truth.[14] Accepting this descriptive claim amounts to conceding that cognitive states can motivate: my *belief* that the evidence conclusively points to *p* automatically motivates my adopting another belief, the belief that *p*. Think of my belief (to take an overly original example) that the lipstick stains on his collar and the condom in his pocket are conclusive evidence for his infidelity. Having this belief automatically motivates me to adopt another belief – that he is unfaithful, in the absence of any relevant desire. It is not that I *want* to believe him unfaithful. Indeed, usually one has a desire *not to* believe one's partner unfaithful. And if one lets this desire motivate the belief, the belief is inappropriately formed according to many epistemologists. (It is called, confusingly in this context, a 'motivated belief'.) But in the envisaged case, the belief about the evidence – and not the desire – motivates the belief about his infidelity.

This case suggests the plausibility of anti-Humeanism about motivation, since anti-Humeanism merely claims that beliefs *can* motivate in the absence of desire, and the above example shows that they sometimes do. We have gleaned this result, moreover, for an ordinary descriptive belief (about the evidence), and so have not begged any barbed metaethical questions.[15]

## 3.7 TWO CHALLENGES

The Humean is likely to press two challenges at this point: first, he may deny that the above case involves genuine motivation (section 3.7.1); second, he may concede that the example involves motivation, but maintain that this is only because an unannounced desire is doing the motivation work (section 3.7.2).

### *3.7.1 The spurious-motivation challenge*

The first challenge[16] focuses on the awkwardness of saying that my belief that he is unfaithful is *motivated* by the belief about the evidence. What

[14] The usual normative gloss of this claim is that a belief is correct if and only if the believed proposition is true. At its strongest, the claim is taken partly to constitute the concept of belief itself (Shah 2003; Shah and Velleman 2005; and Wedgwood 2002a).

[15] I am making the fairly unproblematic assumption here that evidence-beliefs of this kind are not normative. See Shah in this volume.

[16] Thanks to Samantha Vice and Francis Williamson for this challenge.

the belief about the evidence is motivating is not an *action* but another *belief*. Yet the fight between Humeans and anti-Humeans was supposed to be about paradigmatic cases of motivation. What goes on in the belief case, the objection would go, is better described as the evidence-belief *causing* the other belief, or something like that. Sadly, I have no knock-down rejoinder to this challenge. But I think it faces a would-be dilemma which should dull its sting.

To begin with, I take it for granted that the challenge is not driven by the Humean thought that *because* the state in question is a belief, rather than a desire, what goes on is not motivation. That would simply beg the question at issue here. The would-be dilemma arises from interpreting the challenge as neutral on the Humeanism question. The dilemma is this: the challenge commits its exponent either to reasons-externalism or to the view that a belief can never be one's reason for another belief. Let me explain.

It is a common assumption that *the evidence* for a proposition *constitutes* a reason for believing the proposition.[17] It is an equally common assumption that my *belief* that there is evidence for this proposition at least sometimes constitutes a reason for me to believe the proposition (or for upping my credence in it, but I will talk of belief, for short). Now, there is an ongoing debate about the relationship among *there being* a reason for an action/belief, there being a reason *for an agent* to act/believe, an agent *having* a reason to act/believe, and so on. (See, for example, Raz 1975.) But I trust that the participants in this debate will concede that a belief sometimes constitutes at least one of these kinds of reason for another belief. I will talk about their sometimes being the *agent's reason*, but the reader is welcome to replace it with the type of reason he thinks a belief about the evidence most plausibly constitutes.

Now, if no belief can ever motivate me, as the current challenge claims, a set of reasons – those constituted by my beliefs – can never motivate me. So I can have a reason for belief and not be motivated by it. And this is (a type of) reasons-externalism. This is the first horn of the would-be dilemma that the spurious-motivation challenge faces.

Alternatively, the challenger could be an internalist about epistemic motivation and think that there is a necessary connection between having a reason for believing something and being motivated to believe in accordance with the reason. But if beliefs can never motivate on their own, as the challenger maintains, then beliefs on their own can never be our reasons

[17] This point is neutral on the debate between epistemic pragmatists and evidentialists. Their fight is about whether *only* evidence counts as reason for belief.

for believing, since a reason must be able to motivate. This is to deny the plausible assumption that at least sometimes my belief that there is evidence for a proposition constitutes my reason to believe the proposition. So this horn is counterintuitive. After all, it is my beliefs about the evidence that I would at least sometimes appropriately cite by way of justifying my belief.[18]

I said that this is a would-be dilemma which does not constitute a knock-down response to the challenge under scrutiny. Why is it only a would-be dilemma? Firstly, externalism would be a horn of a dilemma, only provided that we establish reasons-internalism. As discussed in section 3.5.3, however, the argument from intention does not, on its own, take us that far. Although reasons-internalism was made more plausible than externalism there, I fudged the task of arguing for it. So the externalist 'horn' for the challenger is a provisional one. Similarly with the second horn. Since I have not discussed in any detail the commitments of reasons-internalism, I have left it open to the challenger to avoid the second 'horn' by making the challenge more precise, and by finding a suitable version of reasons-internalism among the large and protean variety available in the literature. But I hope that, would-be as it is, the dilemma at least dampens the spirit of the challenge. So, instead of trying to block ever more elaborate versions of the challenge, I now move on to the second challenge which the Humean is likely to press.

### *3.7.2 The teleologist's challenge*

The second challenge to my argument for anti-Humeanism is this. Granting that a belief about the evidence (about lipstick, etc.) can motivate another belief (about his infidelity), it only does so because a conative attitude is smuggled through the back door. The belief that the evidence conclusively points to *p*, it will be urged, will only motivate me to adopt the belief that *p* if I have the further aim to believe only truths. (Shah and Velleman 2005 call this character 'the teleologist' about the thesis that belief aims at truth.[19]) Having the aim to believe only truths, of course, is a conative, rather than a cognitive matter – I *want* to believe only

[18] Note that this point does not presuppose classical epistemic internalism, which is the view that *all* justifiers are mental and so *only* beliefs can be my reasons. The present view merely assumes that *some* justifiers and reasons are mental. And this is compatible with classical epistemic externalism, which claims only that *not all* justifiers are mental.

[19] Velleman (2000) held this view. For a convincing defence of the view in the face of the most urgent objections to it, see Steglich-Petersen (2006a).

truths – and this conative element, the teleologist's challenge goes, is doing all the motivation work in the above example. So what motivates my belief that he is unfaithful isn't the belief about the evidence on its own, but the desire to believe truths as well.

Now, I do not think that the teleologist's position is overly plausible for all beliefs. Some beliefs just happen to me – I open my eyes and believe that it is a gloomy morning. Even if I did not want to believe truths, I would believe this one (assuming good visibility, fine eyesight, etc.). Such examples make rather strained the teleologist's insistence on a conative element in all cases of motivation. But then again, the teleologist is not committed to the view that *whenever* a belief is formed it is partly the result of the believer's aiming for true beliefs. Indeed, any teleologist worth his salt would deny this.[20] But while this denial makes the teleologist's position plausible, it also ultimately undermines his present challenge to anti-Humeanism.

Let us start by splitting one's beliefs into two groups – those which, according to the teleologist, are partly due to the believer's aiming for true beliefs, and those which are not (still according to him). Then focus on those which are *not* due to the believer's aim, and think of their relationship to motivation. The teleologist seems to have three options:

**(i)** All of these beliefs are motivated by other beliefs.
**(ii)** Some of these beliefs are motivated by other beliefs.
**(iii)** None of these beliefs is motivated by other beliefs. Rather, such beliefs are more appropriately seen as the causal products of evolutionary mechanisms, or something like that.

The first two options defuse the teleologist's challenge to epistemic anti-Humeanism. As long as *some* beliefs can motivate other beliefs, in the absence of a desire, anti-Humeanism is home and dry. But the teleologist can't hold either of these views as long as he is also a Humean about motivation. For him epistemic motivation is a matter of the believer's wanting to believe only truths. Yet we are here focusing precisely on those beliefs which *do not* involve this aim. So the third option is the teleologist's only option: none of the beliefs which don't involve the believer's aim for truth is motivated by other beliefs.

And the third option does seem to present a problem for anti-Humeanism. But unhappily for the teleologist, it is susceptible to the same would-be dilemma that faced the first challenge. The only difference here is that the dilemma holds for a smaller class of beliefs – only those

[20] Thanks to Asbjørn Steglich-Petersen for persuading me of this.

whose aetiology does not involve the believer's aim to believe only truths. The clearest examples of this class are perceptual beliefs. The dilemma then is that the teleologist is either an externalist about perceptual reasons, or he thinks that our perceptual beliefs never give us reasons for other beliefs. We replace the evidence beliefs with perceptual beliefs, and get the dilemma. The first horn, then, goes like this. Again, we start with a common assumption: some perceptual beliefs constitute reasons for me to believe. But according to the challenge, no perceptual belief can motivate on its own (since it does not involve the believer's aim for truth). So, a set of reasons – those constituted by my perceptual beliefs – can never motivate me. So I can have a reason for belief and not be motivated by it. This is reasons-externalism.

Alternatively (and this is the second horn), the challenger could be an internalist: my reasons necessarily motivate me. But according to the challenge, no perceptual belief can motivate on its own (since it doesn't involve the believer's aim for truth). So, perceptual beliefs on their own can never be our reasons for believing, which is to deny the common assumption with which we started. So my belief that I see raindrops on the window, say, can never be my reason for my belief that it is raining.

The upshot is that the teleologist's challenge also miscarries, and for the same reason as did the spurious-motivation challenge. For the teleologist's challenge to work, he must claim that a particular class of beliefs – beliefs which do not involve the believer's aiming for true beliefs – never motivates. And this commits him either to being a reasons-externalist, which is made *prima facie* implausible by the argument from intention; or to thinking that our perceptual beliefs can never be reasons for one to adopt a belief, which goes against a widely shared assumption.

## 3.8 FROM EPISTEMIC MOTIVATION TO THE NATURE OF EPISTEMIC JUDGEMENTS

As far as motivation goes, then, our metaethics of belief should be anti-Humean about the nature of epistemic motivation, and internalist about how epistemic judgements motivate us. Much more needs to be said, no doubt, in support of these arguments. But instead of saying it, I wish to conclude by dwelling on how these views on motivation illuminate the upper reaches of the proposed metaethics of belief.

The upper-level result on which I want particularly to focus here is that the dual stance on epistemic motivation makes *prima facie* plausible (soft) cognitivism about epistemic judgements. How so? Concisely put, the dual

commitment to internalism and anti-Humeanism, which I have programmatically defended, rules out non-cognitivism about epistemic judgements. To see how, suppose that we were (1) anti-Humeans about motivation and (2) non-cognitivists about epistemic judgements. Using 'desire' and 'belief' to stand for non-cognitive and cognitive attitudes respectively, we get:

**(1)** There is no necessary connection between desire and motivation (anti-Humeanism).
**(2)** Epistemic judgements are desires (non-cognitivism).
**(C)** So, there is no necessary connection between epistemic judgements and motivation.

And this conclusion just states externalism. So, one cannot be an internalist, anti-Humean, and a non-cognitivist.

Conversely, suppose we were (1) internalists about the connection between epistemic judgements and motivation, and (2) non-cognitivists about epistemic judgements. We then get:

**(1)** There is a necessary connection between epistemic judgements and motivation (internalism).
**(2)** Epistemic judgements are desires (non-cognitivism).
**(C)** So, there is a necessary connection between desire and motivation.

And this conclusion is incompatible with anti-Humeanism, which claims precisely that there is no such necessary connection, since beliefs on their own can motivate.

So whichever way we look at it, non-cognitivism is incompatible with the dual commitment to internalism and anti-Humeanism. If cognitivism and non-cognitivism are mutually exclusive and exhaustive, then the above arguments for internalism and anti-Humeanism effectively establish cognitivism about epistemic judgements. Now, there is no doubt that the two positions are mutually exclusive, but they certainly do not exhaust the field, as I have defined them. Cognitivism was defined as the view that *all* epistemic judgements are cognitive; and non-cognitivism, as the view that *all* epistemic judgements are non-cognitive. There is an obvious third option – that some are cognitive and some non-cognitive. This option blocks us from deducing cognitivism from the dual stance on motivation. Now, some authors actually define cognitivism precisely as this third option (e.g. Cuneo 2007). If this is correct, then cognitivism about epistemic judgements is vindicated. But whatever we might wish to *call* the position, given anti-Humeanism and internalism, it follows that at least some of our epistemic judgements are cognitive. This conclusion may be unsatisfying for the hard-nosed cognitivist, but it is much to get from a stance on motivation.

## 3.9 CONCLUSION

I hope to have made attractive here a particular kind of metaethics of belief. Methodologically, the kind in question is one which builds itself up from theoretically lower-level commitments concerning motivation. In this way we get a substantive picture of how epistemic judgements motivate us. This picture, which involves the dual commitment to internalism and anti-Humeanism, is then used to adjudicate at the upper levels of our metaethics of belief. We should think that at least some epistemic judgements are cognitive in nature – they express our beliefs, rather than mere sentiments, about how we should conduct our doxastic lives.[21]

[21] Many thanks to Ward Jones, Asbjørn Steglich-Petersen, Herlinde Pauer-Studer, an anonymous CUP referee, and Rhodes University Philosophy Department for comments on earlier drafts. Thanks to the Andrew Mellon Postdoctoral Fellowship, Rhodes University, and Universität Wien for their financial support.

CHAPTER 4

# *Error theory and reasons for belief*

*Jonas Olson*

## 4.1 INTRODUCTION

According to moral error theory there are no moral reasons. Moral reasons would be metaphysically queer entities not recognizable from the perspective of philosophical naturalism. Consequently, claims about what there is moral reason to do are never true. Recently several philosophers have highlighted the structural similarities between moral reasons and epistemic reasons, or reasons for belief.[1] This raises the question whether, from the error theorist's naturalistic perspective, epistemic reasons should be deemed equally queer as moral reasons.[2] Should the moral error theorist also accept an error theory about epistemic reasons? But isn't error theory about reasons for belief obviously unappealing? These are questions I intend to pursue in this essay.

Some philosophers have argued that epistemic reasons are an embarrassment for moral error theorists. Using a 'companions in guilt' strategy, some of these philosophers have argued that error theory about epistemic reasons is implausible and that therefore error theory about moral reasons should be deemed equally implausible.[3] Some have gone as far as to claim that error theory about epistemic reasons is self-defeating.[4] Terence Cuneo, who has offered the most elaborate and thorough examination of the parallelisms between moral and epistemic reasons, maintains that epistemic error theory implies at least three 'undesirable results' that should prompt

[1] Among them are Cuneo (2007); Kelly (2003); Putnam (1990); Railton (1997); Scanlon (1998); Stratton-Lake (2002).

[2] Terence Cuneo (2007) notes that epistemic error theory is not well represented in the philosophical literature and he takes this to be 'testimony to the fact that philosophers have frequently overlooked the implications that [. . .] various [error theoretical] arguments have for epistemology' (114).

[3] For a critical in-depth survey of the employment of 'companions in guilt' arguments as a general strategy in metaethics, see Lillehammer (2007).

[4] Cuneo (2007: ch. 4); Stratton-Lake (2002: xxv).

us to reject the theory, and with it moral error theory.[5] I shall investigate how undesirable these results really are.

I proceed as follows. In section 4.2 I briefly outline the alleged structural similarities between moral reasons and epistemic reasons and give a more precise characterization of epistemic error theory. I distinguish between *transcendent* norms, which imply categorical reasons, and *immanent* norms, which do not imply categorical reasons. For the sake of argument I accept the view that both moral and epistemic norms are transcendent and that both moral and epistemic reasons are categorical, which is to say that I accept the central claim of Cuneo *et al.*[6] However, error theorists can afford the luxury of being non-committal on the categoricity of epistemic reasons. Perhaps a naturalistic account that denies categoricity is more defensible about epistemic than about moral reasons.[7] Still, error theorists would surely be in a stronger position if they could grant their critics the view that moral and epistemic reasons are both species of categorical reasons. I shall therefore grant this view and set out to defend error theory by arguing that epistemic error theory does not lead to the undesirable results its opponents allege. I will do so mainly by critically examining Cuneo's arguments. This is the task to be undertaken in sections 4.3 and 4.4.

Before I get started I need to make a brief preliminary remark about formulations of error theories. It is routinely said that according to moral error theory, first-order moral claims are uniformly false. First-order moral claims are claims that entail something about what some agent morally ought to do or not to do, what would be morally permissible or impermissible for some agent to do or not to do, what there is a moral reason for some agent to do or not to do, and the like; or what would be morally good (bad) or morally (un)desirable, and the like. Hence, if error theory is correct, the claim that, for example, torture is wrong is false. But the law of the excluded middle entails that if it is false that torture is wrong, it is true that torture is *not* wrong. And the negative claim that torture is not wrong appears to entail that torture is morally permissible, which is clearly a first-order moral claim. This suggests that contrary to the contentions

[5] Cuneo (2007: ch. 4). Cuneo speaks mainly about *nihilism* rather than *error theory*. I prefer the term 'error theory' since 'nihilism' has too many misleading connotations both within and outside philosophical parlance. For instance, in moral philosophy 'nihilism' is often taken to signify the first-order moral view that anything is morally permissible. But error theory is purely a second-order, or metaethical, view. See further below and Olson 2010.

[6] I hasten to add that the term 'reason' is notoriously ambiguous. There are senses of 'reason' in which error theorists can recognize that there are reasons for agents to believe certain propositions, but as I explain in section 4.3.1 below, these reasons are non-categorical.

[7] See Heathwood (2009) for this claim.

of many moral error theorists, moral error theory does have distinctive first-order moral implications.[8] And rather vulgar ones at that; anything turns out to be morally permissible!

I shall not attempt to resolve this issue here.[9] I note merely that one possible strategy for sidestepping this problem is to opt for a version of moral error theory that says that first-order moral claims rest on false *presuppositions* and are therefore uniformly *neither true nor false*.[10] In order not to exclude this version of moral error theory, I shall take moral error theory to be the view that no first-order moral claim is true. Analogously, I shall take epistemic error theory to be the view that no first-order epistemic claim is true, where first-order epistemic claims are claims to the effect that there are epistemic reasons for some agent to believe or not to believe some proposition or that believing some proposition is or would be permissible or impermissible, and the like.

## 4.2 Moral and epistemic reasons; transcendent and immanent norms

In what follows I shall talk about *norms* and *reasons*. A norm, I shall assume, is a fact that can be expressed by a universally quantified sentence that implies that there are, for some class of agents in some set of circumstances, reasons to behave in a certain way, or that for some class of agents in some set of circumstances, some form of behaviour would be (in)correct or (im)permissible. A reason is a fact that explains why some agent ought (pro tanto) to behave in certain ways, or why some form of behaviour would be (in)correct or (im)permissible. I shall allow for the possibility that some norms are self-explaining – that some norm holds might itself be a reason to behave in certain ways (see footnote 16 below).

Many philosophers hold that it is a commonplace that moral norms are *categorical*. This is not merely the claim that moral norms are typically non-hypothetical or unconditional in grammatical form.[11] The thought is rather

[8] Mackie states in his seminal discussion that moral error theory is meant to be a second-order view and that as such it is logically independent of any first-order moral view (1977: 16).

[9] I do so in Olson (2010: 68–70).

[10] See, for example, Oddie (2005: 8–13); Sobel (MS) for this version of moral error theory.

[11] Mackie (1977: 28f.) notes that moral norms may be conditional in grammatical form although they are categorical in the sense that their reason-giving force does not depend on the desires, ends, roles, or activities of agents. For instance, the moral norm 'if you make a promise, you ought to keep it' is conditional in grammatical form but it is categorical in the sense that it is meant to imply that anyone who has made a promise to do something has a reason to do that thing, irrespective of her desires, ends, roles, or activities.

that the reason-giving force of moral norms is not entirely dependent on agents' desires or ends, and neither on agents' roles, such as professional or institutional roles, or their engagements in rule-governed or goal-oriented activities.[12] For instance, the norm that it is morally wrong to inflict pain on sentient beings for the mere fun of it entails that there are reasons for all agents, irrespective of their desires, ends, roles, or activities, not to do so. Philosophers who maintain that epistemic reasons are an embarrassment for moral error theorists hold that epistemic reasons are categorical in parallel ways. According to these philosophers, what there are reasons for agents to believe does not depend entirely on agents' desires, ends, roles, or activities.

The exact way in which epistemic reasons are categorical is not an uncontroversial matter but I suggest that if on the commonsense conception epistemic reasons are categorical, they are so at least in the following sense: if an agent, A, believes that *p*, or considers whether *p*, and is presented with evidence that not-*p*, A has epistemic reason to reject or weaken the belief that *p*, irrespective of whether A wants to believe that *p* or has desires, ends, or roles (e.g. in virtue of being engaged in some rule-governed or goal-oriented activities) whose fulfilment would be promoted by A's believing that *p*. Peter Railton makes a stronger claim:

> On the usual view of things, two agents in the same epistemic situation (same evidence, same background beliefs) would have the same reason for believing any given proposition, regardless of possible differences in their personal goals. (Railton 1997: 53)

Railton's characterization of the categoricity of epistemic reasons is stronger than the one I suggested since it allows that an agent can have an epistemic reason to believe that *p* irrespective of whether she already believes that *p* or considers whether *p*. Railton's characterization is supported by Tom Kelly:

> [F]rom the fact that some subjects are matters of complete indifference to me it does not follow that I will lack epistemic reasons for holding beliefs about those subjects. If, despite my utter lack of interest in the question of whether Bertrand Russell was left-handed, I stumble upon strong evidence that he was, then I have strong epistemic reasons to believe that Bertrand Russell was left-handed. Indeed, my epistemic reasons will be no different than they would be if I had acquired the same evidence deliberately because I *did* have the goal of finding out whether

[12] For defences of the commonsense view that moral reasons are categorical, see, for example, Mackie (1977); Joyce (2001). A well-known dissenter is Foot (1972). Finlay (2008) claims that the idea that moral reasons are categorical is 'the error in the error theory'. I respond to Finlay's arguments in Olson (2010: section 4). Joyce (forthcoming) also responds to Finlay (2008).

Bertrand Russell was left-handed. Once I come into possession of evidence which strongly supports the claim that *p*, then I have epistemic reasons to believe that *p*, regardless of whether I presently have or previously had the goal of believing the truth about *p*, or any wider goal which would be achieved in virtue of my believing the truth about *p*. (Kelly 2003: 625, emphasis is original).

Finally, consider Terence Cuneo:

[S]ome epistemic facts are, imply, or indicate categorical reasons for agents to behave in certain ways [. . .] regardless of whether these agents care about conducting their behaviour in a rational way, whether they belong to a social group of a certain kind, or whether they have entered into social agreements with others. [. . .] So, for example, [. . .] whether I have an epistemic reason, say, to believe a proposition for which I have compelling evidence is not contingent upon whether I care about believing what is true. [. . .] Like moral reasons of certain kinds, a wide range of epistemic reasons stand in sharp contrast to the norms that govern activities such as chess and tennis. (Cuneo 2007: 59)

Although these formulations of the categoricity of epistemic reasons differ somewhat in their details they all seem to concur on the point that for any agent A, A's epistemic reasons are not entirely determined by A's desires, ends, roles, or activities. In Cuneo's words, both moral and epistemic norms are *authoritative* (2007: 58f.).[13] I take it that part of Cuneo's aim in calling moral and epistemic norms authoritative is to distinguish them from institutional norms of, for example, chess, soccer, grammar, or etiquette. It is obvious that there are no reasons for those who are not engaged in playing chess or soccer to comply with the norms of chess or soccer, and it is almost as obvious that there are no reasons for agents to comply with the norms of grammar or etiquette irrespective of their desires, ends, roles, or activities.

I am not sure, however, that it is right to say that institutional norms are less authoritative just because they do not apply to agents categorically. As Philippa Foot once said, institutional norms do not 'fail to apply' to an agent just because she does not care about them.[14] As I suggested above,

[13] Cuneo highlights a number of other structural similarities between moral and epistemic reasons but the idea that both types of reasons are categorical or authoritative is the central one. It suffices to raise a problem for moral error theory, which is why I shall limit my discussion to this similarity. It should be noted that Cuneo focuses on parallelisms between moral and epistemic *facts*. In Cuneo's usage, 'epistemic facts' covers facts about what would be *rational* for some agent to do or believe. I am far from sure that facts about rationality are normative in the same sense that facts about moral and epistemic reasons are normative (see footnote 26 below) and I shall therefore focus more narrowly on epistemic reasons.

[14] Foot (1972: 308). Joyce (2001: 36) calls such institutional norms *weak*, as opposed to *strong*, categorical imperatives. But I doubt that *categoricity* varies in degrees of strength, and even if it did it is hard to see why weak categorical imperatives would be less mysterious than strong ones.

a norm can be expressed by a universally quantified sentence that implies that there are for some class of agents in some set of circumstances reasons to behave in certain ways. A norm, N, applies to an agent A just in case A is a member of the class bound by the universal quantifier in N. In this sense, norms of, for example, etiquette may apply to, or be authoritative for, an agent even though it is not the case that there is categorical reason for that agent to comply with the norms.

This is why I find it more useful to distinguish between *transcendent* and *immanent* norms. The former apply to agents categorically; their reason-giving force transcends agents' desires, ends, activities, or institutions. Immanent norms, by contrast, are those whose reason-giving force depends on agents' desires or ends, their engagement in certain rule-governed activities, or their occupation of certain roles; the reason-giving force of immanent norms does not transcend desires or ends, or rule-governed activities or roles, which is why immanent norms imply only *non-categorical* reasons.

Another way of putting it is to say that while immanent norms determine correct behaviour according to what is required to satisfy agents' desires or promote their ends, or according to rules or institutional standards, it does not follow that there are categorical reasons to comply with these norms. For transcendent norms, it does follow that there are categorical reasons to comply with them. As we have seen, it is a common view that on the commonsense conceptions of moral and epistemic norms these are examples of transcendent norms, whereas the norms of, for example, chess, soccer, grammar, and etiquette are prime examples of immanent norms. For instance, to say that a norm is a moral or epistemic norm is to say that there are reasons for any agent to comply with that norm, irrespective of her desires, ends, roles, or activities. To say that some norm is a norm of etiquette or chess, by contrast, is not to say that there are categorical reasons to comply with it, but rather that some sort of behaviour would be incorrect relative to a certain standard of etiquette or relative to the rules of chess.[15] Hence I shall say that *norms* can be transcendent or immanent and that *reasons* can be categorical or non-categorical.[16]

[15] It is possible, of course, that some transcendent (e.g. moral) norms require compliance with some immanent (e.g. etiquette) norms.

[16] In the context of semantic normativity, Anandi Hattiangadi (2007) distinguishes between *normativity* and *norm-relativity* in a way that resembles my distinction between transcendent and immanent norms and the correlative distinction between categorical and non-categorical reasons. In a similar vein, John Broome (2007b) distinguishes between on the one hand *reasons*, and on the other *requirements of convention*. For instance, Broome's view is that the norm that one ought to shake hands with one's right hand *requires* one to do so. My view is that the norm implies a (conclusive) non-categorical *reason* to do so. That the reason to shake hands with one's right hand

Error theorists object to the existence of transcendent norms and categorical reasons. It is metaphysically puzzling how there can be norms that imply categorical reasons for agents to behave in certain ways irrespective of their desires, ends, roles, or activities. Error theorists do not object to the existence of immanent norms and non-categorical reasons, however.[17] There is nothing metaphysically queer about the fact that there is (conclusive) non-categorical reason for chess players not to move the rook diagonally, since this is just the fact that moving the rook diagonally is incorrect according to the rules of chess. But insofar as epistemic norms belong to the category of transcendent norms, they should be deemed equally metaphysically queer as moral norms in at least two respects. First, just as moral facts about what there is categorical reason to *do* would have to supervene on other facts, epistemic facts about what there is categorical reason to *believe* would have to supervene on other facts. Mackie famously finds moral supervenience a metaphysically queer relation and the supervenience of epistemic facts would seem to be no less queer. To illustrate, just as there might be moral reasons to perform an action *because* it would promote pleasure, there might be epistemic reasons to believe a proposition *because* it is supported by a sound deductive argument or by strong inductive or abductive arguments. Mackie asks rhetorically about the moral case, 'just what *in the world* is signified by this "because"' (Mackie 1977: 41, emphasis in original), and that question looks equally well motivated about the epistemic case.

Second, just as moral facts would tell agents what to *do* by virtue of having 'a demand for [. . .] action somehow built into [them]' (Mackie 1977: 40), epistemic facts would tell agents what to *believe* by having a demand for belief somehow built into them. In other words, both moral and epistemic facts would be *intrinsically prescriptive*. But according to error theorists, intrinsic prescriptivity is a metaphysically queer feature.[18] In these

is non-categorical means that shaking hands with one's right hand is correct relative to a norm of etiquette. The norm that one ought to shake hands with one's right hand is thus self-explaining – it implies that there is a (conclusive) non-categorical reason to shake hands with one's right hand.

I prefer to talk about non-categorical *reasons* rather than *requirements* since the reasons implied by immanent norms need not be conclusive, i.e. they need not be requirements. For instance, a student may have non-categorical reason to submit an essay in each week of term without being required to do so. To say that there is a non-categorical reason for the student to submit an essay in each week of term is just to say that if she does she will be more likely to meet the standards of being successful in her role as student. Finally, I prefer to talk about *non-categorical* reasons rather than reasons of *convention* since some non-categorical reasons depend on agents' desires rather than on conventions; see section 4.3.1 below.

[17] See Mackie (1977: 25–30, 77–82); Joyce (2001); Olson (2010).

[18] See in addition to Mackie, Joyce (2001) and Olson (2010), for example.

ways, and possibly in others too, sceptical doubts about moral reasons spill over to the domain of epistemic reasons. As Philip Stratton-Lake puts it:

> The difference between [moral] and epistemic reasons is [. . .] that they warrant different things. [Moral] reasons warrant pro-attitudes and actions whereas epistemic reasons warrant beliefs. If, therefore, one has doubts about the normative (warranting) relation itself, these doubts could not be localized in such a way as to avoid scepticism about epistemic as well as [moral] reasons. (Stratton-Lake 2002: xxvi).[19]

It seems then that moral error theorists who hold that no first-order moral claim is true are committed to accept *epistemic error theory*, according to which no first-order epistemic claim is true. As I have already indicated, I shall not question the claim that moral and epistemic norms are transcendent norms and nor shall I rehearse or defend further the familiar error theoretical arguments for why transcendent norms and categorical reasons are queer. For the sake of argument, I accept the claim that error theory cannot be localized to the moral domain so as to avoid error theory in the epistemic domain. In what follows, I shall discuss some recent arguments to the effect that epistemic error theory is untenable.

## 4.3 EPISTEMIC ERROR THEORY: HOW UNDESIRABLE?

Terence Cuneo lists three 'undesirable results' of epistemic error theory, each of which he takes to provide sufficient grounds for rejection. I shall argue that epistemic error theory does not have the alleged results.

### *4.3.1 'Epistemic error theory is either self-defeating or polemically toothless'*

The first objection comes in the form of a dilemma: either epistemic error theory is self-defeating or it is polemically toothless. To uncover the first horn, recall that according to epistemic error theory, there are no reasons for belief. But insofar as the error theorist claims that her arguments, for example, the queerness arguments, are reasons to believe that there are no reasons for belief, the theory becomes patently self-defeating. In short, if the queerness arguments are correct and if error theory is true, there are no reasons to believe in the error theory (Cuneo 2007: 117f.).[20]

[19] Stratton-Lake focuses widely on *practical* reasons rather than moral reasons more narrowly, but that is inessential in the present context.

[20] Other writers have also put this argument against moral error theory forward. For instance, Stratton-Lake holds that '[epistemic error theory] seems to undermine itself, for it [says] that we have reason

The error theorist avoids the first horn of the dilemma by distinguishing between arguments to the effect that *p* is true and arguments to the effect that there are reasons to believe that *p*.[21] In particular, the error theorist is offering arguments to the effect that the error theory is true. She is not offering arguments to the effect that there are reasons to believe that the error theory is true. She is thus not committed to the patently self-defeating position that there are epistemic reasons to believe that there are no epistemic reasons.

But now Cuneo claims that if error theorists hold that there are no reasons to believe in their theory they are impaled on the second horn of the dilemma: epistemic error theory becomes 'polemically toothless in the [. . .] sense [that] [n]o one would make a rational mistake in rejecting it' (Cuneo 2007: 117). But since error theorists are not in the business of offering arguments about what would be rational to believe or about what there is reason to believe there is no harm in conceding that error theory is toothless in *those* debates. What is important is that error theoretical arguments have bite in debates on where the truth lies in metaethics and metaepistemology; these are the debates with which error theory is concerned. And given that the aim of metaethical and metaepistemological enquiry is to get at the truth, error theory will not be polemically toothless in those debates.

It should be noted, however, that the term 'reason' is notoriously ambiguous and although error theorists are committed to denying that there are epistemic – and hence categorical – reasons for belief, they can maintain that there are *other* senses of 'reason' in which it might well be true that there are reasons for some agents to believe certain propositions.[22] These senses pick out non-categorical reasons for belief. I shall discuss briefly two kinds of non-categorical reasons: *hypothetical* and *role-* or *activity-dependent* reasons.

First, *hypothetical* reasons for belief. For the error theorist, to say that there are hypothetical reasons for some agent to believe that *p* is simply to

to be sceptical about reasons, and it implies that it is false that we have reason to be sceptical about reasons' (2002: xxv).

[21] See also Danielsson and Olson (2007); Leite (2007); Olson (2009) for this distinction. Guy Fletcher (2009: 366) makes the same point in his review article on Cuneo (2007).

[22] To clarify, I do not claim that Cuneo and other critics of epistemic error theory are committed to denying that 'reason' is ambiguous in the ways I claim. Note also that I do not suggest anything in the way of conceptual revision or reform. My claim is that the term 'reason (for belief)' as used in ordinary discourse is ambiguous in that it can signify either categorical reasons for belief (of which there are none, according to error theorists) or non-categorical reasons for belief (of which there are plenty, according to error theorists). Thanks to Bart Streumer for pressing me on this point.

say that that agent has some desire or end that would be fulfilled were she to believe that *p*.[23] For instance, for agents who want to have true beliefs on matters of metaethics there are hypothetical reasons to believe in the error theory, since believing in the error theory would satisfy that desire. This claim is tolerable for the error theorist since it is purely non-normative; it ascribes to an agent a desire and specifies how it would be satisfied. Hypothetical reasons are thus non-categorical.

Second, error theorists can recognize reasons for belief that apply to some agents in virtue of their roles, or in virtue of being engaged in some rule-governed or goal-oriented activities. One might well hold that there is a norm that chaplains believe the essentials of the Bible, in which case there are reasons for chaplains qua ocupying the role of chaplains to believe that God exists (irrespective of whether it is actually true that God exists and irrespective of whether there are undefeated arguments to the effect that God exists). One might also hold that since the goal of many intellectual endeavours, for example, metaethics and metaepistemology, is to get at the truth, there is a norm that people engaging in such activities have true beliefs on the subject matter, in which case there are reasons for metaethicists and metaepistemologists, qua being engaged in metaethical and metaepistemological enquiries, to believe in the error theory.

But of course, error theorists maintain that these norms are immanent since these norms have reason-giving force for some agents only in virtue of their roles or the activities they are engaged in. Role- and activity-dependent reasons are thus non-categorical reasons. (Note that since role- and activity-dependent reasons apply to agents in virtue of their roles or activities and not in virtue of their desires there need be no perfect correlation between role- and activity-dependent reasons and hypothetical reasons. Clearly, an agent can occupy a role she has no desire to fulfil or be engaged in an activity she has no desire to carry out successfully.) For chaplains and metaethicists and metaepistemologists to fail to comply with role- and activity-dependent reasons is not to violate a transcendent norm

[23] According to Mackie's general characterization of hypothetical reasons, such reasons are contingent on agents' desires (1977: 28). It is not entirely clear to me what Mackie meant by saying that hypothetical reasons are 'contingent' on desires. (To be wholly accurate, Mackie spoke in this context of categorical and hypothetical *imperatives*, and he claimed that in the case of the latter, 'the *oughtness is contingent* upon the desire' (1977: 28, my emphasis).) He could not have meant that there is, for each agent, reason for that agent to satisfy her intrinsic desires since that would entail that hypothetical reasons ultimately rest on categorical reasons. The error theorists' view must be that the claim that there is hypothetical reason for some agent to $\phi$ simply reduces to the claim that the agent desires to $\Psi$ and were she to $\phi$, she would $\Psi$. This is, of course, a purely non-normative claim. Cf. Mackie 1977: 65f., 75. See also Hampton (1998: 127–30); Lillehammer (2007: ch. 3). I discuss this issue further in Olson (2010: 77–8).

but simply to fail to live up to standards of what it takes to be successful in their roles or activities.

To sum up this subsection, the charge that epistemic error theory is self-defeating rests on the assumption that arguments to the effect that *p* amount to arguments to the effect that there are reasons to believe that *p*. But error theorists can distinguish between arguments to the effect that *p* and arguments to the effect that there are reasons to believe that *p*. In particular, they can hold that their error theoretical arguments are arguments to the effect that the error theory is true and not to the effect that there are reasons to believe in the error theory. This also shows why there is no harm in conceding that epistemic error theory is polemically toothless in debates about what there are reasons to believe.

### *4.3.2 'Epistemic error theory implies that there can be no arguments for anything'*

Before coming to Cuneo's second 'undesirable result' I shall discuss the third since that can be dealt with in a way similar to the first. Cuneo claims that (A) '[a] statement's being offered as evidential support for a conclusion [. . .] is just a matter of its being offered as a *reason* for accepting that conclusion' (Cuneo 2007: 121, emphasis in original). I assume here that by 'reason' Cuneo means 'epistemic (i.e. categorical) reason'. From (A), Cuneo infers that (B) 'if [epistemic error theory] were true, it would be impossible that there were premises of an argument that provide evidential support for its conclusion' (Cuneo 2007: 121). And from this – along with what he claims to have established as the first 'undesirable result' – he concludes that (C) '[e]ither epistemic [error theory] is self-defeating or it implies that there could be no arguments for anything' (Cuneo 2007: 121).

I have already argued in subsection 4.3.1 above that epistemic error theory is not self-defeating. The tenability of (B) and of the second disjunct of (C) depends on the tenability of (A), i.e. that evidential support for *p* amounts to reasons to believe that *p*. Clearly, my response to the claim that epistemic error theory is self-defeating rests on a distinction between on the one hand arguments to the effect that *p*, or evidential support for *p*, and on the other hand reasons to believe that *p*. Cuneo implicitly rejects that distinction when he claims that evidential support for *p* just is a reason to accept that *p*, and he says at one point that he uses 'the terms "evidence" and "(epistemic) reasons" more or less interchangeably' (Cuneo 2007: 192, n.12). The basic disagreement, then, concerns whether the notion of evidence, or the evidential support relation, is itself normative.

I concede that in colloquial contexts 'reason to believe that *p*' and 'evidence that *p*' are sometimes used interchangeably but there are plausible explanations of this that do not conflict with epistemic error theory. First, in these contexts we might believe (mistakenly, according to error theory) that evidence that *p* provides categorical reasons to believe that *p*. But this does not establish that evidence is normative since properties that provide reasons, and facts that are reasons, need not themselves be normative.[24] Second, 'reason to believe that *p*' and 'evidence that *p*' might sometimes be used interchangeably because we believe that there are non-categorical reasons (of the kind discussed in section 4.3.1 above) to believe in accordance with evidence with respect to whether *p*, i.e. we believe that believing in accordance with evidence with respect to whether *p* in these contexts would satisfy some desire, promote some end, or help us fulfil some role or activity of ours. Clearly, this does not show that the notion of evidence is normative. Hence the fact that the phrases 'reason to believe that *p*' and 'evidence that *p*' are sometimes used interchangeably does not establish that they express identical concepts.

Tom Kelly (2006, 2007) distinguishes between normative and non-normative notions of evidence and maintains that there is an everyday notion of evidence that is normative. Kelly writes that

> on the view that evidence has no normative force of its own, it is mysterious why an explicit judgement to the effect that one's evidence strongly supports some proposition typically results in a belief that that proposition is true. (Kelly 2007: 468–9)

The alleged mystery dissolves as soon as we take into account a point just made: explicit judgements to the effect that one's evidence strongly supports *p* typically result in belief that *p* because, typically, when we assess evidence for and against *p* we do so because we recognize non-categorical reasons to have true beliefs as to whether *p*; we want to know whether *p*, we are interested in whether *p*, and we believe that were we to believe in accordance with evidence with respect to whether *p* we would come to know whether *p*. Note also that the idea that evidence has 'normative force' does not license the conclusion that evidence is normative. Compare: the idea that pain has normative force in the sense that the fact that an action would bring about pain is a reason against performing the action does not license the conclusion that pain is normative.

[24] Cf. Olson (2009).

Kelly also asks us to consider

> a standard Bayesian explication of evidence, according to which to treat something as confirming evidence is to treat it as a reason to increase one's confidence that that hypothesis is true, while to treat something as disconfirming evidence is to treat it as a reason to decrease one's confidence. Given such an explication, there is an internal connection between recognizing something as evidence and recognizing it as a reason to change one's present view. (Kelly 2007: 467–8)

But this does not show that the concept of evidence is normative in any sense that would spell trouble for error theorists, i.e. it does not show that the concept of evidence involves transcendent norms. In fact, Kelly's point serves to illustrate a point I make elsewhere in the chapter. Failure to revise one's beliefs appropriately in the light of new evidence simply amounts to failure to meet the standards of Bayesian belief revision. This means that there are non-categorical reasons for some agents to comply with the dictates of Bayesian belief revision, namely those agents who want to meet Bayesian standards, or have ends that would be promoted by meeting such standards, or occupy roles or engage in activities that require meeting Bayesian standards. Error theorists can recognize such reasons.

In distinguishing between on the one hand arguments to the effect that *p* or evidential support for *p* and on the other hand reasons to believe that *p*, I have relied on a non-normative notion of evidence that Kelly (2006: §3, 2007: 469–71) calls *indicator evidence*: *q* is evidence that *p* just in case *q* reliably indicates that *p*. As I understand the notion, indicator evidence can be logically conclusive, in which case the premises of a sound deductive argument are indicator evidence for its conclusion. It can also take the form of inductive or abductive arguments. It is true that in order to determine whether some argument is a *strong* inductive or abductive argument one has to invoke norms of theory assessment, but these norms are non-categorical rather than categorical; they are whatever norms are typically adopted by scientists and philosophers in their roles and professional activities. Hence error theorists can maintain that it is not in principle more difficult to determine whether some inductive or abductive argument is *strong* than to determine whether some move in chess is a *good* move. They will have to concede, however, that there is no non-categorical reason to accept the standards of theory assessment and inference rules typically adopted by scientists and philosophers, e.g. inference to the best explanation, rather than some other standard or inference rule, e.g. 'inference to the worst explanation'. This parallels the way in which moral error theory recognizes

the wrongness of torture relative to UN declarations, but rejects categorical reasons to comply with such declarations.

Now, error theorists can maintain that the arguments from queerness indicate that error theory is true since they show that moral and epistemic discourses commit ordinary speakers to entities that are hard to square with a naturalistic world-view. That these supposed entities would be supervenient and intrinsically prescriptive explain why they are hard to square with a naturalistic world-view. The point is that in trying to establish that epistemic error theory implies that there can be no arguments for anything, Cuneo assumes that the notion of evidence is normative in the sense that it involves categorical reasons; this is why he claims that a statement's being offered as evidential support for a conclusion just is a matter of its being offered as a *reason* for accepting that conclusion. The epistemic error theorist can retort that a statement's being offered as evidential support for a conclusion just is a matter of its being offered as a reliable indicator that the conclusion is true, where this involves only non-categorical reasons in the ways suggested above.

At this point one might worry that epistemic error theory has made itself otiose. If claims about evidence are naturalistically kosher, what is the error that permeates first-order epistemic discourse and why doesn't epistemic error theory collapse into epistemic naturalism?[25]

In response, it is worth pausing to recapitulate the dialectic. Critics of moral error theory argue that moral and epistemic reasons are alike in that they are both species of categorical reasons. For the sake of argument, I have granted the critics this view. Now, a common view in first-order epistemic discourse is that there are epistemic reasons to believe in accordance with evidence. As I have argued in this subsection, epistemic error theory holds that claims about evidence are not normative in the sense of entailing claims about categorical reasons, but that claims to the effect that there are epistemic reasons to believe in accordance with evidence are. The error in first-order epistemic discourse is precisely the supposition that there are epistemic, i.e. categorical, reasons to believe anything. Compare: moral error theory holds that it is a non-normative claim that some act is an act of torture but that it is a normative claim that there are moral reasons not to torture. The latter kind of claim is never true. Epistemic error theory holds that it is a non-normative claim that a proposition, $q$, is evidence that a distinct proposition, $p$, is true but that it is a normative

[25] Terence Cuneo, Hallvard Lillehammer, Folke Tersman, and an anonymous reviewer for CUP all raised variants of these worries.

claim that *q* is an epistemic reason to believe *p*. The latter kind of claim is never true.

### *4.3.3 'Epistemic error theory rules out the possibility of epistemic merits and demerits'*

Cuneo's second objection is that if epistemic error theory is true, none of our beliefs can be based on reasons and consequently none of our beliefs can display epistemic merits of being justified, warranted, rational, and the like. As I will explain in this subsection I believe that the first step of this argument equivocates on the slippery term 'reason'. Again, what error theorists deny is simply the existence of transcendent norms, and consequently the existence of categorical reasons for belief. Error theorists may well maintain that believing *p* while lacking evidence that *p*, e.g. as an *idée fixe* or as a result of wishful thinking, is to fail to meet the standards of being a responsible believer (with respect to *p*).

It might be objected that to call someone a responsible (or irresponsible) believer is to commend (or criticize) that agent in a way that presupposes that there is a transcendent norm that one be a responsible believer. Plausibly, when we charge someone with being epistemically or morally irresponsible we mean to imply that there are reasons for the person to change her ways, and that these reasons are categorical. This follows from the fact that both epistemic and moral norms are transcendent. But it is far from clear that the only proper use of the term '(ir)responsible' is a normative one, i.e. one which invokes transcendent norms and categorical reasons. The error theorist can maintain that she uses 'responsible' (and 'irresponsible') in a purely descriptive fashion that indicates that the agent meets (or fails to meet) the standard for being a responsible believer, where the standard for being a responsible believer can be cashed out in purely descriptive terms. Such a standard might involve not consciously believing contradictions; believing the logical implications of one's beliefs if and when one contemplates these beliefs and the logical relations between them; not believing propositions one takes to be unsupported by evidence; and so on.[26] According to this purely descriptive use of the term

[26] One might object that to fail to meet such standards is to violate requirements of rationality and that such requirements are normative in the sense that there is categorical reason to comply with them. But it is in fact highly doubtful whether there are categorical reasons to comply with rational requirements. For an influential argument to the effect that rationality is not normative in this sense, see Kolodny (2005). Broome (2007b) is officially agnostic on the matter; he concludes that he is unable to find an argument that establishes that rationality is normative.

'responsibility', the norms associated with epistemic responsibility are immanent rather than transcendent. To be clear, the suggestion is not that according to the folk understanding of epistemic responsibility, epistemic norms are immanent rather than transcendent; the suggestion is that there is a purely descriptive notion of epistemic responsibility that error theorists can invoke.

To give an analogy meant to show that such a purely descriptive use of the term 'responsible' is not an *ad hoc* stipulation, suppose that the standard for being a responsible Mafioso involves not letting squealers go unpunished. Then to call someone a responsible Mafioso is not necessarily to commend his behaviour but simply to make the descriptive claim that that agent does not fail to punish squealers and therefore meets the standard of being a responsible Mafioso. Similarly, to call someone an irresponsible believer is not necessarily to reproach her for failing to comply with categorical epistemic reasons but simply to make the descriptive claim that she fails to meet the standards of being a responsible believer. Error theory, then, does not rule out the possibility of standards of epistemic merit and demerit. It says that in order for claims about epistemic merit and demerit to be true, they must be understood as purely descriptive claims. Error theory does rule out, of course, categorical reasons to display these merits and avoid the demerits.

Cuneo also claims that epistemic error theory invites a Moorean-style paradox of the following sort (Cuneo 2007: 118):

| | |
|---|---|
| (M) | Epistemic error theory is true, but there is no reason to believe it. |

But as long as the distinction between arguments to the effect that a proposition is true and arguments to the effect that there are categorical reasons to believe that proposition is kept in mind there is from the point of view of error theory no paradoxical nature of (M) to be recognized. Again, if epistemic error theory is true there are no categorical reasons to believe anything, including the theory itself.

Critics might insist that (M) is surrounded by an air of paradox that calls for explanation. Perhaps the fact that by the theory's own lights a claim such as (M) 'cannot even rise to the level of paradox' (Cuneo 2007: 119) is embarrassment enough for epistemic error theory. But provided that the aforementioned ambiguities of the term 'reason' are appreciated such an explanation is available to epistemic error theorists. They can hold that a speaker who acknowledges acceptance of the first conjunct of (M)

conveys that there are, from the speaker's first-person perspective, grounds on the basis of which she accepts it. In other words, she conveys that there are arguments to the effect that epistemic error theory is true and on the basis of which she accepts the proposition that epistemic error theory is true. The speaker conveys this because if she failed to believe that there are such arguments she would believe a proposition she takes to be unsupported by evidence and she would thus fail to meet the standard of being a responsible believer, as described above. Thus, so as long as 'reason' in (M) means something like 'argument or evidence to the effect that the proposition expressed in the first conjunct is true', the air of paradox surrounding (M) is explained by the fact that anyone who accepts (M) fails to be a responsible believer (with respect to whether epistemic error theory is true). But again, whether a speaker does or does not meet the standards of being a responsible believer is, according to error theory, a purely descriptive, non-normative matter.

## 4.4 Epistemic error theory and the normativity of belief

We have seen that much of Cuneo's argument against epistemic error theory relies on the claim that there is an intimate link between evidence that *p* and reasons to believe that *p*. This suggests that claims about evidence are themselves normative claims. In response, I have argued that it has yet to be shown that evidence is normative, and that even if there is a normative notion of evidence there is also a non-normative notion to which the epistemic error theorist can appeal in order to distinguish between evidence or arguments to the effect that *p*, and reasons to believe that *p*. But perhaps there is an intimate tie between categorical reasons for belief and the concept of belief, which makes the concept of belief itself – and belief ascriptions themselves – normative. This is a view that has become increasingly popular in recent literature and it is also one with which Cuneo expresses sympathy although he does not pursue it in detail (Cuneo 2007: 122).[27] To have a name for it we might call it *normativism* about belief.

Normativists might hold that belief ascriptions are constitutively normative, to the effect that ascribing to some agent A the belief that *p* is not only to ascribe to A the attitude of taking it to be the case that *p*, but also to claim that A is subject to certain norms to the effect that there is reason for

[27] See, for example, Kim (1988). For recent defences of normativism about belief, see, for example, Karlander (2008); Shah and Velleman (2005); Wedgwood (2007b). For recent critiques, see, for example, Bykvist and Hattiangadi (2007); Gluer and Wikforss (2009); Steglich-Petersen (2008a).

A not to believe simultaneously that not-*p*; that there is reason for A not to disbelieve anything that she believes follows logically from *p*; that there is reason for A no longer to believe that *p* in the face of strong evidence that not-*p*; and so on. According to normativists, norms such as these serve to distinguish belief from other propositional attitudes that involve the attitude of *taking it to be the case that*, such as conjecture, imagining, supposition, wishful thinking, and the like.

According to normativism, then, belief ascriptions entail claims about what there is reason for agents to believe and not to believe. Since epistemic error theory holds that no claims about what there is reason for agents to believe and not to believe are true, normativism seems to imply that according to epistemic error theory no belief ascriptions are true. This would indeed be an undesirable result for epistemic error theorists.

The simplest and most plausible response on behalf of the epistemic error theorist is to agree with the normativists that norms of the kind mentioned above are distinctive of belief and belief ascriptions. But the error theorist should insist that these are immanent and not transcendent norms. In order to distinguish belief from other attitudes such as conjecture, imagining, supposition, wishful thinking, and the like, it is not necessary to assume that the norms associated with belief ascriptions are transcendent; it suffices to assume that they are immanent.

Above I suggested that someone who believes that *p* on grounds of what she takes to be insufficient evidence, or even in the face of strong contrary evidence, fails to meet the standards of being a responsible believer (with respect to *p*). We see now that an agent who clings to the belief that *p* in the face of strong evidence that not-*p* might fail to meet the standard not only of being a responsible believer (with respect to *p*) but of believing that *p* in the first place. She might rather be engaged in wishfully thinking that *p*. If she repeatedly violates this norm, and others associated with the standard of belief, not only with respect to her belief that *p*, she might be considered to possess (very) few beliefs and possibly to possess no beliefs at all. But none of this suggests that ascribing to an agent the belief that *p* entails ascribing to that agent a categorical reason to give up the belief that *p* in the face of strong evidence that not-*p*.

By way of analogy, suppose a soccer player (who is not a goalkeeper) picks up the ball with her hands during play. Since she is violating a norm of soccer she is not in that instance playing football or perhaps she is playing outrageously poorly, relative to the rules or objectives of soccer. If she repeatedly violates this norm and others associated with the game of

soccer she will be considered not to be playing soccer at all.[28] It is obvious that relative to the rules of soccer, there are reasons for players who are not goalkeepers not to pick up the ball with their hands during play. Similarly, relative to the norms of belief, there are reasons for agents not to cling to the belief that *p* in the face of strong evidence that not-*p*. But this only means that if an agent violates this norm she might not be counted as believing that *p* (rather than, for example, wishfully thinking that *p*) so the reasons in question are non-categorical. There need be no suggestion that belief ascriptions entail categorical reasons. Epistemic error theorists are thus not committed to the arguably unpalatable result that no belief ascriptions are true.

## 4.5 SUMMARY

I accepted as a starting point of this discussion the claim that if categorical moral reasons are queer, categorical epistemic reasons – reasons for belief – are equally queer. I have explored the question of what error theorists should say about reasons for belief. In particular, I have discussed what Terence Cuneo claims to be three undesirable results of epistemic error theory, namely that epistemic error theory is self-defeating or polemically toothless, that epistemic error theory implies that there can be no arguments for anything, and that epistemic error theory rules out the possibility of epistemic merits and demerits. I argued that epistemic error theory does not have the alleged undesirable results. Finally, I explored the implications for epistemic error theory of the claim that belief ascriptions are themselves normative. I argued that for the purpose of distinguishing belief from other kinds of attitudes the norms involved in belief ascriptions need not be understood as transcendent rather than immanent and consequently there are no worrisome implications for epistemic error theory.[29]

[28] Jamie Dreier gives a similar argument in response to the normativist challenge against metaethical expressivism. See Dreier (2002: 140–3).

[29] Earlier versions of this chapter have been presented at seminars in Oxford and Stockholm. It has benefited greatly from comments and criticisms by Lars Bergström, Michael Blome-Tillmann, John Broome, Krister Bykvist, Alex Erler, Daan Evers, Guy Fletcher, Hallvard Lillehammer, Brian McElwee, Anders Nes, Ragnar Ohlsson, Veronica Rodriguez-Blanco, Howard Sobel, Bart Streumer, Torbjörn Tännsjö, Folke Tersman, the editors of this volume, and an anonymous reviewer for CUP. I am especially grateful to Levi Spectre for helpful discussions and to Terence Cuneo for his very helpful written comments.

CHAPTER 5

# *Can reasons for belief be debunked?*

*Nishi Shah*

## 5.1 THE CLASSIC DEBATE: EVIDENTIALISM VS. PRAGMATISM

Tom has mounting evidence that he has incurable cancer, but he also believes that he would be happier, regardless of the truth, were he to believe that he is healthy. W. K. Clifford, who famously claimed, "It is wrong, always, everywhere, and for anyone, to believe anything upon insufficient evidence," would, depending upon the strength of Tom's evidence, direct him to believe that he has incurable cancer, no matter the results for his happiness. The pragmatist William James, on the other hand, might deem this one of those situations in which it is permissible to follow one's passions, and therefore might advise Tom to trust in his health in the face of the evidence that he is unwell.

What makes the debate between evidentialists and pragmatists so vexing is the difficulty of finding neutral ground upon which to adjudicate it. Evidentialists and pragmatists simply seem to have different philosophical conceptions of what could count as a legitimate reason for belief, and it is difficult to see what could count in favor of or against their competing normative commitments. Consequently, most arguments for one particular side are dialectically ineffective because they start from assumptions that the other side does not accept.

In previous work I have attempted to break this impasse.[1] I noticed that the philosophical debate between evidentialists and pragmatists about the legitimate *grounds* for belief is mirrored in the conflicting mechanisms by which beliefs are *caused*. These uncontroversial psychological facts about the causes of belief, I argued, provide the clue to a resolution of this long-standing philosophical debate.

No one denies that there are many, often competing, psychological mechanisms implicated in belief formation. Surely those who think that

[1] Shah (2006).

evidence *never* influences belief are mistaken, but it *is* entrenched in common sense that evidence does not always have decisive influence on belief. Accusations that preferences influence belief either directly or indirectly – by blocking an agent from seeing evidence contrary to his preferred views or by making him susceptible to a fallacious line of reasoning – are common. Recent debates about how intelligence and poverty are related and about whether human sexual behavior is biologically determined are rife with such accusations on both sides, and while individual charges of prejudice might be disputed, the legitimacy of this type of criticism is taken for granted by all sides.

At the same time, it is an equally obvious fact that when we consciously attempt to figure out what to believe – i.e. when we *deliberate* – we invariably focus our attention on the evidence. If I want to determine whether to *believe* that Obama will win the election, as opposed to deciding whether to fantasize about Obama winning the election, I don't ask myself whether I or anyone else would be better off if Obama were president. Instead, when I am attempting to figure out what to believe, I feel compelled to look for evidence in support of or contrary to the proposition that Obama will win the election: I open the newspaper, turn on the news, or conduct polls. Similarly, when Tom *deliberates* about whether to believe that he has cancer, he recognizes immediately that this question is to be answered by, and only by, evidence whether he does have cancer. While he may realize that it would make him happy to believe that he is healthy, when he actually attempts to form a belief about the matter, he finds himself restricted to thinking about what is true, not what would make him happy. Were William James to advise Tom that he would be better off believing that he is healthy, he would be offering him advice that Tom could not, as a matter of fact, follow.

While both of the above facts are widely accepted, it has not been appreciated how difficult it is to construct an account of belief that accommodates both of them. I believe that the best explanation for these psychological facts decides the debate in favor of the evidentialists. Unlike other belief-forming processes such as sensory perception, in which, for example, one comes to believe that it is snowing outside solely on the basis of seeing that it is snowing outside, deliberation involves the subject's application of the concept of belief. And the concept of belief, I argue, contains within it the normative standard that a belief is correct if its content is true, and incorrect if its content is untrue. Thus when one deliberates whether to have an attitude conceived as a *belief*, one deliberates about an attitude to which one applies this standard, and so, as Clifford insisted, one is

committed to considering it with an eye exclusively to whether it is true. In contrast to other accounts of deliberation, my account acknowledges the indisputable fact that beliefs can be influenced by non-evidential considerations, because my view entails only that one is forced to apply belief's standard of correctness in situations in which one exercises the concept of belief. Not all belief-forming processes require the subject to deploy the concept of belief, so the normative standard of truth that controls deliberation needn't causally control other processes, such as wishful thinking. My account thus explains the exclusive role that truth plays in governing deliberation while allowing for the fact that our passions sometimes influence belief.[2] Since the debate between evidentialists and pragmatists is not about the causes of belief, but instead about what normative standard to apply in determining what to believe, we have no choice but to accept the evidentialist standard. Whether one explicitly recognizes it or not, one is committed to evidentialism just in virtue of asking what to *believe*.

## 5.2 A MORE RADICAL OPPONENT: THE ERROR THEORIST

My concern here is not to re-argue my case against pragmatism, but to confront a more radical opponent of evidentialism. The opponent I have in mind is an *error theorist* about doxastic reasons. He rejects evidentialism not because he thinks there are pragmatic reasons for belief, but because he thinks that there are no reasons for belief at all. He claims that our practice of reason-attributions is defective in the same way that the practice of witch-attributions is defective. Just as no persons have the property of being a witch, so there are no facts or states of affairs that have the property of being a reason for belief. Any judgment that attributes to something the property of being a reason for belief is, therefore, false.

The error theorist may accept the following conditional:

If there are reasons for belief, then they are constituted by evidence.

What he denies is the antecedent of the conditional. In the debate between evidentialists and pragmatists, both sides accept that there are reasons for belief. Arguing for evidentialism in this context thus does not require arguing for the antecedent of the conditional, and most defenses of evidentialism in fact do not attempt to do so. The error theorist may be happy to grant that the evidentialist is correct about what reasons for belief would be if there were any. But this concession does not move him any distance

[2] This is just a brief summary of an argument I give elsewhere. See Shah (2003).

toward accepting evidentalism, just as conceding that if God exists God is omniscient does not move an atheist any distance toward believing in God. Addressed to the error theorist, arguments favoring evidentialism over pragmatism merely beg the question.

The error theorist, though, goes beyond merely denying that there are reasons for belief. He claims that there is a practice that we engage in which commits us to judgments about doxastic reasons, but this practice is based on an error because there are no doxastic reasons. This involves more than the assertion that propositions attributing reasons for belief are all false. Think of a topic about which we are all error theorists, witch-discourse. We think that people who believed (or continue to believe) that certain people are witches were mistaken. To arrive at this claim, we must be able to attribute these beliefs to certain people. Unless we are able to identify practices in which some people held such beliefs, there would be no sense to be made of our claim that anyone has ever been *mistaken* about the existence of witches. At a minimum, to arrive at an error theory about reasons for belief, those who accept the conditional above must argue for the following two claims:

**1.** There is a practice of making judgments of the form "S possesses evidence that gives S a reason to believe that *p*."

**2.** Judgments of the form "S possesses evidence that gives S a reason to believe that *p*" are systematically false.

There are two ways that an error theorist can argue for the second claim. One is to argue that although evidence that *p* is a reason to believe that *p*, we can never be in possession of such evidence for any of our beliefs. Thus any judgments we make that someone has a reason for believing some proposition are false. The other way to argue for the second claim is to leave unchallenged claims that we possess evidence for the truth of some propositions, but to deny that the possession of that evidence gives us a reason to believe those propositions on the grounds that there are no normative properties. If there are no normative properties, then evidence fails to have the normative property of rationalizing belief.

Once one seriously considers the scope of the error theorist's thesis, it is easy to see that arguments of the first type are self-defeating. The error theorist is claiming not merely that there is some class of propositions (e.g. about the external world) that we are evidentially unwarranted in believing, nor is he merely claiming that we are unwarranted in believing all the propositions that we in fact do believe; he is claiming that there are no possible propositions that we could be evidentially warranted in believing. Any argument for this very general thesis is self-defeating on

its face. There could not be good evidence that no possible beliefs could be evidentially justified, since if there were, then there would be good evidence for at least one possible belief, namely, the belief that all beliefs are evidentially unwarranted. If there really is a problem about how there can be good evidence for belief, it is a problem for the error theorist as much as anyone else.

What about the second type of argument for an error theory of reasons for belief, which purports to reject the claim that there are reasons for belief on the basis of an argument that there are no normative truths about what one ought to believe, which itself is likely part of a larger argument that there are no normative truths about anything? This will be the focus of the rest of the chapter. In what follows I argue that my previous defense of evidentialism can be used to undermine error theories that are based on this type of argument. I conclude by discussing whether my refutation of the error theorist is an illegitimate form of transcendental argument.

## 5.3 DEFEATING THE ERROR THEORIST

Someone who argues that there are no true judgments ascribing reasons for belief on the basis of a general rejection of normative facts does not deny that we can ever get evidence for any propositions; in fact, he claims to have evidence that there are no true normative propositions about what one ought to believe. Instead, he rejects the inference from

e is evidence for *p*

to

e is a reason to believe that *p*.

The first claim describes a non-normative relation that one proposition, if true, bears to another proposition; the second is a normative claim that endorses having a disposition to believe some proposition on the basis of some other (true) proposition. This type of error theorist accepts claims of the first type, but argues that the inference is invalid because all normative claims of the latter type are false. The fact that the error theorist's argument for the former claim is itself based on evidence thus does not contradict anything he asserts, since he does not deny that we can have evidence for our beliefs. Nonetheless, I think there is an incoherence lurking in any such argument.

In order for there to even be an inference to be blocked, the claims must be distinguishable. If the concept of evidence is identical to the concept of a reason for belief, then the claim that e is evidence for *p* is not distinct

from the claim that e is a reason to believe that *p*, and the error theorist cannot even begin his argument. Those who think evidence is defined in terms of subjective probability, and that subjective probability is defined in terms of rational credence, will deny the error theorist the ground on which to begin his argument.

But I don't think that evidence can be defined in terms of rational credence. If the proposition that e is evidence that *p* were equivalent to the proposition that e makes it rational to believe that *p*, then the following explanation would make no sense: the fact that the sofa looks green makes it rational to believe that it is green because the sofa's looking green is evidence that it is green. After all, a fact cannot explain itself. But such an explanation does make sense. That is, it makes sense to explain why some piece of information e makes it rational to believe that *p* by pointing out that e is evidence that *p*. So evidence cannot be defined in terms of rational credence. I thus am willing to grant the error theorist the claim that an inference is needed to arrive at claims about reasons for belief from claims about evidence.

Our error theorist must attribute judgments about reasons for belief and demonstrate that these beliefs are systematically false. The question is whether these two conditions can be jointly fulfilled. Let us focus on a particular type of such judgment:

That it visually appears to one that *p* (under conditions C) is a reason to believe that *p*.

What is one doing in making this judgment? Non-cognitivists say that one is expressing a conative (non-cognitive) state that motivates one to believe that *p* when it appears to one that *p*, whereas cognitivists say that one is expressing a belief whose propositional content is that the fact that it visually appears to one that *p* is a reason to believe that *p*. If one claims that the judgment is a non-cognitive state, then the question of its truth or falsity doesn't arise, and claiming that it is false makes no sense. If all judgments about reasons for belief were classified as non-cognitive states, an error theory thus would be unreachable. If one claims instead that the judgment expresses a belief that the fact that it visually appears to one that *p* is a reason to believe that *p*, one must fulfill all the conditions required to make such a belief-ascription. So let us focus then on such a judgment:

John believes that visual appearances are reasons for belief.

What is involved in making this judgment? According to the argument I have given against pragmatism, ascriptions of belief require making normative judgments: to ascribe a belief one must judge, implicitly at least, that the mental state so classified is *correct* if its content is true and incorrect if its content is untrue.[3] Judging that John believes that the fact that it visually appears to him that *p* is a reason for him to believe that *p* thus entails a commitment to the following normative judgment:

John's belief that visual appearances are reasons for belief is correct if its content is true and incorrect otherwise.

This is not equivalent to the claim that John's belief that visual appearances are reasons for belief is true iff it is true. "Correct" does not mean "true." Strictly speaking, propositions, not beliefs, are true or false; beliefs are psychological states whose contents are true or false but are themselves neither true nor false. The claim above states that beliefs are normatively assessable as correct or incorrect in virtue of the non-normative property that their propositional contents have of being true or false.

How are we to understand this commitment to a norm of correctness implicit in the ascription to John of the belief that visual appearances are reasons for belief? If we interpret it as a non-cognitive state, then, as before, the possibility of an error theory cannot arise with respect to the normative judgment expressed by it. To judge that John believes that visual appearances are reasons for belief would not itself be a belief, and thus would be incapable of being true or false. Suppose that the commitment to the norm of correctness is a belief whose content is that the belief one has ascribed has the normative property of being correct if its content is true and incorrect if its content is untrue. Thus, judging that John believes that visual appearances are reasons for belief would require believing the following normative proposition:

John's belief that visual appearances are reasons for belief is correct if visual appearances are reasons for belief and incorrect otherwise.

If one were an error theorist about normativity in general – believing amongst other things that there is no instantiated normative property of correctness and thus no true propositions attributing this property to anything – one therefore would be committed to believing the following two inconsistent claims:

[3] My purpose here is not to defend my previous argument against pragmatism, but to demonstrate that, if correct, it defeats the error theorist as well.

There are no true normative propositions.

John's belief that visual appearances are reasons for belief is incorrect.[4]

The second claim is itself a normative proposition that must be false if the first claim is true. And this case obviously generalizes to attributions of all normative beliefs. Attributing normative beliefs to others commits one to believing normative propositions oneself.

Even someone who attempts to argue for an error theory about doxastic reasons without committing himself to a more general error theory about normativity is in trouble. Unlike an error theorist about normative judgments in general, the error theorist I have in mind does not reject all reason attributions, he only rejects attributions of reasons for belief. Nonetheless, he will have to attribute beliefs about doxastic reasons to others; otherwise he will have no subject matter about which to be an error theorist. In order to attribute the beliefs that he rejects, he must commit himself to the judgment that such beliefs would be correct if their contents were true. Attributing beliefs of any kind, including beliefs about doxastic reasons, thus commits him to the following general claim:

If S possesses evidence that *p*, then S possesses information that indicates that believing *p* would be correct.

A reason for a belief just is a consideration that indicates that the belief would be correct, therefore the above claim entails that:

If S possesses evidence that *p*, S has a reason to believe that *p*.

But this just is the type of claim that the error theorist about doxastic reasons sought to deny. Unlike the error theorist about evidence, he is not an error theorist about judgments such as "S possesses evidence that *p*." Instead, he denies that evidence constitutes a reason for belief, which is just to deny that if someone possesses evidence for a proposition then the person possesses a reason to believe that proposition. By attributing beliefs about doxastic reasons, the error theorist about doxastic reasons thus commits himself to the falsity of his own view.

[4] The inconsistency can be made more transparent if we frame the two claims the error theorist is committed to in terms of properties. Here are the two claims reformulated:

There are no normative properties.
The belief that lying is wrong has the normative property of being incorrect.

Those whose metaphysical scruples cause them to recoil at talk of the existence of properties obviously will reject this formulation of the error theorist's commitments.

The error theorist about doxastic reasons faces a dilemma. He claims that all normative judgments attributing doxastic reasons are false, that there are no normative truths that would make some of our normative judgments about reasons for belief true. He must decide whether these normative judgments express beliefs or not. If he decides that they do not, then he must admit that normative statements about doxastic reasons, contrary to surface appearances, do not express propositions. If normative statements about reasons for belief do not express propositions, then it is not the case that they are false, contrary to the error theorist's central contention. But if the error theorist decides that normative judgments about doxastic reasons express beliefs, and attributing a belief entails making a normative judgment, he cannot consistently claim that there are no reasons for belief.

## 5.4 A TRANSCENDENTAL ARGUMENT?

The preceding argument against an error theory about doxastic reasons depends on the claim that ascriptions of belief involve a commitment to truth as the standard of correctness for belief: to ascribe the belief that *p* entails a commitment to the judgment that the belief is correct if *p* and incorrect otherwise. My argument for this claim is an inference to the best explanation: the best explanation of the fact that deliberating whether to believe that *p* requires determining whether *p* is that one applies the norm of truth to one's reasoning in exercising the concept of belief.

It might appear that this argument, even if correct, only yields the conclusion that first-personal deliberation about what to believe involves an application of the norm of truth; it does not yield the further claim that third-personal ascriptions of belief entail any such normative commitment. After all, the phenomenon I sought to explain is essentially first-personal – our experience of what it is like to deliberate about what to believe. Why should the explanation of it have implications for the contents of the third-personal point of view taken up by the error theorist in ascribing beliefs?

This objection is of a piece with objections to Kantian strategies that seek to ground the objectivity of ethical norms in features that are constitutive of the first-personal experience of being an agent. Even if Kant's claim that one cannot be a practical agent – someone who experiences his choices as free – unless one accepts the Categorical Imperative (CI) is true, it does not follow that we ought to follow the CI or that the propositions about what we have practical reason to do that are derivable from the CI are

true. To show that a certain normative judgment is inescapable from the first-person point of view is not tantamount to showing that it is true. Our experience of freedom may just be an illusion built into the very fabric of our agency.

If Kantian arguments can at best achieve conclusions limited to claims about what must be accepted within the first-person point of view, these conclusions will be vulnerable to error theories. Error theories spring from doubts that arise when we attempt to take up an impersonal point of view on the universe, including that part of the universe that includes creatures with subjective points of view. From the impersonal point of view, our normative judgments can appear to be of a piece with judgments about unicorns or witches, judgments that have no chance of being true. Claims about what must be accepted within the subjective point of view that we inhabit as doxastic and practical agents cannot undermine these doubts. Those who are gripped by such doubts will not see these Kantian claims as a vindication of our moral or epistemic practice; they will see these conclusions as condemning us to inescapable error.

Transcendental arguments seek to vindicate the objective validity of our fundamental normative judgments by bridging the gap between the first- and third-personal points of view. They aim to establish results about how things are on the basis of claims about what we must believe to be either doxastic or practical agents. I, like most philosophers, am skeptical that any such argument can be made to work. Once the Kantian makes the bifurcation between the first- and third-personal points of view, no argument can bring them back together. Psychological claims about what people believe, even if they state necessary truths about agents, do not imply conclusions about the truth of the contents of those beliefs.

As long as the results of Kantian arguments are confined to claims about what norms must be accepted from the first-personal point of view, error theories about those norms appear to be left open. But the long history of failed transcendental arguments is strong evidence that attempts to go beyond these results to prove the "objective validity" of these norms are hopeless. So how can I claim to have refuted an error theory about doxastic reasons on the basis of claims about what one must accept to engage in deliberation about what to believe, which is an essentially first-personal phenomenon?

I reject the claim that my argument is limited to conclusions about the first-personal point of view. On the other hand, I don't think that my argument establishes, as transcendental arguments attempt to do, any substantive normative results. Let me explain.

Why think the best explanation of first-personal *deliberation* about what to believe commits us to claims about the third-personal *attribution* of beliefs? The answer lies in recognizing that while deliberation about what to believe requires deliberating about what is true, other kinds of deliberation, such as deliberation about what to imagine or suppose, do not require deliberating about what is true. When a mystery writer deliberates about whether to imagine that the butler did it, he obviously does not ask himself whether the butler really did it; he asks whether portraying the butler as the guilty party would be the most satisfying resolution of the plot. This difference between doxastic and other kinds of deliberation requires explanation. If we accept that the best explanation of why deliberating whether to believe that *p* gives way to deliberation whether *p* is that in asking the former question one applies the norm of truth, then the question becomes why we apply this norm when we ask what to believe but not when we ask what to imagine or suppose.

Once the question is posed in this way, it answers itself. What distinguishes deliberation about what to believe from deliberation about what to imagine or suppose is that it is deliberation about what to *believe*; that is, asking the question that frames doxastic deliberation, as opposed to other kinds of deliberation, requires exercising the concept of belief. The source of the application of the norm of truth thus must lie here, in the exercise of that concept. It is because the concept of belief includes the norm of truth that when we exercise this concept we apply this norm. But if the norm is contained in the concept of belief, then we are committed to applying this norm whenever we exercise the concept – in first-personal deliberation about what to believe and third-personal attributions of belief alike. To claim that we are committed to applying the norm only in first-personal deliberation is to deny that the norm is included in the concept of belief; and to deny this is to leave it a mystery why one specific kind of deliberation – deliberation about what to believe – and not deliberation about what to imagine or suppose – involves the application of this norm.[5]

[5] Shouldn't the error theorist reject my explanation of doxastic deliberation, as others have, since accepting it prevents him from coherently stating his theory? He may well want to reject my theory, but I doubt that he has a good reason to do so. Remember, an error theory is grounded in metaphysical doubts about the existence of normative facts. But my explanation of doxastic deliberation does not make any attributions of normative facts. All it says is that in order to exercise the concept of belief one must make normative judgments. The explanation does not itself involve any normative judgments. So the error theorist cannot reject my explanation on the basis of his denial of normative facts, since I have not asserted that there are any normative facts. For an alternative explanation of the truth-directed nature of doxastic deliberation see Steglich-Petersen (2006a).

But haven't I just made a transcendental argument, moving from a claim about what one must accept to be an epistemic agent to a conclusion about the correctness of what one accepts? From the claim that one must accept truth as the standard of correctness for belief when one engages in doxastic deliberation haven't I claimed to establish the result that truth is the standard of correctness for belief?

No. I have argued that in order to be an epistemic agent or to attribute beliefs one must accept the norm of truth. But I have not argued that these normative commitments are true beliefs about an independent normative realm, because I have not argued that any attributions of belief are true. What I have argued is that one cannot coherently judge that people have beliefs about the norm of truth or reasons for belief and reject the norm of truth or deny that there are reasons for belief. From this nothing follows about whether anyone has ever believed anything or whether it is even possible for anyone to believe anything; thus nothing follows about whether anyone has ever had or could ever have a reason for belief.

Is there a way to underwrite an inference from the fact that we commit ourselves to normative judgments in attributing beliefs to the truth of those judgments? Well, if normative truths were themselves metaphysically constituted by our normative judgments, as constructivists contend, this inference would be warranted. The most general version of metaethical constructivism takes the following form:

**Metaethical constructivism**:

The property of being F, where "F" stands for a normative predicate, is such that for any x, if x is F, then x is F because S would *believe* that x is F (under non-normative conditions C).

If the requisite non-normative conditions are fulfilled, someone's judging that they believe that *p* and thus that they are in a state of mind that is correct iff *p* makes it the case that they are in a state with a correctness condition. If constructivism is true, mere attributions of belief are sufficient to refute the error theorist.

Whether the error theorist can be refuted in this way depends on whether constructivism really can be combined with the claim that it is constitutive of being a belief that it is correct iff its content is true. Unfortunately, trying to combine these two views leads to a vicious regress.

If the normative property of being correct if and only if its content is true is what distinguishes belief from other attitudes, then being a belief is a normative property and the constructivist owes us an account of it. That

is, the constructivist owes us an account of the property of being a belief. The constructivist schema as applied to beliefs would have the following structure:

**Belief constructivism**:

The property of believing that *p* is such that, for any subject R, if R believes that *p*, it is the case that R believes that *p* because S would *believe* that R believes that *p* (under conditions C).

Since belief facts are themselves normative facts, in order to complete this account, the constructivist needs to provide an account of the fact that S believes that R believes that *p*. The fact that S believes that R believes that *p* therefore itself must be constructed out of further belief facts if a fully general constructivism about normativity is correct. The account thus has the following structure:

The fact that S believes that *p* is constituted out of the fact that T would believe that S believes that *p* (under conditions C), which is itself constituted out of the fact that U would believe that T believes that S believes that *p* (under conditions C), etc. . . .

We end up with a regress, with belief facts being constructed out of further belief facts, which in turn must be constituted out of further belief facts, and so on *ad infinitum*. The problem with such a regress is that it prevents the constructivist from establishing that the normative judgments involved in belief-attributions are metaphysically prior to the beliefs that they attribute. Such a regressive formula at best describes a necessary connection between belief-attributions and beliefs. Such a necessary connection, though, is what constructivism was meant to underwrite, not merely to restate. If this regressive formula is the most the constructivist can provide, then he cannot justify inferring the existence of any beliefs, or any normative truths entailed by the existence of beliefs, from the fact that we attribute beliefs.

Defeating the error theorist, though, does not require achieving such a strong result. The error theorist attempts to take an impersonal, detached perspective on human beings, and from this perspective he attempts to pass judgment on our normative judgments about doxastic reasons. This, I claim, he cannot coherently do. In order to attribute judgments to us that he can then condemn as false, the error theorist must think of them as beliefs. Conceiving of them as beliefs, though, requires applying the norm of truth to them. Given that our error theorist is not skeptical about evidence, he will be committed to the existence of reasons for at least some of our beliefs. He thus cannot coherently find us to be systematically in

error in our attributions of reasons for belief or in our normative judgments in general.

Attributing beliefs about anything, including beliefs about doxastic reasons, is inconsistent with the complete denial of doxastic reasons. But this is just what the error theorist attempts to do. The error theorist does not just assert that there are no reasons for belief; he goes further and claims that our practice of making judgments about reasons for belief is therefore defective. Unless our practice commits us to believing that there are doxastic reasons, the former claim does not entail the latter claim. To reach his conclusion, the error theorist thus must attribute beliefs about doxastic reasons to us, but, as I have argued, this very attribution commits him to the existence of doxastic reasons. The error theorist cannot reach his threatening conclusion that our practice of making normative judgments about beliefs is defective without denying the claim on which this conclusion is based. By attributing beliefs to us, the error theorist must abandon his normatively detached point of view and accept the existence of reasons for belief.

PART II

# *Reasons and epistemic justification*

CHAPTER 6

# *Reasons and belief's justification*

*Clayton Littlejohn*

## 6.1 INTRODUCTION

There's little to say about justification that cannot be said in terms of reasons. If you Φ and thereby do all that the reasons demanded, it cannot be that you oughtn't have Φ'd. If it's not the case that you oughtn't have Φ'd, there's no further question as to whether your Φ-ing was justified. Having done all that you should, there's no reason left standing to demand that you refrain from Φ-ing. There's nothing that could stand in the way of the justification of your actions or attitudes. Here's our first connection. If you've done all that the reasons demand, nothing prevents your actions or attitudes from being justified. On the other hand, we shouldn't believe or act without adequate justification. When you oughtn't to Φ, there's an undefeated reason that demands that you refrain from Φ-ing. It's because of this reason that Φ-ing is beyond justification and the most you could hope for is an excuse. Think of cases of conflicting reasons. The only way to justify Φ-ing in the face of reasons that count against Φ-ing is to point to considerations that defeat these reasons. If the reasons you need cannot be found (i.e. reasons that defeat the case against Φ-ing), we can say that Φ-ing cannot be justified because there are reasons that demand that you not Φ that are not defeated. Here's our second connection. Fail to do what the relevant reasons demand and your actions and attitudes are beyond justification.

If we assume this much, there's little that distinguishes doing all that the reasons require from acting or believing with justification. So, we ought to be able to work from an account of reasons and their demands to an account of justification.[1] To do this, we have to settle a question about

[1] The conclusions I'll defend concern the justification of belief, the examples will typically concern reasons for action. Like Alston (1988), I think we have a firmer grip on what reasons for action require. Also, at a certain level of abstraction we ought to expect that reasons for action and belief demand similar sorts of things. As Gibbons (forthcoming) puts it, the similarities between reasons for action and belief have a built-in explanation the dissimilarities lack because both are *reasons*.

reasons. We know that reasons are demanding things, but what do they demand? Maybe reasons are reasons to conform.[2] If there's reason for you to Φ, its demands are met iff you Φ. Critics say that this account misrepresents their demands. According to one objection, it would be unreasonable for reasons to demand full conformity, so they must demand something less.[3] According to another, mere conformity doesn't ensure that your actions or attitudes are justified because you can conform to some reason quite in spite of your deliberative efforts.[4] Reasons demand more than conformity.

In the first part of the chapter, I'll look at some competing accounts of reasons and their demands. It's fair to criticize the conformity account on the grounds that justification involves more than merely conforming to the demands of the undefeated reasons. Normative appraisal is not wholly unconcerned with the reasons for which we act and believe, a fact that's hard to square with the conformity account. It's a mistake, however, to say that normative appraisal is concerned only with the reasons for which we act and believe. Reasons demand full conformity but don't demand compliance. The right account occupies a middle ground between the conformity and compliance accounts.

In the second part of the chapter, I'll argue that there's something wrong with two influential approaches to epistemic justification. I'll focus on evidentialism and the knowledge account because they face structurally similar problems.[5] They're either wrong about which norms govern belief or mistaken about what the reasons associated with these norms demand. There's more to a belief's justification than the evidentialist maintains and less to a belief's justification than the knowledge account says.

## 6.2 REASONS AND DEMANDS

According to the conformity account:

| | |
|---|---|
| I: | In Φ-ing, S does all that the reasons require iff S conforms to the relevant reasons. |

[2] Gardner (2007) and Raz (1990) defend the view. Gardner rejects the view that justification is simply a matter of conforming to undefeated reasons.

[3] Audi (2001), Bird (2007), Cohen (1984), Conee and Feldman (2004), Gibbons (forthcoming), Herman (1994), Langsam (2008), and Wedgwood (2002b) each in their own way suggest that there's no real difference between the justified and the reasonable.

[4] A point stressed by Gardner (2005: 111) who insists that justification isn't just a matter of showing that someone's actions made the world better, but also a matter of showing that the subject's deeds or beliefs could be attributed to the subject's excelling at rationality.

[5] See Adler (2002), Conee and Feldman (2004), Shah (2006), and Steup (2001) for defenses of evidentialism. Williamson (2000) defends the knowledge account as do Bird (2007), Huemer (2007), and Sutton (2005, 2007). Adler (2002) might defend both views.

If you have a reason to Φ, you conform to that reason iff you Φ. We talk as if we have reasons for bringing about certain (external) states of affairs. It's possible to fail to conform to such reasons without thereby being anything less than fully reasonable or responsible. Those who think that reasons cannot demand that we bring about states of affairs on the grounds that someone might try but fail to do so without being anything less than rational or responsible often say that normative evaluation ought to be concerned only with the qualities of someone's deliberative efforts. They likely opt for a view along these lines:

| | |
|---|---|
| II: | In Φ-ing, S has done all that the reasons require iff S is no less than fully reasonable and responsible for having Φ'd. |

Coming from a slightly different direction, some might object to (I) on the grounds that it fails to represent the full range of demands reasons make. While the results of our deliberative efforts do matter, reasons are also there to guide the way we reason, and the conformity account fails to do justice to this. So, someone might opt for the compliance account:

| | |
|---|---|
| III: | In Φ-ing, S has done all that the reasons require iff S complies with the relevant reasons. |

To comply with some reason to Φ is to Φ for that very reason. On this third account, normative appraisal is concerned both with the quality and results of our deliberative efforts. Sympathetic as I am to the idea that normative appraisal is concerned with the quality of our deliberative efforts, I shall argue that it's a mistake to say as (II) does that this is the sole concern of normative appraisal and argue that (III) is wrong in the way it takes the justification of our actions and attitudes to depend on the reasons for which we act and believe. The right account is found somewhere in between (I) and (III).

### *6.2.1 Reasons and rationality*

Some object to the conformity and compliance accounts on the grounds that they represent reasons as making unreasonable demands.[6] It's said that (I) fails to do justice to the deontological character of justification and we're urged to accept (II) in its place. Those who defend the deontological

[6] This complaint is sometimes voiced by those who work on moral luck and deny that consequential luck can determine whether we've lived up to our obligations. See, for example, Nagel's (1979: 31) remark that strict liability might have its uses in the law but is "irrational as a moral position."

theory of justification often say that theirs is a view on which a belief is justified when it is "epistemically permissible, a belief for which the subject cannot justly be blamed, or a belief the subject is not obliged to drop."[7] The argument would be this. If we opt for (I), we have to allow for cases in which a subject who oughtn't to Φ reasonably judges she ought to Φ. Suppose, if only for *reductio*, that such a case is possible. That she oughtn't to Φ would mean that there's an undefeated reason that demands that the agent refrain from Φ-ing. When the agent acts in accordance with her judgment about what she ought to do she will thereby be in breach of her duties or obligations. But, the objection continues, this cannot be. It cannot be that her duty or responsibility was to refrain from Φ-ing because we cannot properly accuse her of any irresponsibility for having Φ'd. Thus, reasons do not demand conformity, as (I) states. If, as a result of our deliberative efforts, we cannot be said to be anything less than fully reasonable or responsible, we cannot fail to do what the reasons demanded from us.

There's little to be said for this argument. The term "responsibility" has a backward-looking and a forward-looking sense. The argument conflates these. It's true that if we oughtn't to Φ, we have an undefeated reason to refrain from Φ-ing. In that sense, refraining from Φ-ing is a responsibility of ours. Here, "responsibility" is synonymous with "duty." It's true that if we non-culpably judge that we ought to Φ and Φ accordingly, we cannot be accused of any irresponsibility. Here, however, "responsibility" is being used with its hypological inflection.[8] To say that we cannot be accused of any irresponsibility in this sense is to say that blame is inappropriate. In the absence of a reason to think that anyone who acts responsibly thereby meets her responsibilities, this argument gives us no reason to accept (II).

To see why reasons demand more than (II) suggests, consider three cases:

| | |
|---|---|
| C1: | Green sees Peacock aim her revolver at Plum and knows that the only way he can prevent Peacock from shooting Plum is by hitting Peacock with a candlestick. He hits her with the candlestick, thus saving Plum's life but injuring Peacock. |
| C2: | White sees Green raising a candlestick with the intention of hitting Peacock but does not realize that Green is trying to protect Plum. Believing that Green is trying to murder Peacock, he picks up a wrench and swings it at Green, mistakenly believing that this is justified on grounds of protecting an innocent party. |
| C3: | In an unrelated incident, Mustard stabbed White. He had gone mad. |

[7] Steup (1999: 375). The argument sketched would be endorsed by those who take the reasonable to be the mark of the permissible.

[8] Hypological judgments pertain to accountability, liability, responsibility, culpability, and the like. See Zimmerman (2002).

Our ordinary scheme of normative appraisal recognizes a distinction between denials of responsibility (e.g. (C3)) and excuses (e.g. (C2)).[9] It also recognizes a distinction between excuses and justifications where a justification depends upon showing that an action is right or permitted (e.g. (C1)). It's possible to maintain these distinctions and thus allow that these three cases illustrate three distinct ways of removing blame only if it's possible for an agent to be no less than fully rational or responsible even if the agent's action is an instance of excusable wrongdoing. An action can only be an instance of excusable wrongdoing, however, if the agent who performs the action fails to do what she ought to have done. An action can be all-things-considered wrong only if the agent who performs them has failed to do what the reasons required. Hence, it's possible for an agent to be fully reasonable and responsible while failing to do all that the reasons required. Thus, we must reject (II).

To defend (II), someone would have to show that our three cases do not illustrate three distinct ways of removing blame. It's clear that (C2) and (C3) differ significantly. To remove blame in (C3), we would have to show that Mustard lacked the capacities necessary for being properly held accountable for his deeds. As White can be held responsible for his deeds, if we are going to argue that White cannot be blamed for his deeds, it seems our only other option is to argue that White's actions, though wrongful, were nevertheless what we would expect from a fully reasonable agent. Given how things seemed to him at the time, we might say, we can see that White's response is no indication that he is anything less than fully virtuous. Presumably it is this that defenders of (II) must deny. They must deny that we excuse someone from wrongdoing by showing that their actions or attitudes were reasonable and that they were fully responsible for having acted or believed as they did. They must insist that this is a justification because defenders of (II) take the mark of permissibility to be an agent's being fully reasonable and responsible for having acted as she did.

This response puts defenders of (II) in a bit of a bind. The mark of permissibility isn't what they say it is. To defend their view, they have to classify the actions in (C1) and (C2) as right, thus denying that (C2) is a case of excusable wrongdoing. This doesn't sit well with intuition. Think about justified intervention. One striking difference between (C1) and (C2) is that it seems that an informed bystander could justifiably use force to assist Green but only justifiably use force to interfere with White. The natural

[9] Strawson (1962) refers to denials of responsibility as "exemptions," but I'll use Gardner (2007) and Horder's (2004) terminology.

explanation for this difference would seem to be that there is a difference in the deontic statuses of Green's and White's actions.[10] If we say both acted rightly, we run out of moral distinctions to use to explain this difference. We should classify (C2) as a case of excusable wrongdoing and reject the link between being fully reasonable and doing what the reasons demand.

### *6.2.2 Conformity or compliance?*

Reasons demand conformity. The question is whether their demands end there. Critics of the conformity account insist that they demand more. Critics of the compliance account insist that they demand less than compliance. The critics are right. We ought to reject (I) and (III). The problem with the conformity account is that it denies that the quality of our deliberative efforts could have any bearing on whether we have done what the reasons required of us, that we could do what the reasons require quite in spite of our deliberative efforts. That's a mistake. Justification does depend on explanatory reasons, but not in the way the compliance account suggests.

According to the compliance account, a reason to Φ demands that you Φ for that very reason. Doing all the reasons require depends on whether you conform to those reasons from those very reasons. This can't be right. Consider cases of overdetermination. Suppose there are two perfectly good reasons for Mustard to turn Plum in to the police. Are we to say that Mustard did anything less than what the reasons required if he turned her in to protect White? Surely not.[11] But, Mustard did not comply with an undefeated reason. If we think that any failure to do what an undefeated reason demands without any reason for so this failure is *prima facie* wrongful, we would have to think of his actions as being *prima facie* wrongful. Or, consider cases where the good reasons to act are explanatorily idle. If Mustard turns Plum in so that he can use the reward to buy himself something rather nice, what of it? Again, his actions aren't *prima facie* wrongful, but he has not complied with any particularly good reason.[12] The lesson is that the failure to comply with undefeated reasons does not by itself show that the agent's actions or attitudes are wrongful. This seems hard to square with the idea that reasons demand compliance.

Those who defend the compliance account insist that reasons demand more than mere conformity. If we know Plum moved her arm because she

[10] As Robinson (1996) argues. Husak (1999) argues that there are problems with the details of the argument and I address his concerns in Littlejohn (MS).

[11] See Raz (1999: 91). [12] See Raz (1990: 180).

tried to sock Green in the jaw, it seems this alone is enough to show that Plum's action is *prima facie* wrongful. We know it is wrongful even if we do not yet know whether she made contact. The point is obvious enough. We can show that someone's actions or attitudes are wrongful even if we do not show that they acted against some undefeated reason. We only have to show that the agent's deliberative efforts were of sufficiently poor quality by showing that the agent either was willing to bringing about the bad or indifferent to the prospect of doing so. The malicious, negligent, and reckless agent fails to do what the reasons demand of them. This seems hard to square with the conformity account because we can easily imagine that there is no undefeated reason to which the agent fails to conform even though the agent is malicious, negligent, or reckless.

The considerations that cause trouble for the conformity and compliance accounts suggest that normative appraisal is concerned with both the quality and results of our deliberative efforts. An account that seems to do justice to this is a modified version of the compliance account that insists that you ought to always act or believe for some undefeated reason.[13] It allows that you needn't act for every undefeated reason. If the reasons for which you Φ correspond to some undefeated reason, you have done all that the reasons require. Do anything less, and you fail to do all the reasons require.

The view is nicely tailored for addressing the problems that beset the conformity and compliance accounts, but the theoretical motivation for the view might seem obscure. It denies that reasons individually demand compliance but insists that there is nevertheless always a demand to comply with some undefeated reason or other, so from where does this demand come? One suggestion is that if we always act or believe on an undefeated reason, we better conform to the demands reasons place upon us.[14] The problem is that it seems to give an instrumental justification for the principle and that's too weak for our purposes. If we say there's overall reason to Φ, it's true that anyone who acts for an undefeated reason will conform to the demands of the undefeated reason. However, it's also true that someone can act for the wrong sorts of reasons knowing the act happens to be what there's overall reason to do. We cannot appeal to an instrumental principle to explain why we regard their actions as wrongful. Suppose Mustard sees Green take aim at Plum with a revolver and clubs him with a pipe knowing

[13] Gardner (2007) defends the view that justification requires acting for an undefeated reason and Raz (1990: 40) defended the view that we always ought to act for some undefeated reason when he first published *Practical Reason and Norms*. He no longer defends the view.

[14] Gardner (2007: 101).

that this sort of thing would be justified on grounds of defending another. Suppose Mustard acts only with the motive of injuring his old enemy. Knowing what he knows, he knows that his act conforms to an undefeated reason. If the principle that you ought always act for some undefeated reason is just an instrumental principle, it doesn't explain why Mustard's action is wrongful.

One problem with the proposal, then, is this. As the motivation for the view cannot stem from a description of the demands individual reasons make, its motivation must derive from elsewhere. The justification that has been offered is instrumental and an instrumental principle is too weak. We want an account that accommodates the idea that normative appraisal is concerned with both the results of our deliberative efforts and the quality of those efforts. The modified compliance account doesn't quite do that, either. For the account asserts that we must always act for some undefeated reason, that the reason for which we Φ corresponds to a genuine reason to Φ. The lack of such a correspondence is not itself any indication that the agent's deliberative efforts are defective in any way. Plum knows she's promised to meet Green for lunch but hears that her former colleague Peacock has fallen ill and rushes to see her. Suppose we learn that Plum acts from a sense of religious duty rather than direct concern for Peacock's welfare. Plum is pious in a way many of us are not, and many of us can see no reason for discharging what she regards as her religious duties. We might say that she's reasonable for having judged that she ought to visit Peacock and no less than fully reasonable or rational for having both acted as she did and for the reasons she did. But, we would be chary to say that her reason for acting corresponded to any real reason for acting. So, we could not say it corresponded to any undefeated reason. Nevertheless, I see no reason to say that she did anything less than what the reasons required of her or that her actions were less than fully justified.

Note two things. First, if she did nothing wrongful, we have a perfectly good counter-example to the principle that states we ought always act for undefeated reason. The reason for which she acted isn't a valid reason. Second, it shows that the rationale for that principle cannot be that normative appraisal is concerned with both the quality of the agent's deliberative efforts and the results of those efforts. For that principle asserts that there must be a correspondence between the reasons for which an agent acts and an undefeated reason. The mere lack of such a correspondence is no indication that there was any defect in the way the agent deliberated. Nor does it indicate that the subject failed to conform to undefeated reasons. So, if the view's motivation stems from the thought that normative appraisal is

concerned with both the quality and results of an agent's deliberative efforts, the view's motivation seems to support only the weaker view that we ought always conform to undefeated reasons while refraining from deliberating in ways that show disrespect for the reason's status as a reason.

Perhaps that's the right view.[15] There are always two ways to go wrong, either by failing to conform to an undefeated reason or by deliberating in ways that manifest willingness or indifference to acting against a reason. Don't go wrong in these ways and we can't fail to be justified. Doing all that the reasons require thus depends on both the explanatory and normative reasons. What distinguishes the modified compliance account from the present account is not this point, a point on which they agree, but the way these accounts take the normative status of our actions and attitudes to depend on explanatory reasons. The present account takes the normative status of our actions and attitudes to be *negatively* dependent on explanatory reasons. The modified compliance account insisted on a correspondence between the reasons for which you Φ and some undefeated reason for you to Φ. It thus makes the normative status of our actions and attitudes *positively* dependent on the reasons for which we act or believe what we do. Remember that the theoretical motivation for these accounts is the same. It's the observation that normative appraisal is concerned with the agent's contribution to action and belief. Objective features of the situation may determine what reasons there are, but they don't tell us that the agent's response to what she took these reasons to be is flawed, defective, or blemished in any way. If we were to say that the normative status of our actions and attitudes were *positively* dependent on explanatory reasons, insisting that we must always act for some genuine undefeated reason, this would have to derive from somewhere else. But we've seen that it can't derive from the assumption that reasons demand compliance, because they don't. So, the view that best harmonizes with intuition and the theoretical motivations offered for the rival accounts seems to be the view that a reason to Φ demands conformity and demands refraining from deliberating in ways that show disrespect for that reason's status as a reason.[16]

An example should help. Plum and Green take aim at Mustard and fire their revolvers. One of their bullets strikes, but one misses. The agent that

[15] Sverdlik (1996) defends the view that sufficiently bad motives and intentions can make otherwise justifiable actions wrongful. Like him, I want to say that normative appraisal is concerned with the reasons for which an agent Φ's because it is interested in the quality of the agent's deliberative efforts.

[16] On this view, reasons can both count against Φ-ing while excluding other reasons, namely those that we cannot deliberate from if we take due care to avoid Φ-ing.

shot Mustard acted against the reason associated with the duty of non-maleficence. It's clear that both agents' actions were wrongful. It seems we don't need to think about different features of Mustard to understand why both actions were wrongful. It's not as if the duty of non-maleficence demands that you refrain from shooting Mustard because he's sentient and then some other duty distinct from the duty of non-maleficence demands that you refrain from trying to shoot him or failing to take due care to avoid shooting him in virtue of some features of Mustard other than sentience so that if there were some sentient being that lacked this further feature it would be wrong to shoot it but not wrong to try. There's just the one set of features of Mustard in light of which we can say that they oughtn't to have shot him, tried to shoot him, or fail to take due care to avoid shooting him. If we don't need two sets of features of Mustard to understand why these actions were wrongful, do we really need two different kinds of reasons where Plum failed to meet the demands of the first and Green failed to meet the demands of the second? I think not. We don't need to wait for the ballistics report. We need only the duty of non-maleficence and the reason associated with it to see that both agents' actions were wrongful and why they were. When there's reason not to Φ, the reason demands that you refrain from Φ-ing and demands that you don't deliberate in a way that shows disrespect for the reason's status as a reason. Beyond that, however, reasons demand nothing further.

## 6.3 JUSTIFICATION

Let's assume, if only as a working hypothesis, that reasons typically place upon us a pair of conceptually related demands. If you have reason not to Φ, in the absence of reasons to the contrary, you oughtn't to Φ and oughtn't deliberate in ways that show disrespect for that reason's status as a reason. In other words, you ought not to Φ and exercise due care to see to it that you do not Φ. Combine this account with the account of the connection between reasons and justification outlined in the chapter's introduction, and it seems we ought to be able to work from a description of the norms of belief to an account of justified belief. Associated with any norm governing belief is a reason that demands, *inter alia*, that we conform to that norm. For any type of epistemic reason, assume there's a norm stating the conditions under which there is such a reason. A belief's justification depends on whether a belief conforms to the norms governing that belief and whether the subject can reasonably assume that the belief so conforms. Its justification shouldn't depend on much else.

If there were some consensus view as to which norm or norms governed belief, our work would be done. There is no such consensus view. I'll look at two of the more influential approaches to epistemic justification and argue that these approaches are flawed. The first, the evidentialist account, asserts that a belief's justification is determined entirely by relations between this belief and the evidence an individual has on hand.[17] Facts that do not supervene on an individual's evidence might figure in some kinds of epistemic appraisal, but not appraisal concerned with a belief's justification. The second, the knowledge account, says we ought to characterize justification using the concept of knowledge and work from the assumption that knowledge is the norm of belief.[18] These views, I'll argue, rest on either mistaken assumptions about the kinds of things that reasons demand or mistaken assumptions about the norms governing belief and the kinds of epistemic reasons there are.

### *6.3.1 Evidentialism*

Evidentialists say the considerations that bear on whether to believe *p* consist of considerations taken to bear on the truth of *p* and can only consist of such considerations. Considerations that bear instead on, say, the practical benefits of believing *p*, however, do not give reason to believe *p*. Why don't practical considerations bear on whether we ought to believe? What explains the hegemony of evidence in doxastic deliberation?

Here's one possible explanation.[19] First, it's said that belief is governed by the truth norm:

| | |
|---|---|
| T: | You ought not believe *p* unless *p*.[20] |

Second, it's supposed to be in virtue of our grasp of this normative truth that considerations taken to be irrelevant to the truth of a claim are excluded from deliberation so that only considerations taken to bear on the truth

[17] Adler (2002), Conee and Feldman (2004), Shah (2006), and Steup (2001) defend evidentialist views that differ in various ways but seem to agree that a belief's justification turns entirely on the evidence the believer has that bears on the truth or falsity of that belief.

[18] See Williamson (2000).

[19] Adler (2002) and Shah (2006) offer this sort of explanation. Steglich-Petersen (2006a) argues that you don't need to assume that truth is the norm of belief to explain belief's transparency to truth.

[20] Owing to the factivity of knowledge, anyone who thinks that knowledge is the fundamental norm of belief thinks that truth is a norm of belief. Boghossian (2003b), Velleman (2000), Wedgwood (2002a), and Williams (1973) also defend the view that belief is governed by the truth norm.

of the relevant proposition are included in doxastic deliberation. Third, nothing can be a reason unless it can figure in reasoning.[21] If we add this assumption that a consideration can constitute a reason to believe or act only if it is capable of being a reason for which we believe or act, we have our argument for evidentialism. What we ought to believe is a function of the reasons there are, and considerations unrelated to truth have just been disqualified as potential reasons. Thus, "only evidence for and against the truth of *p* is relevant to answering the doxastic question whether to believe that *p*."[22]

I think the argument fails as an argument for evidentialism even if it rules out the possibility of practical reasons to believe. If belief is governed by the truth norm, there's a reason to refrain from believing *p* if ~*p*. Whether a belief misrepresents how things are is not something that supervenes on an individual's evidence. So, the evidentialist cannot say *both* that belief is governed by the truth norm and that only evidence for and against a belief bears on whether to believe some proposition. The evidentialist cannot motivate the evidentialist view by saying that the truth norm governs belief and then deny that the falsity of a belief gives us reason *not* to believe. If they say that a belief's justification depends on the evidence and not also on the truth, they would have to say that this reason demands something less than full conformity. The problem with this response is predictable: reasons demand conformity.

I don't think that the evidentialist view is that beliefs that contravene the truth norm are justified because evidence gives overriding reason to contravene that norm. The evidentialist seems to think that the justification of our beliefs depends on the evidence and nothing further *because* belief is governed by the truth norm. Norms don't typically give overriding reason for their own violation. There's an important difference between cases of conflicting reasons and the case in which good evidence supports a false belief. In the case of conflicting reasons, we can be fully aware of both sets of reasons and knowingly judge that we ought to act on the undefeated reason. No one thinks that we can be aware of the evidence, aware that there is reason to refrain from believing *p* because of how that belief contravenes the truth norm, and knowingly judge that we ought to believe what the

[21] Shah (2006: 484) attributes this assumption to Williams (1980).

[22] Shah (2006: 498). According to Owens (2000), the amount of evidence necessary for properly settling the question whether to believe might depend on the practical significance of adopting a belief on the matter and closing deliberation. Fantl and McGrath (2002) later defend a similar claim. Shah might say that his conclusion is not threatened by these authors' conclusions by saying that only evidence is relevant in the way *reasons* are to answering the question whether to believe *p*.

evidence supports. If the evidentialist wants to deny that the truth of our beliefs is necessary for their justification, they had better deny that belief is governed by the truth norm.[23]

Some evidentialists seem to do just this.[24] They say that a belief's truth is unnecessary for its justification without having to say that a belief's failure to conform to the norms of belief is irrelevant to the belief's justification. Rather than say that the reasons demand less than conformity, some evidentialists will say that *only* evidential norms govern belief.[25] This move avoids one set of problems, but introduces new ones. Briefly, here are two. First, the truth norm plays an important role in explanations as to why belief is transparent to truth and why only what the agent takes to be truth-related considerations can figure in doxastic deliberation. Second, think about the connection between an agent's beliefs about what she should do and the actions she performs in light of these beliefs. The agent's beliefs about what she should do rationalize intentions that rationalize actions. I think that the intention can rationalize the action, in part, because once the case for intending to Φ is settled, there is no further question to settle to determine whether to Φ.[26] I think the belief that you should Φ is a belief that can rationalize the intention, in part, because once the case for believing that you ought to Φ is settled, there's no further question to settle to determine whether to intend to Φ.[27] If that's right, that suggests that when there is sufficient reason to believe that you ought to Φ, there's sufficient reason to

[23] This comes with costs. It undermines the explanation that has been given for the transparency of belief to truth and the aim of belief.

[24] See Conee and Feldman (2004).

[25] Feldman (1988a: 247) says we might have truth as a goal, but insists that the truth or falsity of a belief has no bearing on what we ought to believe because a belief could be false without our knowing it. This suggests that he thinks that reasons can bear on what we ought to believe only if we have access to these reasons. I'd say that the truth of a belief cannot give you reason to believe unless you're aware of it, but the falsity of a belief can exert its normative force whether you're aware of it or not. It appears Feldman's reason for denying there's a truth norm is that it implies we could fail to do all the reasons require without being in a position to know this. Maybe he adheres to (II).

[26] For discussion, see Hieronymi (2005) and Shah (2008).

[27] Concerning the toxin puzzle, we might say that the reason that you cannot form the "prize winning" intention is that you know that by forming the intention to Φ, there is nothing further to consider in settling the question as to whether to Φ. However, as you know that there *is* something further to consider (i.e. that you would make yourself sick for no reason), you cannot form the intention without adding reasons or manipulating yourself to lose sight of them. If another eccentric billionaire came along and challenged you to come to believe that you ought to drink the toxin without forming the intention to drink, I think we just reproduce the original case one level up and the same points apply. You cannot form the prize winning belief because you know that if you believe you ought to Φ, there is nothing further to consider in settling the question as to whether to intend and whether to act. But, there *is* something further to consider and so you cannot form the prize winning belief without doing the sorts of things that would allow you to form the intention and violating the terms of the arrangement.

intend to Φ and act accordingly. Of course, if there is a decisive case against Φ-ing, there's not a sufficient case for Φ-ing. So, *if* there's a decisive case against Φ-ing, the reasons that oblige you to refrain from Φ-ing would seem to oblige you to refrain from intending to Φ and from believing that you ought to Φ. Since the facts that determine whether you ought to Φ sometimes include facts that do not supervene upon the evidence, these external facts that oblige you to refrain from Φ-ing oblige you to refrain from concluding deliberation by forming the belief that you ought to Φ. That's so even if you have the same evidence as someone in some possible world who knows that they should Φ. Provided that you shouldn't Φ, you shouldn't conclude deliberation by judging that you should Φ.

An evidentialist could deny this, but then they have to give us some explanation as to why beliefs about what ought to be done rationalize intentions.[28] If distinct considerations determine whether to believe you ought to Φ and whether to intend to Φ, it's hard to see why beliefs would play this rationalizing role. If what motivates the evidentialist to deny that belief is governed by the truth norm is the thought that the truth norm is unreasonable in demanding that believers refrain from believing on strong evidence, I think the evidentialist is in trouble. To block the argument above, the evidentialist has to say that facts about what you ought to believe about what you should do are fixed by facts about your evidence even if facts about what you ought to do are fixed by further facts and allow for pairs of cases where some subject ought to believe she ought to Φ and ought to believe she ought to Ψ where she ought to act in light of one belief but not the other. This view is just as unreasonable as the view that states that belief is governed by the truth norm. Imagine an advisor who, upon learning that you do not know what to believe about the choice between Φ-ing and Ψ-ing, says that you ought to think of Φ-ing as a necessary evil. Imagine this same advisor who, upon seeing you Φ, criticizes you for failing to Ψ. You object but the advisor says that she was being perfectly consistent. One bit of advice was about belief and another about action. If that is what your advisor advises, you need a new advisor.

We've already seen that among the facts that determine whether someone should act are often facts about the situation that do not supervene upon the facts that determine what the agent's evidence is. The evidentialist insists that it's only facts that strongly supervene upon the agent's evidence

[28] Feldman (1988c) claims that it's possible for someone to justifiably believe that they ought to Φ when they oughtn't Φ. Along with Broome (2001), I think there's a normative requirement in light of which you oughtn't both: believe that you ought to Φ but not intend to Φ.

that determine what the agent should believe. If facts about how an agent should act do not supervene upon the evidence but facts about what an agent should believe do supervene upon the evidence, it should be possible for the agent to believe she ought to Φ, be obligated to refrain from Φ-ing, but have no reason at all to refrain from believing that she ought to Φ. It's one thing to say that conscience is fallible, but quite another to say that there is an undefeated reason to refrain from acting on a normative judgment that is epistemically impeccable. If the resistance to saying that belief is governed by the truth norm is motivated by the thought that such a norm demands the unreasonable from agents who have good evidence that supports false propositions, surely the norm that enjoins us to refrain from acting on beliefs about what we ought to do even if those beliefs are the beliefs we ought to have is equally unreasonable. So, while I think many of us would want to say that just as when there is decisive reason not to Φ, there is decisive reason not to intend to Φ, when there is decisive reason not to Φ or intend to Φ, there is decisive reason not to believe that you ought to Φ.

The problem facing the evidentialist is this. On the one hand, if you try to motivate the view by appeal to the truth norm, you can consistently say that certain considerations that don't bear on the truth of our beliefs cannot constitute reasons for belief, but defending the further claim that only pieces of evidence can constitute reasons requires a view of the demands of reasons we have already rejected. It requires thinking of reasons as demanding something less than conformity. We can be excused for acting against some undefeated reason, but that is the most we could ever hope for. On the other hand, those unwilling to embrace the conclusion that the falsity of a belief constitutes a conclusive reason to refrain from holding it take such a consequence to be a conclusive reason for denying that belief is governed by the truth norm. Instead, they'll say that belief is governed only by evidential norms and the falsity of a belief is said to have no normative significance. False beliefs are failures, to be sure, but not failures that have epistemic normative significance. The problem with this is that it seems hard to square with the idea that a belief's justification depends on whether it's fit for the purposes of practical deliberation. When a subject cannot both rationally believe a set of claims and refrain from acting, if the agent oughtn't to perform the action, the agent oughtn't to hold these beliefs. A belief is fit for practical deliberation only if it should not be excluded from practical deliberation, and whether a belief should be excluded from deliberation does not depend solely on the evidence the subject has on hand.

### *6.3.2 The knowledge account*

The problem with the evidentialist view is that it restricts the scope of epistemic evaluation to relations between a belief and the evidence an individual happens to have on hand. The knowledge account seems to do justice to the idea that the scope of epistemic evaluation includes relations between a belief and the states of affairs it represents by insisting that belief is not governed just by evidential norms, but also by the knowledge norm:

| | |
|---|---|
| K: | You should not believe *p* unless you know that *p*. |

Because knowledge is factive, the knowledge account takes epistemic evaluation to be concerned with more than relations between beliefs and the evidence we have on hand. It would also be concerned with relations between these states of mind and the states of the world they represent. While this view overcomes a difficulty the evidentialist view cannot, it suffers from difficulties structurally similar to those that beset the evidentialist view. The view is either at odds with independently motivated claims about the norms or belief or at odds with the account of reasons defended in the first part of the chapter.

In its most popular form, the knowledge account asserts that knowledge is the norm of belief but denies that knowing *p* is necessary for justifiably believing *p*:

> Anything short of knowledge is failure. But some failures are worse than others. And in particular some failures can be laid at the door of the believer, because the source of failure is one or more of the believer's mental states, and some failures can be ascribed to mischance, in that the failure is due to some mentally extraneous factor. The role of the concept of justification is to mark the difference between these different sources of failure.[29]

Focus on the last sentence. This can't be right. Anyone who adopts this attitude toward justification has adopted a view on which justification doesn't involve showing that a justified action or attitude does all that the reasons require or has adopted the view that reasons demand something less than conformity. For surely *if* knowledge is the norm of belief, there's a reason to refrain from believing what you don't know that constitutes a conclusive reason unless it's overridden by some stronger reason to contravene the knowledge norm. As no one thinks there is such a reason, a defense of the

[29] Bird (2007: 95).

knowledge account just is a defense of the idea that there is such a reason to refrain from believing what you don't know to be true. So, the only way to make sense of this view is to either say that we can successfully justify Φ-ing in the face of undefeated reasons not to Φ or assert that reasons demand less than conformity. The first option is incoherent. The second option is indefensible for reasons we've discussed already. You either have to deny that knowledge is the norm of belief or accept that a belief's failure to constitute knowledge constitutes a conclusive reason for abandoning that belief. If knowledge is the norm of belief, knowledge that *p* is true is necessary for having a justified belief that *p* is true.

Some embrace the idea that you can't justifiably believe what you don't know, taking this, as I do, to be a consequence of treating knowledge as the norm of belief.[30] The problem with this view is not that it's incoherent. Problems for this view emerge if we think about its implications for Gettier cases. If knowledge is the norm of belief, subjects are obliged to refrain from believing in Gettier cases and so can't believe with justification in such cases. The most the subject could hope for in such cases is an excuse. But, this seems highly counterintuitive. Consider one of your run-of-the-mill Gettier cases. Having just finished her lunch, Plum hands the waitress a ten-dollar bill. She reasonably believes that she's paid her bill. She has paid her bill with genuine currency. However, unbeknownst to her, she's dining in the land of fake ten-dollar bills. She doesn't know she's handed the waitress a genuine ten-dollar bill because there are fake bills in the other diners' pockets that would easily pass for the real thing. So, she doesn't know that she's really paid her bill. Think about her belief that she has paid her bill. Do we really think, knowing what we know, that there is something to be excused here? Knowing what we know, we wouldn't say that she oughtn't believe what she does. There was no breach to excuse,

[30] Sutton (2005, 2007) defends the view. Hawthorne and Stanley (2008) express some sympathy for the idea that *p* is a reason for belief iff *p* is known. In Littlejohn (2009b), I criticize their arguments for the view that it's proper to treat *p* as a reason iff you know *p*. I think they're right that it's proper to treat what you know as a reason and improper to treat *p* as a reason if ~*p*. The reason, I think, that it's proper to treat what you know as a reason is that it's proper to treat what you justifiably believe as a reason and knowledge requires justification. If we adopt a closure principle on which having a justified belief ensures having a reason to believe that which is an obvious consequence of what is believed, belief can be justified only if it is true. Here's the argument. Assuming our closure principle, if S's belief that *p* is justified, S has a reason to believe that which is an obvious consequence of *p*. That *p* is true is as obvious a consequence of *p* as anything could be. Thus, it follows that if S's belief that *p* is justified, *p* is an epistemic reason of S's. But, *p* is an epistemic reason of S's only if *p*. (In Littlejohn (forthcoming), I argue that justifying reasons are factive.) Thus, if S's belief that *p* is justified, *p* is true.

much less justify. The knowledge account says that there are conclusive reasons to refrain from believing where there are none.

### *6.3.3 A remaining option*

We might sum up the problems for the knowledge account and evidentialism as follows. We can think of a belief's justification in terms of our doing all that is demanded by the norms that govern belief. We do all that is demanded if our beliefs conform to the undefeated reasons associated with a norm and have taken all due care to see to this. The problem that the evidentialist faces is that it seems that if they take truth to be the norm of belief, they have to insist that the justification of belief involves more than just relations between that belief and the evidence. If, however, they insist that a belief's justification involves just relations between the belief and the evidence, they either deny that reasons demand conformity or deny that belief is a state governed by the truth norm. They should say neither of these things. The knowledge account faces a structurally similar problem. If they take knowledge to be the norm of belief, they have to say that if you fail to know for any reason, there's a conclusive reason to abandon the belief. To bring that in line with intuition, they have to say that reasons demand less than full conformity. As reasons do demand conformity and there's less to justification than knowledge, knowledge is not the norm of belief.

The norms of belief are concerned with more than just relations between beliefs and bodies of knowledge or evidence but less than that which turns a belief into knowledge. So, what is the norm of belief? Perhaps the fundamental norm of belief is the truth norm, the norm that says that we oughtn't believe $p$ unless $p$ is true. It's often said that belief aims at the truth, and some have suggested that we ought to unpack this metaphor in normative terms. I'm quite happy to do this. True beliefs can do what beliefs are supposed to do. False beliefs cannot. What beliefs are supposed to do is represent how things are so that we might rely on them for the purpose of deliberation. They're supposed to give us the reasons from which we can then reason, and if reasons are the facts represented by beliefs, false beliefs cannot do what beliefs are supposed to because they cannot contribute reasons from which we might deliberate. Rather, they pass off non-reasons as if they were reasons.

Some have argued that you cannot build a theory of justification starting from just the assumption that truth is the norm of belief. That assumption, they say, is far too weak to do the work it must.[31] If we say that truth is

[31] Williamson (2000: 245). I address his argument in detail in Littlejohn (2010).

the norm of belief, it seems that any true belief will be "correct." But then it seems that any true belief will be justified, and this is clearly false.[32] It would be clearly false to say that any true belief is justified, but this isn't what the truth account says. It's hardly news that there can be undefeated reasons to refrain from Φ-ing even if Φ-ing fulfills some goal at which it's permissible to aim. If I'm right about what reasons demand, the reason associated with the truth norm does demand that you refrain from believing falsehoods. In the absence of overriding reason to believe falsehoods, false beliefs will be unjustifiable. If I'm right about what reasons demand, the reason associated with the truth norm *also* demands that you refrain from believing if your believing indicates a lack of due care for conforming to the truth norm. That is to say, it's because you oughtn't believe what's false that you oughtn't hold those beliefs only an irresponsible or irrational person could hold. Insofar as this requires having beliefs backed by the sort of evidence we'd think of as the sort of evidence a reasonable and responsible person would have prior to believing, there's no reason to think that a belief's truth suffices for its justification.

If truth is the norm of belief, we can say that there's reason to refrain from holding those beliefs that fail to faithfully represent how things are and reason to refrain from holding those beliefs for which you are properly faulted for holding. If a belief's justification can be cashed out in terms of a believer doing all that the reasons required, we can say that the justified belief is the faultlessly held belief that faithfully represents how things are. If we think of justified belief as the inner analogue of warranted assertion, then perhaps the following gives some motivation for the view. Plum asserts that Green had planned on killing Mustard for years but missed the opportunity when Mustard died of natural causes. If Green never planned any such thing, Green can say truthfully that Plum shouldn't have asserted this.[33] If we know that Plum had no evidence to back this assertion, this fact alone gives us sufficient warrant to say she oughtn't have asserted this and we can fault her for the assertion. Is there some third way for

[32] Vahid (2006: 305) says that a consequence of Wedgwood's (2002a) view is that a belief is correct (in a normative sense) iff the proposition believed is true and this commits him to saying we ought to believe all truths. I think this is easily avoided. Correctness doesn't entail the non-existence of reasons to refrain. Think about true assertions that express unreasonably held beliefs. Vahid also objects on the grounds that Wedgwood's view implies that no false beliefs are justified. In Littlejohn (2009a), I explain why an externalist view with this implication is consistent with our ordinary intuitions about justification ascription.

[33] Suppose truth is the norm of belief. If assertion wasn't subject to the truth norm, it should be possible for me to permissibly assert *p* knowing that you'd accept my say-so even if it's wrong for you to accept *p*. That seems unacceptable. It's odd to think to yourself that it's not wrong to sincerely assert knowing that it would be taken at face value when it's wrong for others to take the assertion at face value and thereby believe it.

Plum's assertion to be wrongful? I can't think of one. Perhaps an assertion is unwarranted if false, unwarranted if unsupported by the evidence, and not unwarranted otherwise. It's tempting to think that the features in virtue of which she oughtn't have asserted *falsely* that *p* are the features in virtue of which she oughtn't have asserted *without evidence* that *p*. So, it's tempting to think that there's some single set of considerations in light of which something demands that she refrains from asserting the false and asserting without evidence. We have one ground with two conceptually related demands. Further, it's tempting to think of assertion and belief in parallel. If the norms of belief permit belief, they permit asserting that the belief is true. If the norms don't permit the belief, the norms of assertion won't then let you assert what you oughtn't believe. If there's nothing more and nothing less to warranted assertion than what I've suggested, a belief's justification will not involve less than the belief's faultlessly and faithfully representing how things are. Maybe there's not much more to it, either.[34]

[34] Thanks to Robert Howell, Errol Lord, Luke Robinson, Brad Thomson, Steve Sverdlik, and Mark van Roojen for discussing the issues addressed in this chapter. I also want to thank Asbjørn Steglich-Petersen and an anonymous referee for their helpful written comments.

CHAPTER 7

# *Perception, generality, and reasons*

*Hannah Ginsborg*

In the last two decades there has been much debate about whether the representational content of perceptual experience is conceptual or non-conceptual. Recently, however, some philosophers[1] have challenged the terms of this debate, arguing that one of its most basic assumptions is mistaken. Experience, they claim, does not have representational content at all. As they see it, having a perceptual experience is not to be understood on the model of thought or belief, as a matter of the subject's taking things in the world to be this or that way, or of having this or that feature. Nor is it to be understood on the model of receiving testimony about how things are in the world, as a matter of its being represented to us that things are this or that way. Rather, it is simply a matter of our being presented with things. To have a perceptual experience of an object is to stand in a certain kind of relation to it which makes it available to us to be represented in thought or belief, but which does not itself involve our representing it, or its being represented to us.

Much of the motivation for rejecting the view that experience has representational content springs from a concern to do justice to what is distinctive about perceptual experience in contrast to thought and belief. In particular, perceptual experience seems to be more "primitive" than thought, in that it seems natural to appeal to the character of perceptual experience to explain the possibility of thought about objects rather than the other way around. There is thus a concern that we will deprive perceptual experience of its explanatory role with respect to thought if we think of perceptual experiences as representational states on the model of thoughts and beliefs. This kind of concern has played a role in motivating the view that the content of perceptual experience is non-conceptual rather than conceptual. But on the more radical approach at issue, the distinctive

[1] For example Brewer (2004, 2006, 2008); Campbell (2002); Gupta (2006); Martin (2002, 2004); and Travis (2004).

character of perception can be accommodated only by rejecting the idea that experience has content *überhaupt*. And for at least some defenders of the approach, this is required not only to do justice to the explanatory priority of experience over belief, but also to respect the intuitively appealing idea that perception offers immediate acquaintance with the constituents of reality.

My main aim in this chapter is to consider the implications of this approach for the widely acknowledged idea that perceptual experiences can stand in rational relations to belief. Can we deny that experiences have content, while still doing justice to the kind of reason-giving or justificatory relation which has typically been assumed to hold between perceptual experiences and beliefs? I shall argue that we cannot. Taking as my focus Bill Brewer's attack on what he calls the "content view," and his proposal to replace it with an alternative "object view," I shall argue that perceptual experiences must have content if they are to rationalize beliefs, at least in the sense that has traditionally been invoked in the context of epistemological debates about empirical justification. More constructively, however, I shall go on to suggest a version of the content view which aims to address some of the concerns motivating the challenge raised by Brewer and others. I shall argue that we can construe perceptual experiences as capable of standing in rational relations to thought and belief, hence as having content in the same sense that thoughts and beliefs do, while still respecting the distinctive character of experiences in contrast to thoughts and beliefs.

## 7.1 PROBLEMS FOR THE CONTENT VIEW

In this section I want to look briefly at how Brewer understands the content view and at what he thinks is wrong with it. The "basic idea" of the content view, he says, is that "perceptual experience is to be characterized . . . by its representational content, roughly, by the way it *represents* things as being in the world around the perceiver" (2006: 165). The representational content here, he thinks, is modeled on that of "a person's *thought* about the world around him, as . . . expressed in his linguistic communication with others, and registered by their everyday attitude ascriptions to him" (2006: 166). Specifically, what Brewer calls the "initial model" for content, a model which serves as the basis for the content theorist's understanding of the content of perceptual experience, is a subject's thought, about an object *a* in his environment, that it is F. The content theorist arrives at a conception of the content of perceptual experience by modifying or qualifying

this initial model. On the version of the content view which Brewer takes as his target, the initial model is qualified in three respects, two of them corresponding to features of John McDowell's version of the content view, the third corresponding to a feature introduced in Brewer's own, earlier, articulation of the content view.[2] First, unlike that of a thought, the content of a perceptual experience is entertained passively: one finds oneself, in McDowell's terms, saddled with the content that *a* is F, as opposed to actively putting that content together oneself. Second, the singular component of the content is an object-dependent demonstrative sense: the content one is saddled with is, for example, that *this* man or *that* table is F, a content which depends on the existence of the object demonstrated. Third, the predicative component of the content is also demonstrative and, in Brewer's terms, world-dependent: it is articulated through a doubly demonstrative expression of the form "that is thus," for example "that (man) is thus (in facial expression)."

Now the point of these qualifications is to secure the contrast between experience and thought. To summarize, they specify that experiential content is unlike thought-content in being passively entertained rather than actively put together, and in being demonstrative and thus world-dependent both with regard to the object perceived and to the features which it is perceived to have. Is this specification sufficient to do justice to the distinctive character of experience in contrast to thought and belief? Brewer thinks not. The qualifications to the initial model, as he sees it, are "too little too late" (2006: 166), since that model retains two objectionable features which vitiate any attempt to modify it so as to capture the content of perceptual experience. The first feature is that content, on the initial model, allows of falsity, so that perceptions can be misleading. While this is often taken to be an advantage of the content view, Brewer argues that it is in fact a liability: truth and falsity should be understood as belonging only to one's *thought* about a particular object out there, not to one's perception of it.

The second, and, as it turns out, more fundamental of the two features, is that content on the initial model involves generality. On the initial model, "a particular object, *a*, is thought to be *a specific general way*, *F*, which such objects may be and which infinitely many qualitatively distinct objects are" (Brewer 2006: 173). This feature of the model, according to Brewer, requires the content view to construe perception as involving two kinds of selection or abstraction. Suppose, to take Brewer's example, that you are perceiving a

[2] McDowell (1994b); Brewer (1999).

particular red football. Your perception "must begin by making a selection" from among the various dimensions in which different such things can vary, for example selecting the color and shape of the ball from among a range of dimensions which also includes its weight, age, and cost. Second, and more importantly, it must select a determinate general way in which the ball is represented as being along each of the already selected dimensions, so that the ball is represented, say, as having a determinate shape and color. Such a determinate way the object is represented as being is general in that a range of objects might be that same way. Your perception thus classifies or categorizes the object by representing it as having something in common with other actual or possible objects.

According to Brewer, this is objectionable: "[h]owever automatic, or natural, such general classification may be, it still constitutes an unwarranted intrusion of conceptual thought about the world presented in perception into the [content] theorist's account of the most basic nature of perception itself" (2006: 174). Brewer suggests two reasons why this is so. One is that, in allowing perception to involve the exercise of abstract general thought, the content view deprives experience of any role in explaining the possibility of abstract thought. As Brewer puts it: "Perception itself constitutes the fundamental ground for the very possibility of any such abstract general thought about the physical world subjectively presented in it" (2006: 174). This objection is similar to John Campbell's argument for favoring what he calls the relational view of experience over the representational view. Campbell puts it like this: "we cannot view experience of objects as a way of grasping thoughts about objects. Experience of objects has to be something more primitive than the ability to think about objects, in terms of which the ability to think about objects can be explained" (2002: 122). The other reason is that the generality which the content view ascribes to perception prevents it from doing justice to the idea that experience consists in openness to the objects in the physical world. For Brewer, the content theorist's account of perceptual experience "trades direct openness to the elements of physical reality for some intellectual act of classification or categorization. As a result [the content view] loses all right to the idea that it is the actual physical objects before her which are subjectively presented in a person's perception, rather than any of the equally *truth-conducive* possible surrogates" (2006: 174). So the content view fails to accommodate the intuitively appealing idea that what constitutes the character of perception is the actual objects we perceive, as opposed to the general features which we represent those objects as having.

## 7.2 A PROBLEM FOR THE OBJECT VIEW

We have seen some considerations motivating the rejection of the content view. What can be said in its defense? It is often thought that a major motivation for the view is the need to account for perceptual illusions, since on the face of it these would seem to consist in experiences which represent the world as being otherwise than it actually is. Accordingly, Brewer and other philosophers attacking the view have devoted considerable attention to showing how we can make sense of illusion without invoking representational content.[3] But I am concerned with another source of motivation which has received less attention, at least in the context of this debate. This is the idea that a satisfactory account of perceptual experience should make sense of how perceptual experiences can stand in rational or justificatory relations to belief and judgment: how, in particular, our beliefs and judgments can be rationally intelligible in the light of our perceptions. This demand has been emphasized by John McDowell in his defense of the view that the content of perception is conceptual, and indeed by Brewer himself in earlier work defending the conceptualist position.[4] It has also been widely accepted by non-conceptualist critics of McDowell, who have typically chosen not to reject the demand, but rather to argue that it can be satisfied even on a non-conceptualist construal of experiential content. We would expect, then, that philosophers who reject the view that experience has content, and specifically who endorse a position like Brewer's "object view," would also want their position to satisfy this demand. But, on the face of it, the demand seems to require that we ascribe some kind of representational content to perceptual experience and, in particular, content with the kind of generality which Brewer finds objectionable. For on a certain natural line of thought, perception of an object cannot rationalize a belief, that is, make it rationally intelligible, unless it presents the object as being a certain way, that is, as having a certain general property or feature.

To get this line of thought into focus, let us begin with an example. Suppose that you are approaching your house, see a package on the front porch, and form the belief that the books you ordered have arrived. Does your perception make rationally intelligible your belief that the books have arrived? In other words, is it plausible, in the light of your having that

[3] See for example Travis (2004) and Brewer (2008). Brewer describes this as the "obvious challenge" facing the object view (2004: 70).

[4] See in particular Brewer (1999).

perception, to suppose that your belief was formed through a rational process? If we assume that in seeing the package you see it *as* a package, and more specifically see it *to be* a package,[5] then the answer, albeit with some qualifications, would appear to be yes. Assuming the right background of beliefs, for example that you ordered books, that they are due to arrive around now, that no other deliveries are on the way, and so on, then it does seem as though, when you see the package on the porch, it is rational for you to form the belief that your books have arrived. Your perception, as I shall put it for short, "rationalizes" your belief. But on the object view as Brewer understands it, your seeing the package as a package or taking it to be a package is not part of your perception proper, but rather a part of what he calls your "classificatory conceptual engagement" with the package. So if your perception rationalizes your belief it cannot be in virtue of your seeing the package as a package. What plays the justificatory role must be something more basic: that is, the mere fact of the package's being present to you perceptually, regardless of what you see it as being or take it to be.

On the face of it, though, your merely seeing the package does not seem to be the kind of thing which could rationalize the belief in question. For suppose that, on seeing the package, you take it to be, not a package, but rather a patch of sunlight or a pile of newspapers; or that you merely register its presence without taking it to be anything in particular at all. Even if your perception causes you to form the belief that your books have arrived, say by an unconscious process of association, there does not seem to be any ground for supposing that the process of belief formation counts as rational in the light of your perception. On the contrary, there seems to be something paradigmatically non-rational about the formation of your belief under those circumstances. Your belief seems on the face of it to be no more rational than if it had just popped into your head as you were approaching your house without your having seen the package at all.

Now Brewer does not say explicitly that perception on the object view stands in a rational relation to belief. But he does say that perception *grounds* belief: "the course of perceptual experience . . . provide[s] the subject with the grounds for her . . . beliefs about the world . . . not by serving up any fully formed content . . . but, rather, by presenting her directly with the actual constituents of the physical world themselves" (2006: 178). And it seems reasonable to understand the "grounding" here as rational grounding of the kind which figures in the demand that perceptions rationalize beliefs.

[5] While I argue that the latter is required if perception is to rationalize belief (Ginsborg 2006c), the distinction does not matter for present purposes.

The question raised by our example, then, is how the direct presentation of the package can, on its own, provide rational grounds for believing that your books have arrived. A natural answer is that while it cannot provide such grounds directly, it can provide grounds for more basic beliefs which ascribe properties to the package. Assuming that the package is not disguised in some way, and more specifically that it is visibly package-like – brown and rectangular, say – presentation of the package can rationally ground the belief that there is a package in front of you, or, at the very least, that there is something brown and rectangular in front of you: beliefs which can in turn rationally ground the belief that your books have arrived. In the example where, as I put it, you "see the package as a patch of sunlight," but still form the belief that your books have arrived, the failure of rationality does not indicate that your perception did not provide rational grounds for any beliefs at all. Your perception did provide rational grounds for the belief, say, that there was a package in front of you. Your failure of rationality lay in your having come to believe that your books had arrived without the mediation of that more basic belief.

Let us consider, then, how seeing a package in front of you might provide you with rational grounds for believing that there is a package in front of you. A possible answer is suggested by Brewer's claim that perception presents its objects in a way which makes "experientally accessible" the perceptible properties which they actually have. Brewer makes this claim in the context of a discussion of the Müller-Lyer illusion. Contrary to the position that would presumably be held by adherents of the content view, namely that perceptual experience presents the lines in the diagram as unequal, Brewer holds that it presents "the very lines out there, distributed in space as they actually are" (2004: 70). Their identity in length is thus "made experientially accessible to the normal subject" (2004: 70). The subject feels an inclination to judge that the lines are unequal in length, but the identity in length is still a "perceptible feature in part constitutive of normal subjects' experience of them" (2004: 71). This is indicated both by normal subjects' capacity to point to the endpoints of the lines, and the fact that if the misleading arrow-heads and -tails are gradually removed, what happens is not that the subject experiences the lines changing in length, but rather that the identity in length "becomes gradually more obvious" (2004: 71). Now the Müller-Lyer illusion is of course a special case, but the point can be generalized to normal perception. When you see the package, your experience might be said to present the very package out there, colored and shaped as it actually is. Its properties of being brown and rectangular, and, we might also suppose, its functional character of being a package, are

thus experientially accessible to you, allowing the experience which makes them accessible to rationalize the belief that there is a package, or at least something brown and rectangular, in front of you.

But what does it mean to say that the properties are "experientially accessible"? It cannot mean that the package is presented *as* brown and rectangular and *as* a package, so that its possession of these features is actually recognized within the subject's perceptual experience. For this is precisely what Brewer wants to deny. So it must mean something weaker, something which is, in particular, compatible with the claim that identity in length is "experientially accessible" in the Müller-Lyer illusion even though, intuitively, we represent the lines as unequal in length. But what this seems to amount to in the Müller-Lyer case is just that we are capable, under appropriate circumstances, of coming to represent the lines as having the property of identity in length: for example, we can come to represent them as identical in length if the misleading arrow-heads and -tails are removed. By the same token, the color, shape, and functional kind of the package are experientially accessible in this weaker sense if the package has them, and if we are capable of coming to see that it has them. And saying that they are experientially accessible in this sense does not seem to add anything to the claim that the package presented to us is in fact brown, rectangular, and a package, since these are features of a kind which we can, typically, come to see things as having.[6] So it does not help with the question of how having a particular brown, rectangular package perceptually presented to us can rationalize the belief that it has those properties of being brown, being rectangular, and being a package.

Another answer is suggested by Brewer's appeal to what he calls "visually relevant similarities" to account for perceptual illusion (2008: 171ff.). Even though the Müller-Lyer diagram makes experientially accessible the equality of the lines, its being visually similar to paradigmatic diagrams of unequal lines may either intelligibly lead us to take the lines to be unequal, or, in the case where we are not deceived, to "notice the intelligible applicability" of the concept of inequality to the lines (2008: 176). Again, the point can be generalized to ordinary perception: in the straightforward case where you take the package to be a package, the visual similarity of the package to paradigm packages makes intelligible your taking it to be a package. If intelligibility here means "rational intelligibility," then the

[6] While there may be a question about how to draw the line between properties which things can and cannot be perceived to have, I am assuming that we can perceive something to be brown, rectangular, and a package, but not, say, to have been seen by Gustave Flaubert on a Wednesday afternoon.

suggestion is that the similarity of the package to paradigm cases of packages might rationalize your forming the belief that it is a package. Here again, though, it is not clear how the similarity is supposed to figure in your perceptual experience. It cannot be that your perceptual experience proper represents the package *as* similar to paradigm cases of packages, since similarity to paradigm packages is no less a general property than color, shape, or packagehood themselves. But the only alternative would seem to be that the package is in fact similar to paradigm packages, where that similarity might or might not become salient to a given perceiver in a given context. And while that might make it possible to understand why, as a matter of psychological fact, your perception of the package leads you to believe that the package is a package, it does not help with the question of how your perception can make your belief "rationally intelligible" in the sense at issue.

## 7.3 Can states without content rationalize beliefs?

I have been arguing that we cannot make sense of a perception as rationalizing a belief unless we take the perception to represent its object *as* having some general feature or other, and hence as involving generality in just the way that Brewer finds objectionable. But I might be accused here of relying uncritically on the claim, defended most prominently by McDowell, that representational states must have conceptual content in order to rationalize beliefs. Philosophers on both sides of the debate about non-conceptual content have challenged this claim, arguing that perceptual states can rationalize beliefs even if their content is non-conceptual. That might seem to open a route to the more radical view that perceptual states can rationalize beliefs without having representational content at all.

Consider, for example, Peacocke's defense of the view that states with non-conceptual content can have rational bearing on beliefs. For McDowell, such rational bearing requires that we be able both to "scrutinize" the relations between experience and belief, and to "articulate" the reasons which experience supplies for belief. Peacocke accepts this point, but maintains, against McDowell, that these conditions can be satisfied even if the content of experience is non-conceptual. Regarding the demand for rational scrutiny, he points out that "a thinker can ask 'Is something's looking that way a reason for judging that it's square?'" even if the demonstrative expression "that way" refers to something non-conceptual (2001: 255). Relatedly, a thinker who comes to believe that something is square on the basis of how it looks, where its looking that way to her amounts to her

being in a state with non-conceptual content, can articulate her reasons: "'I believe it's square because it looks that way'" (2001: 256).[7]

Now Peacocke, of course, is assuming here that the reason-giving states have representational content. But it is not clear that this feature is essential to his challenge to McDowell.[8] For even if a perceptual state lacks content altogether, it still seems open to a subject to scrutinize its putative reason-giving relation to belief, asking for example "Is my perceiving this package a reason for judging that my books have arrived?" Similarly, saying "Because I see a package" seems like a perfectly good way of articulating at least one possible reason one might have for believing that one's books have arrived, namely that one sees a package. It is perhaps less natural to suppose that someone might regard her perception of the package as a reason for the more basic belief that there is a package in front of her. While it is not impossible to imagine someone's asking "Is my perceiving this package a reason for believing that there is a package here?" this would seem to be at most a limiting case, given that someone who can describe herself as perceiving a package is already committed to there being a package there.[9] Still, this lack of naturalness does not invalidate the suggestion that one's perceiving a package can be a reason for believing that there is a package in front of one, and it is perfectly plausible that someone might reply, when

[7] A similar line of argument is offered in Byrne (2005: 240–242) and Speaks (2005, 374–375). Both Byrne and Speaks endorse conceptualism but reject the view that it is required to accommodate the reason-giving character of perceptual experience.

[8] This is not to say outright that it is *not* essential. For Peacocke, part of what secures the rational relation between perceptions and beliefs is that there is an internal connection between the "way" in which a perceiver perceives a thing, and the "way" the perceiver believes the thing to be. It is, in part, because the thing looks *square* to the perceiver that its looking the way that it does is a reason for the perceiver to judge that it is square. So on the face of it, Peacocke's response to McDowell does seem to depend on the assumption that perceptual states have content. The same holds for Byrne's more articulated version of the response (2005: 240–242). Byrne finds it puzzling that a perceptual state without content, such as a "mere sensation," could provide a reason to believe that a thing is blue. But a perceptual state in which an object looks blue can provide such a reason because it has content, and more specifically, according to Byrne, content which strictly implies that the object is blue.

I have argued (2006a: 414) that Peacocke's assumption of an internal relation between a thing's looking a certain way to a perceiver and its being a certain way rests on an equivocation on the notion of a "way of being perceived," and for that reason should be rejected. But it does not seem to me that Peacocke needs to rely on this internal relation in order to argue that we can cite perceptual states as reasons for beliefs. Perhaps, as in Gupta (2006), one can cite one's perceptual experience E as a reason for the belief that *p* not because there is an internal connection between the content of E and *p*, but because one has knowledge of external facts about the world and one's sense-organs which imply that experiences of kind E typically occur only when *p* is in fact the case.

[9] I am here using "perceiving an F" in the same sense in which it is used on the object view, namely in a sense where you do not count as perceiving an F unless there is in fact an F present to you. For the distinction between this use of the expression, and a use which does not involve commitment to the presence of the F, see for example Price (1932: 22–24) and Harman (1990: 36–37). Travis objects to the second use (2004: 85).

asked why she believes that there is a package in the vicinity, "Because I see it." It might seem, then, that the line of thought in the previous section is mistaken: that there is nothing at all mysterious about the idea that one's perceiving a particular package, as construed on the object view, can be a reason for, or rationalize, beliefs concerning either the package or the implications of its presence.

In response to this objection, I want to look more closely at the notions of "reason" and "rationalize" that figure in the relevant debates. To fix ideas, let us consider a particular case of belief-acquisition: one where you look out of the window, see that the streets are wet, and come to believe that it rained. Assuming that this happens in the ordinary way, rather than, for example, because the sight of wet streets is associated with some trauma on account of which your present experience triggers delusional flashbacks replicating childhood memories of rain, this looks like a paradigm case of rational belief formation. In particular, it seems to involve a paradigm case of rational transition between beliefs: from the perceptual belief that the streets are wet (which is either a part of, or a consequence of, your seeing that they are wet) to the inferential belief that it rained. What, in this transition, should we describe as your "reason" for believing that it rained? According to T. M. Scanlon's characterization, a reason for an attitude is a "consideration counting in favour of [it]" (1998: 17). What counts in favor of your belief that it rained, in the scenario we are considering, is the fact, or perhaps the proposition, that the streets are wet.[10] When you are asked why you believe that it rained, and you reply "Because the streets are wet," it is this fact or proposition which you are citing as your reason. But there is also a philosophical usage in which your "reason" for believing that it rained is not the fact or proposition that the streets are wet, but rather the belief which has this fact or proposition as its content. This usage is exemplified in Donald Davidson's claim that that "nothing can count as a reason for holding a belief except another belief" (1983: 310). Davidson clearly does not mean to deny the possibility of your citing the fact of the wet streets in support of your belief that it rained. So he must be using "reason" in a different sense, one in which the "reason" is not the fact which you take to support the belief, but your belief that the supporting fact obtains.

I am inclined to think that the first sense of "reason" is more fundamental, and that talk of reasons in the second sense can be paraphrased away in terms of reasons in the first sense. But it is convenient to allow both uses to

[10] The distinction between these options does not matter for present purposes.

stand, and I will distinguish them in what follows by referring to reasons in the first and second senses respectively as "reasons$_1$" and "reasons$_2$."[11] Can we distinguish corresponding senses of the term "rationalize"? Here it seems to me that the most natural construal of that term is cognate with talk of reasons$_2$ rather than reasons$_1$. It is your belief that the streets are wet, not the fact that the streets are wet, which is most naturally described as making rationally intelligible, or "rationalizing," your belief that it rained. And this is the sense in which I have been using the term up to this point. But some philosophers might be inclined to an understanding of the term "rationalize" in which it is the fact of the streets being wet, in this situation, which has the "rationalizing" force (even if only in virtue of its being believed by you). So let us stipulate also two senses of "rationalize." In the case we described, your belief that it rained is rationalized$_2$ by your belief that the streets are wet, and rationalized$_1$ by the fact, or proposition, that the streets are wet.

With these distinctions in mind, let us return to the question of what is required for perceptions to stand in rational relations to belief. When Peacocke cites the possibility of a thinker's scrutinizing the putatively reason-giving relation between a perception and a belief by asking "Is something's looking that way a reason for judging that it's square?," or of her articulating her reason for a belief by saying "I believe it's square because it looks that way," which sense of "reason" does he have in mind? It seems to me that it must be the first sense. Her asking whether the thing's looking a certain way is a reason for judging that the thing is square, or citing its looking that way as a reason for her belief that it is square, is analogous to her asking whether the streets' being wet is a reason for believing that it rained, or saying that she believes that it rained on the grounds that the streets are wet. The fact that the thing looks a certain way to her, i.e. that she has a certain perceptual experience of it, is figuring in her reason-giving in the same way that the streets' being wet figures in our paradigm case of rationalization: namely as a consideration which potentially favors the belief in question, or a reason$_1$.[12] Putting the point another way, the perceptual experience plays the same rationalizing role as the wet streets: it is something whose presence, or the fact of whose presence, can serve as a reason$_1$. So if this is the model of rationalizing that the defender of the object view intends to exploit, then perceptual experiences as construed by the object view will

[11] I discuss this distinction in Ginsborg (2006c: section I).

[12] This is clear in Byrne's presentation of the non-conceptualist response to McDowell's articulation argument, since Byrne has the non-conceptualist identify reasons with propositions (2005: 238–241).

also provide reasons$_I$. That you have an actually existing package presented to you perceptually will be a consideration which can tell in favor of your forming certain beliefs, just as, in our paradigm example, the streets' being wet is a consideration which can tell in favor of your belief that it rained.[13]

However, I do not think that this way of construing the reason-giving relation does justice to the motivations underlying the requirement that perceptions be capable of rationalizing beliefs. To address the worries about coherentism which motivate McDowell's insistence on the requirement, perceptions must be capable of being reasons for beliefs in the same sense in which beliefs are (typically) reasons for beliefs: that is, they must rationalize$_2$ beliefs. If it were sufficient for them to be, or more precisely to figure in, reasons$_I$, then it would be possible to avoid the threat of coherentism without making any claims about perceptual experience at all. We could avert the coherentist threat, that our beliefs constitute a self-contained system without rational grounding from anything outside them, by pointing to any example in which a belief is rationalized$_I$ by a fact which does not itself involve someone's having a belief. For example, we could point out that, in our paradigm example of rational belief formation, the fact that the streets are wet serves as a reason$_I$ for the belief that it rained. But this would clearly be too easy. Or, to put it another way, it would miss the point of Davidson's "coherentist" principle. When Davidson says that nothing but a belief can be a reason for another belief, he does not want to rule out that the fact or proposition that the streets are wet can be a reason$_I$ for believing that it rained. Nor, for that matter, does he want to rule out that sensations or other psychological states that are not themselves beliefs can figure in rationalizations$_I$ of beliefs. We can cite the occurrence of our sensation of a green light flashing, just as we can cite any non-psychological fact which might indicate that there is a green light flashing, as a reason$_I$ for believing that there is a green light flashing. What Davidson wants to rule out is the possibility that anything other than a belief could play the

[13] Gupta's account (2006: 188–189) of how perceptual experiences entitle us to judgments only in combination with certain beliefs, about e.g. the functioning of our sense-organs and the prevailing perceptual conditions, suggests that this is how he understands the reason-giving role of experience. On a separate point: there is a question about whether, on this kind of model, we can say that the experience itself serves as a reason$_I$. Because I take reasons$_I$ to be facts or propositions, I think that strictly speaking it is not the experience but the fact of one's having the experience which is the reason$_I$. However, we often do not draw a distinction, in ordinary talk, between someone's experience and the fact of their having it. So the thought that (the fact of) your seeing *x*, or (the fact of) *x*'s looking F to you, is your reason$_I$ for believing *p*, could be expressed by saying that your experience of *x*, or your experience of *x* as F, is your reason$_I$ for believing *p*. (We might also say that the wet streets are your reason$_I$ for believing that it rained; here again, strictly speaking, your reason$_I$ is [the fact of] the streets' being wet.)

kind of rationalizing role that is paradigmatically played by beliefs – as I have put it, that anything other than a belief could rationalize$_2$ another belief. So if the claim that perceptions stand in rational relation to beliefs is to cut against Davidson, then it must be interpreted as the claim that they rationalize$_2$ beliefs, and not just that their occurrence can provide reasons$_1$ for beliefs.

Might the object theorist maintain, then, that your perception of the package is a reason$_2$ rather than a reason$_1$ for believing that there is a package in front of you? Here again, we might look, for a possible model, to discussions of non-conceptual content. When Richard Heck, for example, defends non-conceptualism by arguing against McDowell that perceptions need not have conceptual content in order to be reasons for belief, he makes clear that he takes them to be what I have called reasons$_2$ rather than reasons$_1$: "perceptions justify beliefs . . . and provide reasons for belief . . . pretty much the same way beliefs do" (Heck 2000: 509).[14] But Heck's model cannot simply be taken over by the object view, since, as he makes clear, it depends on the claim that "perceptions are . . . *like* beliefs insofar as to be in a perceptual state is to hold an . . . attitude towards a certain *content*" (2000: 509, my emphasis). Perception for Heck, like belief, "purport[s] *to represent how the world is*," and that is why perceptions can play the same reason-giving role. So the object theorist is in the more difficult position of having to show that perceptions can play the same kind of reason-giving role as beliefs even though, unlike beliefs, they do not have representational content.

While I cannot argue that this task is impossible, I can try to bring out the difficulty it faces by contrasting the present case with our paradigm case of rational belief formation. On the approach we are considering, the perception of the package has to play the same kind of role in rationalizing the belief that there is a package present that the belief that the streets are wet plays in rationalizing the belief that it rained. In our paradigm case it is plausible to suppose that the belief plays that rationalizing$_2$ role because it, so to speak, makes available to you a reason$_1$ for your belief, namely the

[14] See also Heck (2000: 517–519), where Heck considers a different kind of reply to McDowell. As I see it, this reply, which Heck rejects, is a version of the strategy which understands perceptual states as reasons$_1$. The "linguistic tangle" to which he refers (Heck 2000: 519 n.30) corresponds to the ambiguity I aim to resolve with talk of reasons$_1$ and reasons$_2$. (Heck tries to address it by distinguishing what one believes or how things appear to one, on the one hand, from the fact that one believes what one does, or that things appear that way to one, on the other, but I do not think that this goes far enough in capturing the contrast.) Alan Millar's view that there is a "quasi-inferential" link between experiences and beliefs (1991: 111–122) also seems to identify experiences with reasons$_2$. However, the account of perceptual justification which he gives (2007: 185–187) seems to me to cast the having of perceptual experiences in the reason$_1$ role instead.

consideration that the streets are wet. When you form your belief that it rained, you have "in view" the fact or proposition that the streets are wet, and you are in a position to cite that fact or proposition as a reason$_I$ for your belief that it rained. So if your perception of the package is to play the same kind of role with respect to your belief that there is a package in front of you, then it is plausible to suppose that it must also bring into view a consideration which supports the belief that there is a package in front of you. But what consideration could that be? If, contrary to the object view, your perception were taken to involve the representation of the package as having package-like features, say as being brown or rectangular, or indeed if your perception were taken to represent the package as, simply, a package, then there would at least be a case to be made for saying that it brings into view a consideration supporting the belief that there is a package in front of you. That the thing in front of you is brown or rectangular is a reason$_I$ for believing that there is a package in front of you, and the same might be said, as a limiting case, of the fact that it is a package. However, according to the object view, your experience does not bring into view any facts about the package, but only the package itself. And the package itself cannot count in favor, either of the belief that it is a package, or of the belief that there is a package in front of you. Not being a fact or proposition, it is simply not the right kind of thing to serve as a reason$_I$ for a belief.

A possible response here would be to broaden the "object view" by including items other than objects among the constituents of reality with which perceptual experience can present us. Perhaps, under favorable conditions, our experience can put us directly in touch not only with the package, but also with its brownness and rectangularity, or even, simply, with its package-like character. Such a broadening would preserve the spirit, if not the letter, of the object view as long as it were allowed that these features, no less than the objects which possess them, were genuine elements of the world.[15] But at the same time, the response goes, it would allow us to construe perception as rationalizing$_2$ belief insofar as these features might

[15] Brewer himself rejects the suggestion that the view could be broadened in this way, since he denies that general properties are "features of the physical world on a par with the objects themselves which have them" (2006: 180 n.9). We should also note that the sense in which a property is perceptually presented on this broadened view is different from the sense in which, on Brewer's view, a property is accessible in perception. Brewer does think that the properties of objects as well as the objects themselves are in some sense available in perception, and, relatedly, determine the subjective character of perception, but he takes these properties to include properties which, intuitively, we do not "see the object as having" (for example, equality in length in the case of the Müller-Lyer diagram). On the present proposal, though, the properties which perception presents to us would be just those properties which the content theorist would view as belonging to the representational content of perception, so that the brownness of the package would not be presented to us in the case where, say, we mistook the package for a patch of sunlight.

constitute reasons$_1$ for belief. In seeing the brownness of the package, say, we would be presented with a consideration favoring the belief that it was a package.

Now one difficulty with this response as it stands is that one's being presented with an item and with a feature which it has does not yet add up to one's being presented with the kind of connection between the item and the feature which would seem to be needed if the experience is to play this rationalizing$_2$ role. We can see this by thinking of cases in which you see an object and one of its features, but without seeing the feature as belonging to the object. If the package is behind a screen door, and you see its brown color as belonging to the door rather than the package, then there does not seem to be, in what you see – the package, and its brownness – any more of a reason$_1$ for believing the item to be a package than was provided by the package on its own. What seems to be needed, in order for your perception to rationalize$_2$ that belief, is that it present you not just with the package and the brownness, but with the package and the brownness in some kind of predicative or at least proto-predicative relation: something which might be expressed as *the package's being brown*.[16] A view of perception which apparently meets this need, while still remaining within the spirit of the object view, is defended by Mark Johnston, for whom perception presents us with states and events as well as objects and stuffs. We can perceive such things as the snubnosedness of Socrates, or the astringency of the calvados, where these are conceived of as states or conditions of Socrates or the calvados, and as being on a par with events such as a particular chiding of Socrates by Xanthippe (Johnston 2006: 280–281). These states and events, Johnston emphasizes, are not to be confused with facts or propositions. There is a "difference between such things as the snubnosedness of Socrates – a certain state or condition of Socrates – and the true proposition that Socrates is snubnosed" (2006: 281). For this reason, just as on the object view, perception lacks truth-evaluable representational content: "sensory awareness is 'presentational' not representational" (2006: 284). At the same time, though, the states and events which we perceive might be thought to serve as reasons$_1$ for belief. Perhaps the "brownness of the package" on Johnston's construal, that is the state or condition of the package constituted by *its being brown*, can count in favor of believing the package to be a package in a way that the package itself, or the co-presence of the package and its brownness, cannot. And in

[16] The difference between seeing a feature of *x*, and seeing *x* having the feature, is emphasized by Mark Textor (2009).

that case your perception of the package's brownness – of its state of being brown – can rationalize$_2$ the belief that it is a package. Alternatively, your perception can rationalize$_2$, as a limiting case, your believing the package to be brown, where that belief in turn can rationalize$_2$ your believing it to be a package.

While I cannot pursue this suggestion in depth here, it seems to me that the aspect of Johnston's account which preserves the spirit of the object view – that is, the distinction between states and conditions on the one hand, and facts or propositions on the other – undermines the thought that perception, on that account, can rationalize$_2$ belief. For whatever plausibility there is to the idea that perceiving Socrates's snubnosedness can rationalize$_2$ the belief that he is snubnosed, seems to rely on the assumption that, in perceiving Socrates's snubnosedness, one perceives *that* he is snubnosed, and hence is presented with a fact or proposition counting in favor (here, as a limiting case) of the belief that he is snubnosed.[17] The point can be brought out most readily in connection with events. It is possible to perceive the event of a chiding of Socrates by Xanthippe without realizing that Socrates is being chided by Xanthippe: one might at the time be capable of describing what one is hearing only as "a muffled voice coming from the next room," and find out only later, if at all, that one had heard Xanthippe chiding Socrates. If that is the way in which one hears Xanthippe chiding Socrates, then, even if one's perception causes one to form the belief that Xanthippe is chiding Socrates, the belief is not rationalized$_2$ by the perception. It is the fact or proposition that Xanthippe is chiding Socrates which can serve as a reason$_1$ for, or count in favor of, this or that belief, not the event of Xanthippe's chiding Socrates. If we follow Johnston in assuming that states and conditions are to be understood on the same model as events,[18] the same must hold true for them. Seeing the package's brownness, or the package's being brown, can rationalize$_2$ the belief that the package is brown, or a package, only if it involves seeing the package to be brown.[19] The state of the package's being brown, in

[17] Textor (2009) makes a related point.

[18] The question whether we should lies outside the scope of this chapter.

[19] For a case where it does not, consider an apparatus designed so that a light flashes just in case a brown object is placed within range of it. Someone who sees the flashing light might be said to see the state of the object's being brown whether or not she is aware of the purpose of the apparatus, and hence of the fact that the object is brown. While this might seem to strain the idea of "seeing an object's being brown" or "seeing the brownness of an object," I think that this is because we typically understand these expressions as implying that the object is seen *to be* brown, and hence that we grasp the fact or proposition that it is brown.

contrast to the proposition that it is brown, cannot itself serve as a reason$_1$, so the perception of that state cannot, simply as such, serve as a reason$_2$.[20]

The object theorist could at this point offer a quite different kind of response, namely that rationalizing$_2$ does not necessarily require the presentation of reasons$_1$. Perhaps it is a merely contingent feature of the way in which beliefs rationalize$_2$ that they present us with reasons$_1$, so that other psychological states, such as perceptions, can rationalize$_2$ beliefs in a way which does not involve this feature. The rationalizing$_2$ relation, so construed, would be a more primitive relation than the relation between reasons$_1$ and the beliefs they support, in the sense that we could understand what it is for one psychological state to rationalize$_2$ another without appealing to the notion of a fact or proposition's counting in favor of a belief. While this line of response again requires more discussion than I can offer here, it seems to me that if we accept it, we lose whatever handle we had on what it is for one psychological state to rationalize$_2$ another. Without further clarification of the rationalizing$_2$ relation, we are not in a position to understand why some beliefs are reasons$_2$ for the beliefs they cause whereas others are not (as in the case where one belief leads to another through free association). And, more importantly for the philosophy of perception, we are not in a position to understand why, among the myriad psychological and non-psychological causes of our beliefs, it is beliefs and perceptions which deserve to be singled out as capable of rationalizing$_2$ them.[21]

## 7.4 A CONTENT VIEW NOT MODELED ON BELIEF

The argument of the previous two sections suggests that, if we want to respect the demand that perceptions be able to rationalize beliefs, in the

[20] What if the view is broadened still further so that the elements of reality which can be presented in experience include facts or true propositions? In that case, your perception of the package could serve as a reason$_2$ for believing that what you see is a package by acquainting you immediately with the fact either that it is brown, or (as a limiting case) with the fact that it is a package. McDowell's (1994b) view can be read along these lines; see also Martin (2002: 399). However, at least if facts are understood as true propositions rather than as truthmakers, it is no longer clear that such a view qualifies as a version of the object view, as opposed to the content view. I have offered an independent argument against this kind of view (2006c: 303ff.); a related objection is raised in Chen (2006: 251–252).

[21] The object theorist might argue here that they need not be singled out: perhaps many different possible states can rationalize$_2$ beliefs. James Pryor suggests that a headache can justify you in believing you have a headache (2005: 192–193), and the context suggests that the kind of justification he has in mind is rationalization$_2$ as opposed to rationalization$_1$. But, while I cannot pursue the point here, I am inclined to think that the occurrence of a headache can at most be (a limiting case of) a reason$_1$ for believing that you have a headache.

sense of "rationalize" which is relevant to worries about avoiding coherentism, then we have to understand perception as involving not just presentation of an object, but presentation of the object as having some specific feature or features.[22] And this seems to require that we ascribe to it representational content with the same kind of generality which characterizes the content of belief and thought. But is it possible to do this while still respecting the intuitions which lead Brewer and others to reject the content view, specifically that perceptual experience is explanatorily prior to belief and thought, and that it acquaints us directly with actual objects? I want to suggest that it is. As we saw in section 7.1, Brewer takes the notion of representational content which figures in the content view to be modeled on that of "a person's *thought* about the world around him" (2006: 166) and, more specifically, on the "initial model" of a singular thought that an object *a* in the environment is F, qualified to yield a conception of perceptual content as a species of thought-content which is passively entertained and doubly demonstrative. The difficulties Brewer sees for the content view derive from elements of the "initial model" – most fundamentally, the involvement of generality – which it retains in its qualified form. But I want to suggest another conception of the representational content of perception, independent of the "initial model," which does not require us to construe the representational content of perception as a modified form of thought-content.[23] This conception, I will suggest, allows us to understand perceptual experiences as having the kind of content which allows them to rationalize beliefs, but without falling foul of the difficulties I have mentioned.

My suggested approach takes as its starting point Brewer's own view of perception, that is the object view, but then introduces two modifications which have as their joint upshot that while perception still presents objects to us, we represent them, in that perception, as having general features. The first preserves fully the spirit of the object view, in that, like the view which Brewer himself proposes, the view under this first modification denies representational content to perception. Where it differs from Brewer's view is allowing more of a role, in determining the character of perceptual experience, to features of the subject who is having that experience. For Brewer,

[22] In Ginsborg (2006c) I argue, against McDowell, for a more demanding condition, namely that the perceptual experience that *x* is F cannot rationalize$_2$ a belief unless it involves commitment to the claim that *x* is F. But the argument which I am presenting against the object view does not require acceptance of this stronger view.

[23] I offer a version of this conception, as part of a defense of conceptualism, in Ginsborg (2006b). A similar conception, articulated in a way which remains neutral on the question of conceptualism, is developed in Ginsborg (2006a).

the "basic idea" of the object view is the idea that "the core subjective character of perceptual experience is given simply by citing the physical object which is its mind-independent direct object" (2008: 171). This basic idea, as he indicates elsewhere, is qualified to some extent so as to take into account factors additional to the actual physical object perceived: we can characterize what it is like for the subject to have the experience "by citing the perceptible features of the specific mind-independent empirical things which are accessible to her in perception, *given her point of view on the world and the relevant perceptual conditions*" (2004: 69, my emphasis). But we might wonder whether the austerity of the basic idea could not be relaxed still further to take into account the subject's psychological constitution, and in particular her capacities (innate or acquired) for perceptual discrimination. For it is not unreasonable to suppose that the subjective nature of experience is determined not only by features of the environment external to the subject (the character of the object, the viewing conditions) and by the subject's location within that environment, but also by the subject's discriminative capacities.

It will be helpful both for the present point, and as a background to discussing the second modification, to have in mind some examples. Imagine two animals who have undergone different kinds of training, one designed to bring about responsiveness to color, the other designed to bring about responsiveness to shape. The first animal always produces a certain characteristic response when shown a red object as opposed to, say, a yellow one, but shows no sign of being able to respond differentially to spheres as opposed to cylinders. The second animal exhibits the reverse pattern of responses: its responses discriminate among objects of different shapes, but not among objects of different colors. So when the two animals produce their characteristic discriminative response to, say, a red ball, we can think of each of them as reacting to a different feature of the object. The first animal's response is prompted by the ball's redness; that of the second animal, by its spherical shape.[24] Now I think it is plausible to suppose, given the difference in their patterns of response, that these two animals, confronted with the same red ball, have subjectively different experiences. In acquiring their respective ways of sorting or discriminating objects of different colors and shapes, the animals have also come to

[24] To qualify: whether we say that the animal is responding to the ball's redness will depend on how finely it discriminates. For example, if it discriminates between red and yellow things, but not between red and orange ones, then it will be plausible to think of it as reacting not to the ball's redness, but to a feature which it has in virtue of being colored either red or orange.

experience those objects in corresponding ways. We might put this by saying that the first animal's experience registers the ball's color and not its shape, whereas the second animal's experience registers its shape but not its color.

For a related example in the case of humans, imagine two people who have different discriminative capacities with respect to the experience of music. One can discriminate among chords of different qualities (for example, major, minor, and diminished triads), the other among the timbres of different but related instruments (cornet, trumpet, saxophone). It is natural to think that each of them, listening to the same major chord played by a saxophone trio, will hear it differently. Because the first is, as we might put it, sensitive to the harmonic qualities of what she is hearing but not to the timbre, her experience will register its character of being a major triad. The experience of the second, conversely, will register the characteristic saxophone sound of the chord, but not its harmonic quality.

It is important to note that nothing I have said commits us to the idea that the experiences mentioned in the examples have representational content. We might indeed say that the animals in the first example see the red ball, respectively, *as* red and *as* spherical, which seems to imply that the animals' experiences involve the ascription to the ball of the corresponding general feature. But this, I think, is just for convenience. The first animal does not "see the ball as red" in a sense which bears any philosophical weight; while it has a characteristic way of perceiving the ball, there is no particular way it perceives the ball *to be*.[25] Rather, saying that it sees the ball "as red" is just a way of saying that it perceives the ball in a way which registers the ball's redness, or that the ball's redness has an influence on the phenomenal character of its experience. The situation is less clear in the second example, since there might seem to be more of a case here for saying that the experiences have genuine representational content. Typically a subject who can discriminate different kinds of chords or the sounds of discriminate musical instruments is also a subject who can perceive a chord as major, or as having the characteristic saxophone timbre, in the philosophically substantial sense which implies some kind of ascription of the corresponding general quality to the thing perceived. I will say more about this shortly, but for now I just want to note that this is not a necessary feature of the example I gave. We might for example notice that an infant – perhaps destined to become the next Sonny

[25] I discuss this distinction in Ginsborg (2000b: 358–360) and Ginsborg (2006a: 413–414).

Rollins – becomes exceptionally alert and attentive whenever she hears the sound of a tenor saxophone. It is reasonable to suppose that her experience registers the timbre of what she is hearing, and in this respect differs from the experience that she has listening to the same tune played, say, on a trombone or on an alto saxophone in the lower part of its range. But we can suppose this without supposing that she represents what she is hearing *as* having the sound of a tenor saxophone, or indeed as having any particular kind of sound at all.

For reasons related to the consideration I just mentioned, the modification I have described does not represent a significant departure from the object view as Brewer describes it. It simply specifies that the range of features which count as "accessible to [the subject] in perception," and hence determine the character of her perceptual experience, is limited not only by the subject's point of view on the object and the conditions of perception, but also by her capacities for perceptual discrimination. The second modification represents a more radical departure from Brewer's view since, as I will go on to suggest, it does commit us to the idea that perceptual experience has representational content. To introduce it, I want to propose a distinction between two very general ways in which a subject can carry out an activity of discrimination or sorting, ways which I will describe by saying that she can do it with or without the consciousness of normativity, and label, even more briefly, by speaking of "normative" and "non-normative" sorting or discrimination. Let us go back to the very simple case of the animal which, when prompted, produces a characteristic behavioral response to red things which (let us suppose) it does not produce to things of any other color. Unless it is a very intelligent animal, we are not likely to be tempted by the thought that, in producing that response to a particular object, it takes itself to be responding appropriately to the object in the light of its previous training. We are likely instead to say that it responds "blindly" to the redness of the object. Part of what this implies is that its response is not guided by its prior recognition of the object as red. Its response is not, for example, prompted by the reflection that, since this object is red and it has been rewarded for responding to red objects in the past, it is likely to be rewarded for responding to this object too. But it also implies something stronger, namely that the animal produces its response without any sense at all of the response being called for by, or appropriate to, its present situation. It is that absence of any sense of appropriateness, and not just the absence of a process of deliberation, which I want to characterize by saying that the animal discriminates red objects without the consciousness of normativity.

By contrast, much of human activity appears to involve a kind of discrimination or sorting which does involve the consciousness, in producing a sorting response to an object, that the response is appropriate to the object. The most pervasive cases involve the use of language, but looking at a simpler case makes it easier to compare this kind of normative discrimination with the non-normative discrimination just described. If you are given a bowl of red and black beads and asked, without further specification, to sort them into two different boxes, you will most likely put the red beads in one box and the black beads in the other. As far as the tangible result of your activity is concerned, we can equate it with that of, say, a pigeon trained to distinguish red things from black. You will produce one kind of response to the red beads, namely, putting them in box A, and another to the black beads, namely, putting them in box B. But what you do will be different from what the pigeon does, I want to suggest, in that, when you put each bead in its respective box, you will take yourself to be responding to the bead appropriately. Your putting a given red bead in the box with the other red beads will be carried out with the sense that this is where the bead belongs.

Now the examples I gave in discussing the first modification of Brewer's object view were of non-normative sorting; or at least, as in the case of the human subjects discriminating timbre and harmony, they left open whether the sorting was normative or non-normative. My point there was to suggest that the subject's capacities for that kind of sorting or discrimination could be seen as determining the subjective character of her experience. But I now want to suggest – and here I am proposing the second modification – that it makes a further difference to the subjective character of the experience whether the capacities in question are for normative or non-normative sorting. And specifically, I want to propose, a subject who is capable of normatively discriminating red things from things of other colors – that is, of discriminating them with the awareness that her discriminative response is appropriate – has experiences of red things which not only register their redness, but also represent them *as* red. The thought here is that what it takes for a subject to perceive something not only *in* a certain way, but also *as being* a certain way, is just for her to have the kind of perception of it that a subject typically has when she is capable of normatively discriminating the presence of the corresponding feature. You count as perceiving a given object as red, as opposed to merely perceiving the thing in such a way that the redness influences the phenomenal character of your experience, if you perceive it in the kind of way which is associated with your being able to sort it with other red things, and where your sorting involves the

consciousness that you are sorting appropriately. This thought is motivated in part by the intuition that taking a set of objects to share a certain general property is in the first instance a matter of taking them, in some sense, to "belong together." In order to see the bead as red you do not first need to be capable of abstract thought to the effect that this or that thing is red. It is sufficient, rather, that you have the kind of perception of it which leads you not just to put it in Box A, but to put it in Box A with the sense that that is where it ought to go. But the thought also relies on the idea that when you take things to "belong together" in the context of this kind of sorting activity, you are not engaged in an exercise of abstract thought which requires you to recognize that the things share a common property. Rather, your "taking the things to belong together" is an aspect of normative sorting regarded as a more primitive activity: an activity which does not presuppose the capacity to recognize the objects being sorted as having features which make your way of sorting appropriate.[26]

This line of thought, if acceptable, yields a modified version of Brewer's "object view" on which perception, while still presenting objects to us, also presents them to us as having general features, and thus has representational content. The subjective character of perceptual experience is given, as on Brewer's view, by citing the object perceived, along with its perceptible properties. But it is also determined by the subject's capacities for discriminating those properties. So the phenomenal character of the experience of a red ball will vary depending on whether the subject is a creature who is incapable of responding discriminatively to red things at all, a creature who responds discriminatively but "blindly" to red things, or finally, a creature who responds discriminatively to red things with the consciousness of her response as appropriate to her situation. In this last case, on my proposal, the experience has representational content: the subject represents the ball as red. So the upshot of the modifications I have suggested is that, in the typical case, perceptual experience, for human beings, has representational content. More specifically, it has content of a kind which makes it capable of rationalizing$_2$ beliefs. For the perceptual state in which you see the ball as red, on my view, is a state in which the ball's being red is made available to you to serve as a reason$_1$ for your subsequently formed beliefs. You might not be in a position to express linguistically the proposition that the ball is red, but insofar as you represent the ball, in your perception, as "belonging" with the objects which you have previously sorted in the same way, you are

[26] I argue for the possibility of this kind of activity in Ginsborg (2006b: 360–363), Ginsborg (2006a: 419–427), and Ginsborg (forthcoming: section II).

*eo ipso* representing it as having a general feature in common with those other objects. Like the linguistically expressible belief that the ball is red, or that the ball has the same color as the previously classified objects, your representational state is capable of rationalizing$_2$ any belief favored by the consideration that the ball has that general feature.

We can think of the view I have presented, then, as a version of the content view. As such, it avoids the difficulty I raised in sections 7.2 and 7.3 for views which deny representational content to perceptual experience, namely that they fail to allow for rational relations between perception and belief. But it differs in a crucial respect from the content view which Brewer takes as his target. As noted, Brewer assumes that the content view takes as its starting point an "initial model" of the content of a linguistically expressible thought. So the content view as Brewer understands it explicitly models perception on thinking: the notion of representational content which it employs is directly derived from the notion of the representational content of linguistically expressible thought. By contrast, the version of the content view which I have proposed is arrived at, so to speak, from the other direction. It begins with the pre-theoretically attractive notion of perception as direct acquaintance with objects, and then modifies it to introduce, first the idea that the same object can be perceived by different subjects in different ways depending on their capacities for perceptual discrimination, and then the idea that our perceptual discriminations can involve a sense of their own appropriateness to the object perceived. The notion of perceptual content which emerges on this view does not rely on the notion of the content of a linguistically expressible thought. Rather, it is constructed out of the notion of perceptual discrimination understood as involving a primitive awareness of normativity. The modifications made to the object view do indeed entail that perception resembles thought in having representational content, and specifically representational content with the generality characteristic of the content of thought. But the ascription of representational content to perception is motivated not by any prior commitment to a kinship between perception and thinking, but rather by considerations derived from reflection on the character of perception itself.

It follows, I think, that the view I am proposing is not vulnerable to the difficulties for the content view described in section 7.1.[27] First, the

[27] There remain further difficulties for the content view, in particular that developed in Travis (2004). Here I will say only that my own version of the content view avoids the difficulty raised by Travis by construing perceptual experience not as what he calls "allorepresentation" (its being represented

generality it ascribes to the content of perception does not undermine the explanatory priority of perceptual experience over belief and thought. For it understands what it is for a subject to perceive an object as having a general feature – as red, say – without appealing to the subject's capacity to entertain, in an abstract way, thoughts involving the content *red*. Relatedly, it does not construe the generality in the content of her experience as a result of her having performed an "intellectual act of selection or categorization" (Brewer 2006: 174) of the kind Brewer describes in connection with the example of the red football. The only element in the account which might qualify here as an "intellectual act" is the subject's consciousness, in making the discriminations to which her nature and training dispose her, that these discriminations are appropriate. But this consciousness does not presuppose an antecedent grasp of the corresponding concepts and, in particular, does not require her to be capable of representing things as having the corresponding features outside the perceptual context. The view I have offered thus leaves open the possibility of explaining her more sophisticated capacity to think of things, in the abstract, as possessing those features, in terms of her more primitive ability to perceive them as having those features when they are perceptually presented to her.

Second, the generality invoked here does not detract from the status of perceptual experience as "direct openness to the elements of physical reality" (Brewer 2006: 174). The view does indeed insist that there is more to the character of an experience than that of the object presented to us: it holds that there are different ways of being open to the same elements of physical reality, and that these involve our ascribing different general features to those elements. If we assume that the only way to understand this kind of generality is by way of a conception of thought drawn from outside the context of perception, then there might indeed be something incoherent about this position: perhaps the idea that our perception involves judging the object to have this or that general property does indeed, on that assumption, undermine the idea that what is presented to us in the perception is the particular object itself, rather than any of a range of surrogates to which that property can also be ascribed. But I have tried to suggest in this last section of the chapter that there is an alternative way of understanding the ascription to objects, within perception, of general properties. And I believe that this enables us to preserve the essential insight of the object

to one that things are thus and so) but rather as "autorepresentation" (one's taking things to be thus and so) (see Travis 2004: 61–62). This opens my view to other difficulties, notably that of accounting for cases of known illusion (for example, seeing the Müller-Lyer lines as unequal in length while believing them to be equal), but space considerations preclude further discussion.

view, while still ascribing to experiences the kind of content which allows them to stand in rational relation to beliefs.[28]

[28] Predecessors of this chapter were presented at the Eastern Division Meeting of the American Philosophical Association in 2006, and at the Humboldt University in Berlin in 2007. I am grateful for comments and discussion, on those and subsequent occasions, to Bill Brewer, John Campbell, Tim Crane, Dina Emundts, James Genone, John McDowell, Mike Martin, and Alan Millar.

CHAPTER 8

# *Immediate warrant, epistemic responsibility, and Moorean dogmatism*

*Adam Leite*

For Jim Pryor, with gratitude, in order to find out exactly where we disagree.

According to the "Moorean dogmatist" response to external world skepticism (notably, Pryor 2000), our sensory experience provides us with *prima facie* (defeasible) immediate justification, warrant, or reason to believe certain propositions about the world. In addition to this core claim, many Moorean dogmatists also hold that there is an acceptable form of reasoning arising from this warrant: an ideal rational agent who considers whether she is being deceived by an evil demon could (1) start from a position which presupposes no beliefs at all about the world, (2) consciously take her current experience as a ground for believing that she has hands and so come to believe that she has hands on that basis, (3) consciously reason from that belief to the conclusion that she is not a disembodied spirit being deceived by an evil demon, and (4) thereby form the latter belief for the first time and in a fully epistemically satisfactory way. The Moorean dogmatist thus hopes to satisfy the perennial epistemological aspiration, post-Descartes, of explaining how an ideally rational agent could arrive at fully epistemically satisfactory beliefs about the world – including the belief that she is not being deceived by an evil demon – by reasoning from an initial position that takes for granted no initial claims about the world.

Many people, both epistemologists and non-philosophers, find the Moorean reasoning unattractive: it seems that the resulting beliefs, formed in that way, would not be fully epistemically satisfactory. It is not easy to locate the source of this dissatisfaction, however. In this chapter I will argue that it arises from requirements of responsibility that apply to rational agents in the course of conscious reasoning or deliberation. The rough idea will be this: even if sensory experiences provide the kind of reasons or warrants that the dogmatist claims, one cannot acceptably reason from them to conclusions about the world unless one already has epistemically

satisfactory beliefs to the effect that one is not, for example, a disembodied spirit being deceived by an evil demon. In working out this rough idea, I will take inspiration from a principle known as "the principle of inferential justification" (Fumerton 1995):

> In order to arrive at a justified belief that *q* on the basis of an inference from some premise *p*, one must have a justified belief to the effect that *p* supports *q*.

Elsewhere I have related such principles to the demands appropriately placed upon mature rational agents and have defended them against charges of vicious regress and other forms of unsatisfiability (Leite 2008).[1] In this chapter, I will defend a related principle. My first point will be that this principle is completely compatible with the core dogmatist claim about sensory experience. However, it will preclude – and explain our distaste for – the forms of reasoning endorsed by many dogmatists. A surprising result is that dogmatism alone isn't enough to evade standard arguments for skepticism about the external world.

## 8.1 Immediate warrant and the demands of epistemic responsibility

### *8.1.1 Epistemic warrant*

Rather than talking indifferently about a person's having warrant, justification, or reason to believe a certain proposition, I am going to talk in terms of *warrant*. The aim of this section is to explain what I mean by the term.

Roughly, a warrant to believe Q is a state or condition that *counts in favor* of believing Q. It is the sort of thing that could provide a correct answer to the question, "Why is Q an epistemically appropriate or acceptable proposition to believe (regardless of whether you actually believe it)?" I am going to assume not only that there are states or conditions that are warrants for belief, but also that we can sensibly ask what it is in virtue of which these states or conditions are warrants.

My main reason for talking this way, rather than using the more colloquial term "reasons" in its normative sense, is theoretical inclusiveness: there are states or conditions that would, on some views, count as warranting a belief even though it would not be natural to say that the person has reason, or a reason, to hold that belief. For instance, process reliabilist and proper functionalist views seem to be trying to characterize conditions

[1] I offer a motivation roughly along these lines in Leite (2008: sections 1 and 2).

under which a person has warrant to hold a certain belief. But the conditions they describe could be satisfied even though it doesn't seem entirely natural to describe the person as having reason, or having a reason, to hold the belief.

An adequate framework for thinking about warrant should allow us to draw two distinctions.

The first is a distinction between *immediate* and *mediate* warrants. On one prominent characterization, an *immediate warrant* is a state or condition whose status as a warrant is not constituted by one's having warrant to believe any other propositions, in the following sense: an adequate account of *what makes it the warrant that it is for believing that proposition* would not need to say anything about one's having warrant to believe any supporting propositions. This requirement's satisfaction is compatible with its being the case that things appealed to in that account *are* warrants to believe other propositions. All that matters is that their being warrants to believe other propositions isn't part of the account of what makes this state or condition the warrant that it is for *this* proposition (Pryor 2005). Epistemologists have debated whether any warrants are immediate in this sense. A number of candidates have been put forward, including visual and other experiential states as well as such things as pains. For the purposes of this discussion, I will grant that there are immediate warrants. My question is what exactly would follow if there were.

An experiential state would not provide immediate warrant to believe every proposition whose truth it supports. Some limitation is needed. A natural limitation is to say that an experiential state cannot provide you with immediate warrant to believe anything other than what it "directly tells you" about the world, which is often interpreted in terms of the content of the experiential state itself (e.g. Pryor 2000). For the purposes of this discussion, I will simply grant this conception of experience as well as this limitation on immediate warrant. I will assume that your visual experience might well give you immediate warrant to believe that there is a hand in front of you. But it will not give you immediate warrant to believe that your experience is reliable, that you are not a brain in a vat, or that you are not a disembodied spirit being deceived by an evil demon.

An adequate framework for thinking about warrant should enable us to draw a second distinction as well. This is the distinction between *prima facie* warrant and *all-in* warrant. One can have warrant to believe one proposition and also have warrant to believe an incompatible proposition. What each, taken singly, warrants one in believing is what one has *prima facie* warrant to believe. What one has *all-in* warrant to believe is the

result when these warrants are weighed against each other or the conflict is otherwise adjudicated. Many warrants are likewise *defeasible*: if certain conditions obtain, then the warrant provided by the state or condition will be defeated. (It is of course possible for the defeating conditions to be defeated in turn.) When defeating conditions obtain, it would still be appropriate to characterize the state or condition as a warrant, but it is only a *prima facie* warrant. Given that defeating conditions obtain, it does not provide *all-in* warrant.

There is debate about what sorts of states or conditions can constitute warrants. One issue concerns the relation of warranting states or conditions to the subject's awareness. Some theories hold that only states or conditions of which the subject is aware can constitute warrants, while others allow that states or conditions of which the subject is not aware can constitute warrants. (Process reliabilist views would be a case in point.) Another issue here concerns the "location" of the warranting states or conditions. Must they be internal to the subject, in some relevant sense of "internal"?

I will not be fully ecumenical on these issues. I assume, first, that warranting states or conditions are the sorts of things of which subjects can become aware, though not necessarily through introspection alone; second, that warrants are not wholly "internal," at least in the following regard – there can be real-world necessary conditions, entirely independent of the subject's psychology or intrinsic physical states, for a state or condition's being a warrant. (At present I see no reason not to accept that mind-independent states or conditions can themselves constitute warrants, but nothing here turns on this.)

I make one further crucial assumption. Warrants are things of a type that is suited to play the role of normative epistemic reasons in an agent's deliberations about what to believe. That is to say, if an agent has a warrant to believe Q and is deliberating about whether to believe Q, that warrant will be the sort of thing that could play the role marked out by the "because . . ." clause in a correct judgment of the form, "I should/may believe that Q, because . . ." This is not to say that whenever one has a warrant to believe some proposition, one could use that warrant in this way; other factors – relevant ignorance, say, or considerations about how one may acceptably proceed in the course of reasoning – may prevent one from doing so on a particular occasion. (This requirement, as I mean to understand it, does not entail that the warrant must itself be a proposition or have propositional content. I will assume that the warranting state or condition is what is picked out by the "because . . ." clause, not the proposition by which we pick it out.)

An epistemic reason in favor of believing a particular proposition supports, tells in favor of, or indicates the truth of that proposition. If warrants are suited to play the role of reasons in an agent's deliberations about what to believe, then warrants support, tell in favor of, or indicate the truth of the propositions that they warrant believing. I will summarize this thought by saying that warrants must be "truth-conducive." There are a variety of ways of interpreting this requirement. Here, I will assume that it involves a demand for *reliability*: roughly, the world has to be such that relying upon the warranting state or condition will, all else equal, lead one to a true belief. This gloss may in fact be stronger than necessary, but it points in the right direction; in demanding reliability, I mean to be invoking our ordinary language term in its colloquial sense. I accordingly want to prescind from the details of particular theories of reliability (statistical, counterfactual, etc.) and highlight the underlying thought that such theories aim to develop.

Certain brands of "internalism" (including Pryor's, I think) will reject this interpretation of the requirement that warrant be truth-conducive. However, it is strongly suggested by our actual epistemic practice. Suppose that someone is brought up to predict the outcomes of battles by reading tea leaves, a method endorsed by everyone in his community. Neither he nor anyone in his community is in a position to understand the considerations showing that there is no reliable connection between the arrangement of leaves in tea cups and the outcomes of battles. This person performs blamelessly if he infers from considerations about tea leaves that a battle will turn out a certain way; he has done everything that can reasonably be demanded of him in order to form a true belief. But at the same time, we feel that there is a shortcoming here. We might say, "His training and circumstances are unfortunate. He really shouldn't believe on that basis that the battle will turn out a certain way; no one should. Regardless of what he thinks, considerations about tea leaves don't actually provide any reason to believe anything at all about the outcomes of battles." When we make judgments like this, what seems to be motivating us is the thought that there is not in fact the right sort of connection between arrangements of tea leaves and the outcomes of battles: the one is not a reliable indicator of the other, and as a result the belief about the outcome of battle is not warranted.

So understood, warrants will often be relative to contingent real-world conditions. This can most easily be seen in cases in which the warrant consists in the possession of evidence. In many cases, evidential support depends upon general facts about how the world works. For instance, we

might grant that as the world works, how matters look to you can be good evidence about how they are. But – to take up the sort of example that is the Moorean dogmatist's concern – if you were just a disembodied spirit in a non-material world being given deceptive visual experiences by an evil demon, then there would not be any sort of reliable link between things looking a certain way to you and their actually being that way. In that case, your visual experience wouldn't in fact be evidence at all about how things are, regardless of what you quite reasonably believe about the matter. Since this evidential support relation thus depends upon your visual system's functioning in a certain way in relation to the world, any warrants you have for beliefs about the world in virtue of possessing visual evidence would likewise do so: they would consequently depend upon the obtaining of certain facts about the world. And since a warrant is a state or condition whose obtaining can serve as a reason or evidence in favor of the warranted proposition, many warrants will be such as to depend upon contingent real-world conditions even if they do not consist in or require the possession of evidence.

I will use the term "enabling condition" for a condition without which – as things are – the right sort of link would not hold between the obtaining of the warranting state or condition and the truth of the warranted proposition. And I will use the term "disenabling condition" for a condition which, if it obtained, would prevent the state or condition from so much as *prima facie* or defeasibly warranting belief in the proposition in question. Here are some examples. My recognition that I just took ibuprofen defeasibly warrants me in believing that my headache will soon lift. But this would not be so if human physiology were such that ibuprofen didn't alleviate pain in anyone. This possibility is thus a disenabling condition for this warrant. Likewise, if visual experience provides warrant for beliefs about the world, then one's being a disembodied spirit deceived by an evil demon will be a disenabling condition relative to this warrant.

As I have characterized a disenabling condition, it is not the same thing as a defeater. A defeater is a condition which is such as to defeat the *prima facie* or defeasible warrant provided by a particular warranting state or condition. It is not such as to prevent that state or condition from being so much as a *prima facie* or defeasible warrant in the first place. Many theorists hold that a subject's other mental states (beliefs, doubts, and the like) can constitute defeaters for the support the subject's experience provides for holding certain beliefs. Taken by itself, these theorists will say, the experience does support holding those beliefs, but all things considered – given the subject's other mental states – the subject is not warranted in

holding those beliefs. A disenabling condition, by contrast, would prevent the state or condition from even defeasibly supporting holding the belief in question.[2]

One last point about the notion of epistemic warrant. A person might have a warrant to believe a particular proposition, believe that proposition, and yet not hold that belief in a reasonable, epistemically appropriate, or epistemically satisfactory way. For instance, the person might base the belief on some other, bad reason, or the person might not hold the belief on the basis of any epistemic reason at all, but instead out of wishful thinking. This is just one sort of example; a variety of things can go wrong in such a way that the person's belief-state is not fully epistemically satisfactory even though the person has a warrant for the believed proposition and so is warranted in believing it. To give a label to the status that is lacking in all such cases, I will say that a person is *doxastically justified* just when the person not only has a warrant for the belief, but also holds the belief in a way that is beyond epistemic criticism in all respects other than truth. A variety of conditions will have to be met in order for a person to be doxastically justified regarding a particular belief on a particular occasion, and I will not attempt to offer a full account of them here. What I want to emphasize is that we have to distinguish these further conditions from the conditions that must be met in order for the person to have warrant. It would be a mistake to think that the factors involved in these further conditions must be factors that constitute the person's having warrant to hold the belief.[3]

### *8.1.2 Some "responsibilist" proposals*

I turn now to requirements on epistemically satisfactory belief that arise from considerations about epistemic responsibility.

[2] It is possible to allow that real-world, external states or conditions can constitute defeaters. In fact, I myself incline toward such a position. Even on such a view, however, we will want to handle defeaters and disenabling conditions separately, since only the latter would prevent the putatively warranting state or condition from even defeasibly supporting holding the belief in question. Admittedly, it may not be a straightforward matter to exhaustively sort possibilities into defeaters or disenablers relative to a given warrant. There may be puzzling cases, and the sorting may depend upon our characterization of the warranting state or of the link between it and the truth of the proposition in question. But none of these difficulties of detail undercuts the basic distinction.

[3] I use the term "doxastic justification" here partly in order to highlight that what I have in mind is a status attaching to a person in relation to an actual belief in virtue of how that belief is held. I don't mean to import a host of presuppositions from any particular epistemological theory or debate; this is just a convenient label that picks up on some features of ordinary and philosophical usage.

A person can have a warrant for a particular belief and yet not recognize the warranting condition or state as such. Very young children are generally in this situation, and mature adults can be in it as well. Mature adults differ from very young children, however, in that they are often in a position to recognize a particular state or condition as telling in favor of the truth of a particular belief. Mature adults – and especially those who have been trained to think critically – also differ from children in another important respect: they are capable of directly forming and modifying their beliefs in response to their evaluations of reasons in the course of explicit, conscious deliberation.[4]

Different abilities warrant different demands (Wright 2001). The above differences consequently open up the possibility of modes of epistemic evaluation that go beyond the question of whether a person has warrant to hold a particular belief. In particular, they open up the possibility of asking whether a person proceeded *responsibly* in the deliberative formation of a particular belief and whether a particular belief – regardless of whether or not formed through explicit deliberation – is held in an *epistemically responsible way*. For if someone is capable of directly forming or modifying a mental state or attitude through deliberation, then it is something for which she is properly held accountable and about which it can be asked whether she has conducted – and is conducting – herself well or poorly.

One central issue for such evaluations is how matters look from the person's perspective, particularly regarding reasons for and against the belief. It is arguable – though disputed – that a mature adult cannot hold a belief responsibly unless she regards something as a good reason for it. Plausibility precludes requiring the person to have explicitly highlighted something as a good reason for holding the belief; only certain dispositions of thought and response can reasonably be demanded. Even so, it might be thought that such a requirement is too strong. But something weaker is hard to deny:

> When in the course of explicit, conscious deliberation or reasoning one bases a belief that *p* upon a particular warranting state or condition W, that belief will not be formed or held responsibly unless one takes W to support (defeasibly tell in favor of) the truth of *p*.

Again, it is surely too strong to require that the person have an explicit, occurrently conscious belief that W supports the truth of *p*. At most what is required is an "implicit" or "dispositional" belief, which – like your

[4] This is not to endorse the voluntaristic claim that we can form beliefs at will.

current belief that the sun is more than 42,798 miles from the earth – may be manifested in nothing more than certain dispositions of thought and response (such as a willingness to assent to it if asked, to treat the claim in question as a premise for reasoning, etc.).

Here is an argument for this requirement. To explicitly base a belief upon a particular state or condition is to link one's attitude to that state or condition in such a way that the acceptability of one's belief-state stands or falls with the adequacy of that basis. It is in effect to say, "It is upon *this* that I stake my claim." Understanding the basing relation in this way, it is hard to see how someone could responsibly base a belief upon a particular state or condition in the course of conscious deliberation or reasoning without taking that state or condition to tell in favor of the truth of the believed proposition. For one thing, to base the belief upon that state or condition is to treat that state or condition as an epistemic reason, and that requires, as any mature believer recognizes, that the state or condition tell in favor of the truth of the proposition. For another thing, to base a belief upon a particular state or condition is precisely to treat it as something other than a hunch or guess, but rather as something *supported*. And it is irresponsible to treat a belief in this way without taking something to tell in favor of the truth of the proposition in question. The basic idea I want to endorse here is consequently this: if someone consciously and explicitly moves in thought from a particular warrant to a particular belief, that belief will be little more than a guess from that person's point of view unless the person takes the warrant to tell in favor of the truth of the believed proposition. And if the belief is little more than a guess from that person's point of view, then it is not responsibly held. We quite appropriately expect more of mature believers.

This requirement engenders further requirements. Consider some possible *disenabling condition* $\Pi$ whose obtaining would prevent the state or condition W from defeasibly telling in favor of the truth of $p$. Suppose that you base your belief that $p$ upon W. As I've just argued, this requires you to believe that W tells (at least defeasibly) in favor of the truth of $p$. And suppose that you recognize that $\Pi$'s obtaining would prevent W from even defeasibly telling in favor of the truth of $p$. Then, you are rationally required to believe also that $\Pi$ does not obtain, at least if you consider the question. For given that you recognize the incompatibility between $\Pi$'s obtaining and W's defeasibly telling in favor of the truth of $p$, requirements of consistency preclude you from endorsing both the claim that $\Pi$ obtains and that W tells in favor of the truth of $p$, and they also preclude you from endorsing the claim that W tells in favor of the truth of $p$ while suspending

judgment or forming no opinion at all about whether Π obtains. So if you consider the question at all, you are rationally committed to endorsing the claim that Π does not obtain. If you failed to do so under such circumstances, your belief that *p* would be rendered irresponsible: either from your own point of view it would be little more than a guess, or it would be held in a way that involves you in a manifest inconsistency. Consequently, the following additional requirement appears quite plausible:

When in the course of explicit, conscious deliberation or reasoning one bases a particular belief upon a particular warranting state or condition, that belief will not be formed or held responsibly unless for each possible condition that one considers and regards as a disenabling condition for the warrant in question, one believes that it does not obtain.

As before, the requisite belief may be implicit or dispositional, manifested in such things as a willingness to treat the claim in question as a premise for reasoning, to assent to it if asked, to dismiss the relevant possibility as soon as one considers it, etc.[5]

Three points should be noted. First, this requirement is limited to cases in which one forms a belief through explicit, conscious reasoning. Second, it is limited to possibilities that one explicitly considers. Third, it concerns only disenabling conditions, not defeaters. It is a vexed matter whether basing a belief upon defeasible evidence commits one to the non-obtaining of every defeating condition that one considers (or perhaps even to the non-obtaining of all possible defeating conditions). But this requirement only concerns certain possibilities relevant to the existence of the defeasible evidential warrant in the first place.[6]

This requirement is supported by simple examples. Here's one. Suppose that on the basis of the consideration that you just took ibuprofen, you believe that your headache will soon lift. You recognize that if ibuprofen

[5] One can regard something as a disenabling condition for a warrant without having the ability to use that precise phrase. It would be enough if one took it (roughly) that if the condition obtained, then the state or condition in question wouldn't support believing the proposition in question.

The considerations adduced in support of the requirement defended in the main text may also support a requirement to the effect that if one bases a particular belief upon a particular warranting state or condition in the course of explicit, conscious deliberation or reasoning, one is thereby committed to taking no possible disenabling conditions to obtain in relation to that warrant. However, the requirement stressed in the main text is what carries weight in the criticism of Moorean dogmatism that will be developed here, and that is why I focus on it instead.

[6] Some epistemologists hold that only states of the subject defeat warrants. They might say that responsible belief doesn't require beliefs about the absence of defeaters, but only the absence of the relevant mental states: one needn't also believe that they are absent. Even so, there is room for a requirement of the sort I've defended so long as one accepts enough of a reliability requirement on warrant to generate the very possibility of disenablers in the first place.

did not at all interact with human physiology in such a way as to remove pain, then the fact that you just took ibuprofen would not tell in favor of the belief that your headache will soon lift. You are asked: "What of the possibility that human physiology is such that ibuprofen does not remove pain?" If you sincerely say, "Oh, I don't have any opinion at all about whether that's so or not, but still: my headache will soon lift because I just took ibuprofen," then – assuming there's no funny business going on – you are criticizable for holding your belief about your headache in the way that you do.

This requirement demands only that one *have* certain beliefs in order to responsibly hold certain other beliefs. However, it seems clear that if these additional beliefs are not responsibly held, then the target beliefs will not be responsibly held either. For instance, suppose that I explicitly base a belief that it is 2:15 p.m. upon the fact that my watch says so. I recognize that if my watch isn't reliable, then the fact that it reads "2:15" wouldn't defeasibly support the conclusion that it is 2:15. Fortunately, I believe that my watch is reliable. Suppose, however, that this latter belief flies in the face of lots and lots of evidence in my possession. Then my belief that my watch is reliable is clearly open to epistemic criticism: I haven't proceeded as I should have in relation to it. The same point will apply, all else equal, to my belief about the time, and it will apply because of my failing in relation to my belief about my watch's reliability.

In fact, an even stronger requirement is quite plausible. Suppose that I have no actual warrant for my belief that my watch is reliable. In that case, this belief will not be fully epistemically satisfactory even if it is responsibly held. In the terminology I used previously, it will not be *doxastically justified*. But under these conditions, my belief that the time is 2:15, based as it is upon what my watch says, will also be open to epistemic criticism; it too will not be fully epistemically satisfactory. For instance, we might say, "Since you have no good reason to believe that your watch is reliable, you shouldn't believe that the time is 2:15 merely on the basis of what your watch says." I am consequently inclined to accept the following requirement:

> When in the course of explicit, conscious deliberation or reasoning one bases a particular belief upon a particular warranting state or condition, that belief will not be *doxastically justified* unless for each possible condition that one regards as a disenabling condition for that warrant and considers at all, one believes that it does not obtain *and* this latter belief is *doxastically justified as well*.

This requirement leaves a lot open, pending a detailed account of doxastic justification. And it may be a little too strong, depending on what doxastic

justification turns out to require. But this much is quite plausible: in the specified conditions, the person must have the relevant belief, it must be held responsibly, and it must be warranted. And even this formulation leaves a lot open. For instance, for all I've said here the requisite belief could be fully epistemically satisfactory even if the person does not recognize anything as a reason for holding it. It might even enjoy a kind of default status, so long as there is an appropriate connection with truth in the ground of that status.

This requirement makes no claim about the nature or structure of the warrant that one must have for the requisite belief. It does not demand that one have an antecedent or independent warrant. So it is not an epistemic priority requirement. It states that if one has a certain sort of doxastically justified belief – explicitly based on certain grounds under certain sorts of conditions – then one also has to have certain other doxastically justified beliefs. It is a further question how the warrants for the latter beliefs might acceptably be related to the warrants for the former. Likewise, this requirement would not generate problematic regresses, since it does not involve higher-order requirements or require that every doxastically justified belief be arrived at through explicit, conscious deliberation or reasoning.[7]

There is no incompatibility between this requirement and the existence of immediate warrants. So one could hold that while there are immediate experiential warrants, no belief deliberatively based on such a warrant can be doxastically justified unless one also has certain other doxastically justified beliefs. On this view, some beliefs would be immediately warranted, and based upon an immediate warrant, but their status as doxastically justified would constitutively depend upon one's possession of some other doxastically justified beliefs. They would accordingly *not* be immediately justified. Immediate warrant and immediate doxastic justification are thus not merely notionally separable for a view along these lines; they are extensionally inequivalent.

This combination of views would be attractive in many ways. It would accommodate the idea that experiential states' warranting power is not constituted by one's having any warrant to believe anything else. It would allow the reliabilist idea that epistemically satisfactory beliefs must be appropriately related to truth. And it would allow the "internalist" idea that epistemic evaluation of mature adults also involves considerations about epistemic responsibility and how matters look from the agent's point of view. The combined view would thus show a way to connect our interest

[7] See Leite (2008 and 2005) for relevant further discussion.

in epistemic warrant with our interest in how things look from the first-person standpoint of an epistemic deliberator, and it would explain how our interest in the first-person deliberative stance relates to our interest in the acquisition of true beliefs.[8]

## 8.2 "MOOREAN DOGMATISM"

### *8.2.1 "Moorean reasoning" and immediate warrant*

The external world skeptic holds that no one is doxastically justified in believing anything about the external world by means of the senses. The Moorean dogmatist diagnoses this skeptical claim as arising from a principle about *experiential warrant*:

> One's possession of any experiential warrant to believe anything about the external world constitutively depends on one's having an independent warrant for believing such things as that one is not dreaming, not a brain in a vat, and not a disembodied spirit being given misleading hallucinatory experiences by an evil demon.

If this principle is correct, then we have no immediate experiential warrants for beliefs about the world outside of our own minds. But the Moorean dogmatist holds that there are immediate warrants for beliefs about the world – for instance, that the visual experience as of a hand gives one immediate warrant to believe that there is a hand. The Moorean dogmatist thus attempts to avoid skepticism by rejecting this principle.

Of course, the Moorean dogmatist grants that we have warrant for believing such things as that we are not disembodied spirits deceived by an evil demon. The Moorean dogmatist proposes that this warrant is *mediate*, stemming from our *immediate* warrant to believe such things as that we have hands. The Moorean dogmatist consequently represents the order of our warrants like this:


[8] The particular label, "doxastic justification," is irrelevant here. The proposal is not intended as an "analysis" of some pre-theoretical concept or term in everyday usage. What would matter is that the package captures a bundle of epistemic desiderata that are of significance to us.

Notice that the claim so far simply concerns the relations amongst the warrants a person can have to believe particular propositions. It is *not* a claim about what a person's beliefs are based upon or may acceptably be based upon, nor is it a claim that a certain inference or course of reasoning would be reasonable or epistemically acceptable.

Standardly, however, Moorean dogmatists also claim that this pattern of relations amongst warrants maps onto a fully reasonable pattern of argument or reasoning by which one could arrive at a doxastically justified belief that one is not a bodiless spirit under an evil demon's sway. Consider, for instance, a course of reasoning which we can represent thus:

Moore-1. [Visual experience of hands]

↓

Moore-2. I have a hand.

↓

Moore-3. I am not a disembodied spirit being deceived by an evil demon.

Pryor urges that if there is immediate perceptual warrant, then this reasoning "can be a way to acquire some warrant to believe Moore-3; and it can be reasonable to believe Moore-3 on the basis of it" (forthcoming: 7). He consequently has spent much effort attempting to explain away the seeming unattractiveness of this reasoning (Pryor 2004).

Acceptance of such reasoning is readily intelligible, given the Moorean dogmatist's aspirations in relation to external world skepticism. The Moorean dogmatist accepts a framing conception of much twentieth-century epistemological theorizing. According to this conception, if we are doxastically justified in believing anything about the world around us, then we should be able to reconstruct how an (ideal) epistemic subject could – without assuming any claims about the external world at the outset – acceptably reason or deliberate her way to warranted conclusions about the world, thereby acquiring doxastically justified beliefs. The Moorean dogmatist's aspiration, then, is to explain how this could be possible, not just for beliefs such as that we have hands, but also for beliefs such as that we are not bodiless spirits being deceived by an evil demon (under the assumption that the latter beliefs are not warranted by, and cannot be acceptably reached from, purely what is given to us through *a priori* reflection.) It is important to be completely clear about the Moorean dogmatist's thesis at this juncture. As Pryor puts it,

> In the happy case where you neither *have* nor *have reason to have* the kinds of doubts the skeptic wants to induce, then the [warrant] your experiences give you for Moore's premise [2] will be undefeated and [rationally] unobstructed. Having that [warrant] for the premise *will* make Moore's conclusion more credible for you; and that justificatory relationship is one that you can rationally endorse in your reasoning. (2004: 369)

The careful dogmatist's thesis is thus this. Suppose one has no background commitments about the world and no relevant antecedent information. One has no doubts, or reason to have doubts, to the effect that one is being deceived by an evil demon. One does not suspect that one is so deceived. And one has no beliefs to the effect that one possesses defeaters for or is not warranted in believing that one has hands on the basis of one's visual experience. If one now considers whether one is being deceived by an evil demon, then it is possible in principle for one to answer the question through the following course of explicit deliberation or reasoning. First, one arrives at a fully epistemically satisfactory belief about the external world on the basis of an immediate perceptual warrant, from there infers that one is not being deceived by an evil demon, and thus arrives for the first time at a fully epistemically acceptable belief that one is not so deceived.[9]

This thesis is not established merely by the thesis that one can have immediate experiential warrant for beliefs about the world. For consider how the responsibilist view sketched above would regard this reasoning. That view allowed that a visual experience as of your hands provides immediate warrant for the belief that you have hands. However, being a disembodied spirit deceived by an evil demon would be a disenabling condition for that warrant. Suppose, then, that you recognize that this is so (though perhaps not in so many words). You are in the position specified by the dogmatist's thesis. You are deliberating about whether to believe, on the basis of your visual experience, that you have hands. Suppose that you go ahead and form this belief on this basis. According to the responsibilist view, the belief will not be responsibly held, since you do not yet believe that you are not a disembodied spirit under an evil demon's sway. (That latter belief is supposed to be arrived at only in the next stage in the reasoning.) Since the belief that you have hands would not be responsibly held under such circumstances, it also wouldn't be doxastically justified. And if you go on to infer from it that you are not a disembodied spirit under an evil demon's sway, that latter belief will not be doxastically justified either. One

[9] In considering this thesis, we are of course supposed to bracket any concerns pertaining to the conditions for possessing contentful beliefs about the world at all.

cannot arrive at a doxastically justified belief by explicitly basing the belief upon another belief that is not doxastically justified. The resultant belief would be open to epistemic criticism.

So the responsibilist view could allow immediate experiential warrants but hold that the Moorean reasoning – if literally followed out in a course of explicit, conscious reasoning – will not enable one to arrive for the first time at a doxastically justified belief that one is not under an evil demon's sway. Again, to focus attention on the key issue: if one literally followed out the Moorean reasoning under the imagined circumstances, there would be a moment, just before one formed the belief that one has hands, at which one's deliberative position would be fully represented like this, "I recognize that if I were a disembodied spirit deceived by an evil demon, that would prevent my visual experience from telling in favor of the truth of the claim that I have hands, and as of yet I have no opinion whatsoever as to whether I am being deceived by an evil demon, but still, I conclude from my visual experience that I have hands." According to the view in question, that position is epistemically criticizable – it is epistemically irresponsible – and this is so even if one's visual experience gives one an immediate warrant to believe that one has hands.

So if this view is correct, there is a perennial epistemological aspiration that dogmatism tries – but fails – to fulfill. The dogmatist fails to explain how an ideal epistemic agent who considers the possibility that she is being deceived by an evil demon could fully explicitly reason from a starting position that takes for granted no claims about the world and thereby arrive at fully epistemically satisfactory beliefs about the world including the belief that she is not being deceived by an evil demon. Requirements at play in our ordinary notion of epistemic responsibility preclude the dogmatist from accomplishing this in the way that he hopes.[10]

This is in no way an argument in favor of the skeptical principle the dogmatist opposes: that principle is incompatible with the existence of immediate experiential warrants. Nor is this an argument against the dogmatist thesis that there are immediate experiential warrants. It isn't even an argument against the suggestion that a course of cognitive processing responsive to these warrants could ultimately yield an epistemically satisfactory belief that one isn't an evil demon's victim. Rather, it is an argument

[10] The Moorean dogmatist would not be helped by maintaining that one's visual experience of one's hands provides immediate warrant to believe that one is not a disembodied spirit being deceived by an evil demon. If that were so, then the Moorean reasoning would collapse into one step. But the objection presented above would still apply.

about what is acceptable in the course of fully explicit, conscious deliberation or reasoning engaged in by a mature rational agent who already recognizes certain things: it is an argument that even if there are immediate experiential warrants, one cannot utilize them in fully explicit deliberation or reasoning to arrive for the first time at doxastically justified beliefs to the effect that recognized disenabling conditions for those warrants do not obtain. This is not to say that we cannot be doxastically justified in believing that we are not victims of an evil demon's deceptions; only that if we are, it will not be by virtue of explicitly arriving at this belief via this sort of reasoning.

Not all Moorean dogmatists accept the reliability constraint on warrant that I appealed to in articulating and defending the responsibilist criticism of the Moorean reasoning. For this reason, they could reject the criticism offered in this section by rejecting part of its motivation. However, I did not mean to offer an internal criticism that every dogmatist would have to accept. Rather, I meant to articulate what I take to underlie our (pre-theoretical) dissatisfaction with the kind of reasoning that Moorean dogmatists endorse. As I see it, then, the ultimate question is this: which total package fits best with our ordinary epistemic judgments and practice? I've urged that the reliability constraint on warrant fits well, and the total package I've offered is strengthened by the fact that this constraint can help explain responsibilist requirements that enable a plausible account of our pre-theoretical resistance to the dogmatist's reasoning.

At the same time, however, the pre-theoretical evaluative judgments that fuel this resistance are independently compelling. In ordinary life we would all agree that (all else equal) a mature adult is open to criticism if she explicitly and consciously reasons from the fact that she has taken ibuprofen to the conclusion that her headache will soon lift while also avowing that she has no opinion whatsoever as to whether human physiology is such that ibuprofen never relieves pain in human beings. As long as we agree that relative to the respective courses of reasoning the possibility that one is being deceived by an evil demon is relevantly analogous to the possibility that human physiology is such that ibuprofen never removes pain, the objection to dogmatism stands regardless of the theoretical machinery that we bring in to explain the relevant similarity. And it is hard to deny that the two are relevantly similar.[11]

[11] It might be charged that I have failed to identify the source of resistance to the Moorean reasoning, because we would regard such reasoning as problematic – and in the same way – even for someone who failed to take the evil demon hypothesis to state a possible disenabling condition for the relevant

### *8.2.2 The upshot for Moorean dogmatism*

My argument has depended on distinguishing considerations about warrant from requirements of responsibility and other factors at play in doxastic justification. With this distinction in view, however, it becomes clear that the dogmatist's encounter with skepticism is not over. Suppose there are immediate experiential warrants. It might still be said that other factors relevant to doxastic justification dictate that a mature rational agent cannot acceptably believe anything about the world on the basis of sensory experience unless she has an independently warranted belief that she is not a disembodied spirit being deceived by an evil demon. This is to demand a warrant for this belief that does not constitutively depend – whether immediately or mediately – upon experiential warrants to believe things about the world (though not to demand it as a condition for having any experiential warrants at all). If one cannot – within the constraints of this requirement – have a doxastically justified belief that one is not being deceived by an evil demon, then this requirement commits one to external world skepticism. And this points out a very important lesson: the "dogmatist" view that there are immediate experiential warrants does not defeat a very familiar form of skeptical argument. More is needed.

Two broad options present themselves. The first is to accept this requirement, but to supplement the account of immediate perceptual warrant with an explanation of how we are non-experientially warranted in believing that we are not an evil demon's victims. The history of our subject

warrants. However, it is not in general true that one may not explicitly reason from a warrant to the non-obtaining of a condition that is in fact (though unbeknownst to one) a disenabling condition for that warrant. Suppose, for instance, that without the presence of a certain brain structure, human vision would be pervasively unreliable. The absence of this structure would thus be a disenabling condition relative to any visual warrants for beliefs about the world. Suppose, moreover, that no one yet recognizes the importance of this structure for the proper functioning of human vision; there is some question as to whether human beings even possess it, though it has already been identified in other primates (but without any understanding of its function or importance). Now consider the first brain researcher to identify this brain structure in human beings. Suppose that on the basis of visual inspections of brain dissections, he concludes that human beings do indeed possess this structure. In doing so, he bases that belief upon his visual warrant; he thereby consciously and explicitly moves in thought from a visual warrant to the conclusion that a condition does not obtain which (unbeknownst to him) in fact would be a disenabling condition for that warrant. Clearly, there is nothing objectionable about this procedure. So there is nothing in principle objectionable about moving from a particular warrant to the conclusion that what is in fact a disenabling condition for that warrant does not obtain. I see no reason not to conclude that what crucially matters for our distaste for the Moorean reasoning is something about the reasoner's position, and in particular, stems from the reasoner's recognition that the evil demon possibility is a disenabling condition relative to the relevant warrants.

suggests pessimism about such a response, though the idea of unearned default warrants may be of some help.

The second option is to reject this requirement. Here again two broad options present themselves.

The first is to claim not only that there are immediate experiential warrants, but also that there can be *immediate experiential doxastic justification*, that is, cases in which a belief is doxastically justified by experience in a way that does not constitutively depend upon one's having any other doxastically justified beliefs. However, I'm inclined to think that for mature believers at least, the demands of epistemic responsibility militate against this possibility, roughly because they require that in every case, one must acceptably take or be prepared to take something as telling in favor of the truth of the belief – which demands, in turn, that one have further doxastically justified beliefs.[12]

For this reason, I think that it is important – especially for those of us who are drawn to requirements stemming from epistemic responsibility – to notice a second option: one can deny this skeptical requirement without claiming that there can be immediate doxastic justification, even if one grants that there is immediate *prima facie* experiential warrant for (some) beliefs about the world. The key is to see that there would be no incoherence in a view which *grants*:

**1.** In order to be doxastically justified in believing anything on the basis of sensory experience, one must have a doxastically justified belief that one is not a disembodied spirit being deceived by an evil demon,

but which *denies*:

**2.** This latter belief must be doxastically justified in a way that does not constitutively depend upon one's having other doxastically justified beliefs about the world.

A view along some such lines would evade the skeptical conclusion by denying that in order to have doxastically justified beliefs about the world, one must have an *independent* warrant for believing that one is not an evil demon's victim. However, even if a view in this direction is coherent, it can be hard to see its possibility. To do that, we have to give up the assumption that if we have any doxastically justified beliefs about the world at all, then we should be able to reconstruct how an ideally reflective subject could – without assuming any claims about the external world at the outset – explicitly reason or deliberate her way to conclusions about the world,

[12] See Leite (2008) for an argument that this demand doesn't give rise to vicious regress.

making only acceptable moves at each step and thereby acquiring doxastically justified beliefs about the world. And to give up that assumption is of course to give up one of the traditional demands of epistemology.[13]

## 8.3 Appendix: responsibilist requirements and "transmission failure"

A prominent criticism holds that arguments such as Moore 1–3 suffer from failure of *warrant transmission* (Wright 2002, 2003, 2004, 2007). The rough idea behind this charge has two parts: first, that it is constitutive of the kind of warrant one has for Moore-2 that one couldn't have that warrant unless one had some warrant – other than any warrant provided by Moore-2 – to believe Moore-3; second, that because of this constraint, one's warrant to believe Moore-2 will not transmit across the inference from Moore-2 (based on Moore-1) to Moore-3.[14]

The "responsibilist" objection to Moore 1–3 is not a charge of transmission failure for warrant. For one thing, the "responsibilist" objection allows that visual experiences can provide *immediate warrants* for beliefs about the world, so it would not grant that one could not (as a constitutive matter) have the kind of warrant for Moore-2 that one has unless one also has some other warrant for Moore-3. Moreover, talk of "warrant transmission" across inferences or through courses of reasoning is itself rather odd, given the account of warrant I've presented. A dogmatist might want to say that the Moorean reasoning "can be a way to acquire some warrant to believe Moore-3" (Pryor forthcoming: 7). But this doesn't seem to be quite right, strictly speaking. Even if one does not go through the reasoning, one has a warrant to believe Moore-3 in virtue of undergoing certain visual experiences. It is a *mediate* warrant – a warrant which one has only in virtue of having warrant to believe certain supporting propositions – but it is a warrant all the same. So reasoning in accordance with Moore 1–3 could be

[13] I would like to thank the participants at the Conference on Inferential Internalism held in Fribourg, Switzerland in June, 2008 (especially Gian-Andri Toendury, Martine Nida-Rümelin, and Jim Pryor) and at the 2008 Midwestern Epistemology Workshop (especially Tyler Burge, Al Casullo, and David Henderson). I would also like to acknowledge my doctoral student Elizabeth Palmer for helping me see the attractiveness of relating considerations about warrant and considerations about responsibility in something like the way explored here. Thanks to Jonathan Weinberg, Ram Neta, Mark Kaplan, and especially Katy Abramson for comments and conversation. And thanks to Jim Pryor for his teaching and guidance, and for providing such a compelling articulation of Moorean dogmatism.

[14] On this characterization, the charge of warrant transmission failure is parasitic on denial that we have immediate perceptual warrant for beliefs about the world. It is a tricky question whether there is a version of the charge that does not assume the denial of immediate perceptual warrant. For discussion, see Pryor (forthcoming).

a way to acquire a warranted belief in Moore-3, but it does not follow that this belief will be epistemically satisfactory: given your visual experience, *any* way in which you acquire a belief in Moore-3 will yield a belief for which you have warrant. The warrant thus doesn't "fail to transmit" in the Moorean reasoning, because it isn't right to describe the warrant as something that could, or could not, *transmit*. On the account of warrant presented here, warrants aren't things that get transmitted through acts of inference or courses of reasoning. It may be that the state or condition of having gone through a course of reasoning can generate a warrant. But that is not a matter of a warrant transmitting from the premise to the conclusion.

Perhaps talk of "warrant transmission" is a fanciful way of expressing the conjunction of a *closure* claim for the relevant warrants and an "in virtue of" claim identifying which warrant is antecedent to the other. If that is what is meant, then we can understand talk of transmission failure in two ways, as claiming a failure of closure under entailment, or as granting closure but denying the relevant "in virtue of" claim.

Take closure first. Perhaps my visual experiences of my hands do not provide any warrant at all – not even a mediate one – for the belief that two plus two equals four, even though on many views "I have hands" entails "two plus two equals four." A similar claim might be made about the warrant situation regarding my visual experiences of my hands and my belief that I am not being deceived by an evil demon. (Certain ways of understanding the reliability requirement in the proposed account of warrant may also guarantee failures of warrant to be closed under entailment.) But the "responsibilist" objection to the Moorean reasoning did not involve any such claim.

Second, one could grant the relevant instance of closure but deny that it is in virtue of being experientially warranted in believing Moore-2 that one is warranted in believing Moore-3. So far as I can see, the only way to make this charge stick (given my earlier assumption that one's experience of a hand does not immediately warrant one in believing that one is not being deceived by an evil demon) is to claim that one could not have an experiential warrant to believe Moore-2 unless one had some other warrant to believe Moore-3.

The "responsibilist" objection does not depend upon any such claim. Rather, the charge is that even if one has a mediate warrant to believe Moore-3, one cannot responsibly come to believe Moore-3 for the first time through a line of explicit, conscious reasoning such as Moore 1–3. This is a charge about *acceptable courses of explicit, conscious reasoning*. It commits

the view sketched here to the following surprising possibility: the existence of mediate warrants which are such that in certain cases it *is not possible* for an ideal rational agent to arrive at a fully epistemically satisfactory belief by explicitly inferring in a way that tracks their structure. But consider the proposition that one currently has no occurrent second-order beliefs. That's a straightforward example of a proposition which someone could have a warrant for believing but which the person *can't* reasonably come to believe by reasoning on the basis of that warrant, though the precise failure in this case is admittedly rather different.[15]

Might we aptly characterize the failing I've ascribed to Moore 1–3 as a failure of *doxastic justification* to transmit through an inference? The core problem is this, call it what you like: in the envisioned circumstances, one cannot responsibly come to believe Moore-2 by consciously basing it on one's visual experience. As a consequence, one would not be doxastically justified in believing Moore-3 on the basis of Moore-2, just as one would not be doxastically justified in believing *any* conclusion on the basis of Moore-2, since one cannot acquire a doxastically justified belief by explicitly drawing an inference from a belief which is not responsibly held.

Suppose that we modify the situation slightly, so that one *is* doxastically justified in accepting Moore-3 at the stage in the deliberation just before one accepts Moore-2. Would the reasoning then exhibit a failure of doxastic justification to transmit through an inference? In such a case, one would acquire a doxastically justified belief in Moore-2. And one would already have a doxastically justified belief in Moore-3. Suppose, then, that one went on to try to base one's belief in Moore-3 on one's belief in Moore-2. We might want to say that at the end of this procedure one's belief in Moore-3 remains doxastically justified. Or we might want to say that it doesn't because this procedure renders one's belief in Moore-3 (at the end) epistemically criticizable. But so far as I can see, not much (if anything) turns on this issue. Either way, something is going wrong.

[15] This example comes from Juan Comesaña. He isn't to be held responsible for my use of it.

CHAPTER 9

# *Primitively rational belief-forming processes*

*Ralph Wedgwood*

One approach to understanding 'reasons' – including both reasons for action and reasons for belief – postulates a fundamental connection between *reasons* and *reasoning*. According to this approach, there is a reason for you to φ – where φ-ing could be either an action (such as *writing a letter*) or an attitude (such as *believing that it is snowing*) – if and only if there is a possible process of sound or rational reasoning that could take you from your current state of mind to your rationally φ-ing.

This approach is familiar in discussions of reasons for action. As Bernard Williams (1995: 35) put it, for there to be a reason for you to φ, there must be a 'sound deliberative route' that leads from your current state of mind to your being motivated to φ. Presumably, Williams's reference to 'a sound deliberative route' comes to more or less the same thing as my own terminology of a 'possible process of rational reasoning'.[1]

It is clearly possible to take the same approach to understanding reasons for belief. You have a reason to believe *p* if and only if there is some possible process of rational reasoning that could lead you from your current state of mind to your being rationally inclined to believe *p*. So one way to make progress in understanding reasons for belief is to investigate which processes of reasoning count as rational.

In this chapter, I shall investigate this notion of rational processes of reasoning. Specifically, I shall focus here on *belief-forming processes* – that is, types of mental process that result in the thinker's forming a belief in some proposition. A full theory would also have to consider other processes as well – such as processes that result in the thinker's *revising* her beliefs in various ways. But for simplicity, I shall focus exclusively on belief-*forming*

[1] Admittedly, there are other approaches to understanding reasons – such as the approach of John Broome (2004). But for my purposes, it is enough if there is at least *one* concept expressed by the term 'reason' that is to be understood along the lines of this approach. (It does not matter for my purposes if there are also *other* concepts that can be expressed by the term 'reason'.)

processes here. I shall try to solve a basic problem that arises when we try to understand how belief-forming processes can be rational.

## 9.1 How can there be fallible primitively rational processes?

The problem that I shall try to solve concerns belief-forming processes that have the following three properties:

**a.** They are *rational* processes, and the beliefs that we form through these processes are themselves rational or justified beliefs.

**b.** Even if in most cases these processes reliably lead to correct beliefs (that is, beliefs in true propositions), they are *not infallible*: it is still logically possible for beliefs that are formed through these processes to be incorrect (that is, to be beliefs in *false* propositions).

**c.** The rationality of these processes is *basic* or *primitive*. That is, the rationality of these processes is not due simply to the availability, purely by means of some *other* rational processes, of a rational or justified belief in the reliability of these processes. As we might put it, the rationality of these processes is not due to the availability of a 'process-independent justification' for these processes.

In this chapter, I shall focus on a specific example of a belief-forming process that has these three properties – what I shall call the process of 'taking experience at face value'.

Very roughly, the process of taking experience at face value is the process whereby one comes to believe a proposition *p* in response to the fact of one's having a conscious experience that has *p* as part of its representational content. For example, if my present experience has the proposition 'I am holding my hands in front of my face' as part of its content, then this process would involve my responding to this experience by forming the belief that I am indeed holding my hands in front of my face.

This initial specification of the process is obviously too rough. I could have an experience of this kind even if it also appeared to me that a demon was talking to me out of the palm of my hand, mocking me with the claim that all of my experiences are complete illusions. In this case, it would clearly *not* be rational for me to form the belief that I really am holding my hands in front of my face. If this process of 'taking sensory experience at face value' is to be rational, it is not sufficient for forming a belief by means of this process that one just forms a belief in *p* in response to *any* experience that has *p* as part of its content. It must also be the case that one's experiences, background beliefs, and other mental states do not contain

any *special defeating reasons* of this kind. So a more precise name for this process would be the 'process of taking one's experiences at face value, *when such special defeating reasons are absent*'.[2] But for the most part, I shall omit this qualification here.

Even with this qualification, however, this is still only a rough description of the process of taking experience at face value. First, this description says nothing about the fact that even when a proposition *p* is part of the representational content of one's experience, the *degree of belief* that one may place in *p* through taking one's experience at face value can vary enormously, depending on how specific this proposition *p* is. The proposition that *there are exactly 26 lights arranged in a circle in front of one's eyes* may be part of the representational content of one's experience; but without careful counting, one should presumably place a much lower degree of belief in this proposition than in the weaker and less specific proposition that *there are some lights before one's eyes*.

Secondly, this description of the process of taking one's experience at face value may also require revision if it is true, as many philosophers have argued,[3] that the primary representational content of sensory experience is *non-conceptual*. However, even if this is true, we could still make sense of what we might call the 'conceptualized upshot' of experience, which would be the result of one's conceptual capacities' operating on the primary non-conceptual content of experience. Then we could refine our description of the process of taking one's experience at face value by saying that it involves forming the belief in a proposition *p* in response to having an experience that has *p* as part of its 'conceptualized upshot'. In general, it will clearly be a difficult and controversial matter to give a fully precise account of the process of taking experience at face value. I believe, however, that the rough description that I have given will be sufficient for present purposes.

It is important, however, that as I have characterized it, this process of taking one's experiences at face value need *not* involve relying on any antecedent belief in the reliability of one's experiences (or indeed on any belief about one's experiences at all). This process involves coming to believe *p* in *direct response* to having an experience that has *p* as part of its representational content (together with the absence of defeating reasons of

[2] In his comments on this essay, Stephen Schiffer argued that my reference to 'the absence of special defeaters . . . stands in for something extremely complicated'. But in fact, it does not 'stand in for' anything. We have a general notion of a defeater – that is, the general notion of a set of mental states that one is in at the relevant time such that, although normally forming a belief in a certain way would count as rational, it does not count as rational at that time because of the presence of these mental states. It is this general notion that I am using here.

[3] For arguments that the content of sensory experience is non-conceptual, see Peacocke (2001) and Burge (2003); for the other side of the debate, see McDowell (1994b).

the relevant sort); it is not required that one should have any belief *about* one's experiences.

It seems plausible that this process has the three properties that I listed above. First, it seems rational to form beliefs by means of this process – as I shall sometimes say, it is rational to 'engage in' this process. Since my current experiences, background beliefs, and other mental states contain no special defeating reasons of the relevant kind (I do not seem to see any demon talking to me out of the palm of my hand or the like), it is rational for me now to form the belief that I am holding my hands in front of my face; and the belief that I form in this way is itself a rational or justified belief.

Secondly, this is obviously a fallible process. If I really were being undetectably deceived by an evil demon, then the overwhelming majority of the beliefs that I would form by means of this process would be false, and so in this case, this process would fail to be a reliable guide to the truth.

Thirdly, it also seems plausible that this process is 'primitively' rational. It does not seem plausible that the rationality of this process is explained merely by the availability of any process-independent justification for the process.[4] Admittedly, it is not as obviously true that this process has this third property as that it has the first two properties that I listed above. But in this chapter, I shall simply assume that this process has these three properties, without making any attempt to justify this assumption. I am invoking these assumptions in order to illustrate a more general problem. If these three assumptions are not in fact true of the process of taking experience at face value, I would only need to pick another process that does satisfy these three assumptions; and it seems indisputable that there are at least some processes of this kind.

Now the idea of such primitively rational belief-forming processes seems to give rise to a problem. How *can* a belief-forming process (i) be rational, given that (ii) the process is fallible and might fail to be leading us to the truth, and (iii) it is not rational in virtue of the availability of any process-independent justification of the process?

The problem seems to arise from the following point: it seems that if a belief-forming process is rational, then forming beliefs by means of the process must be a *reasonable way to pursue the truth* – that is, to pursue the

[4] My claim that the process of taking one's sensory experiences at face value is primitively rational is similar in sprit to the position that James Pryor (2000) has called 'dogmatism'. My claim agrees with Pryor's 'dogmatist' that the rationality of taking your sensory experience at face value is not *explained* by any justification that you have for believing in the reliability of your experiences. However, although I agree with this dogmatist on this point, I do not agree with another thesis that many dogmatists endorse – the thesis that the rationality of this process does not even *entail* the availability of such justification; but unfortunately, I cannot go further into this question here.

goal of believing the proposition in question if and only if that proposition is true. But if a process is perfectly capable, in some possible circumstances, of leading one to believe falsehoods rather than truths, then how could the process be a reasonable way for one to pursue the truth, unless one has some process-independent reason for regarding the process as reliable in one's actual circumstances?[5]

This account of the problem points in the direction where we must look to find a solution. In effect, the solution would involve (a) giving a more precise clarification of the intuitive formulation that I have just used – a 'reasonable way for one to pursue the truth' – and (b) explaining how the process of taking experiences at face value really does satisfy this formulation when it is clarified in this more precise way.

An ambitious version of such a solution would be in a sense *reductive*: that is, it would clarify this intuitive formulation by *analysing* what it is for something to be 'a reasonable way for one to pursue the truth' in entirely *non-normative* terms, and then it would aim to show that the process of taking one's experiences at face value satisfies the conditions that are given in the analysis.

I am not confident that any such reductive analysis of rationality is possible. The rationality of taking one's experiences at face value may simply be a bedrock normative truth, of which no such ulterior explanation can be given. So in clarifying the idea of a process that is a 'reasonable way of pursuing the truth', I shall not offer any such reductive analysis. Instead, I shall simply articulate some crucial features that all rational belief-forming processes must have, without attempting to ensure that these crucial features are picked out in strictly non-normative terms. Then I shall try to dispel the sense of puzzlement that surrounds the thesis that the process of taking one's experiences at face value is rational, by making it plausible that the process has these crucial features.

## 9.2 ASSUMPTIONS ABOUT RATIONALITY

In the discussion that follows, I shall rely on a number of assumptions about rationality. First, I shall assume that the view that epistemologists call 'internalism' is correct. That is, I shall assume that the processes through

[5] Indeed, as we shall see later on, there are even some *infallible* belief-forming processes that seem not to count as reasonable ways to pursue the truth. So the fallibility of the process of taking experiences at face value is not in fact essential to the fundamental problem that we are concerned with here. The fundamental problem is simply that it is not clear how *any* belief-forming process can be primitively rational: focusing on fallible processes serves only to make this problem more intuitively vivid.

which it is rational for a thinker to form beliefs at a given time *t* are always determined purely by the facts that are in some way "internal" to the thinker's mind at *t*. Specifically, I shall assume that whether or not it is rational for a thinker to form beliefs by means of a given process at *t* is entirely determined by the facts about what mental states the thinker has, what mental states the thinker lacks, and what mental abilities and capacities the thinker has, at *t*.[6]

If internalism about rationality is correct, the rationality of a belief-forming process must supervene purely on the thinker's mental states. The belief-forming process itself must be individuated purely in terms of the relations between the internal mental states involved; and the features of the process that make it a rational belief-forming process must consist purely of features that the process has in virtue of its role in the thinker's mind. Nothing that is external to the thinker's mind can be involved in making the process rational.

The most prominent rival to this internalist view of rationality is the 'reliabilist' view. According to the reliabilist view, whether or not a belief-forming process is rational depends entirely on whether it is a reliable way of arriving at the truth.[7] I shall not attempt to defend internalism, or to criticize reliabilism, in this chapter. I shall simply assume for the sake of argument that internalism is true. This internalist assumption makes the problem that I am concerned with here significantly harder to solve. The reliabilist has an easy response to this problem. The reliabilist can just say, 'If, as a matter of fact, the process of taking one's experiences at face value is reliable, then it is a source of justified or rational beliefs; otherwise, it is not a source of justified or rational beliefs.' On the other hand, if the internalist view of rationality is correct, and reliabilism is incorrect, then this easy response is not available.

The second assumption about rationality that I shall rely on here is an assumption that I have already articulated – intuitively and metaphorically – by saying that the rational belief-forming processes must all be 'reasonable ways of pursuing the truth'. In other words, there must be some fundamental connection between rational belief-forming processes

[6] This sort of epistemological 'internalism' is attacked by Goldman (1999a) and defended by Conee and Feldman (2001). I have offered my own defence of internalism elsewhere (Wedgwood 2002b); as I explained there, in characterizing internalism, we need to use a notion of 'mental states' that excludes the so-called 'factive mental states' that play such a prominent role in the epistemology of Williamson (2000).

[7] For this reliabilist conception of justified belief, see Goldman (1979).

and the truth. But exactly *is* it this fundamental connection to the truth that is a necessary feature of all rational belief-forming processes?

There are several possible connections between belief-forming processes and the truth that we could consider. First, some belief-forming processes themselves lead to the conclusion that these processes are reliable guides to the truth; engaging in these processes will itself lead us to believing in the reliability of these processes. As we might put it, these processes 'recommend themselves' as reliable or truth-conducive processes. Still, although this may be a *necessary* condition that rational belief-forming processes must meet, it cannot be a *sufficient* condition for the rationality of such processes – since many crazily irrational processes might recommend themselves in this way.

Secondly, some belief-forming processes are in a sense *infallible*: if one genuinely forms a belief by means of this process, the belief is bound to be true. One example of such an infallible process is the process of forming the introspective belief that one could express by saying 'I am in pain' directly in response to the fact that one is in pain. But as we have seen, the process of forming beliefs by taking one's experiences at face value is an eminently *fallible* process, since one could form a belief by means of this process even if one was undetectably being deceived, so that the belief in question was not true at all.[8]

Thirdly, there are certain other processes that count as rational ways to form beliefs, but only because we have some *process-independent* way of forming a rational belief that the process is a reliable guide to the truth. For example, this seems to be the case with the processes in which we form beliefs on the basis of the use of measuring instruments. It is rational for me (in the absence of defeaters) to form a belief about the temperature of the air on the basis of reading a certain thermometer; but this is only because it is *already* rational – quite independently of the process of relying on this particular thermometer – for me to believe that the thermometer is reliable. However, I am assuming that this is not the case with the process of taking one's experiences at face value. If this is a rational process at all, it must be, as I put it, a *primitively rational* process.

Thus, it seems that none of these three connections to the truth can be a central part of what makes it rational to engage in the process of taking

[8] It may be correct to interpret much of the epistemology that was inspired by Descartes as motivated in part by the ambition to reconstruct as much of our knowledge of the world as possible purely on the basis of these infallible belief-forming processes. I am assuming here that, contrary to the presuppositions that seem to lie behind this tradition, fallible belief-forming processes can be perfectly rational.

one's experiences at face value. Still, it seems that *some* such connection to the truth must be a central part of what makes it rational to engage in this process. We need to find a *fourth* connection to the truth.

## 9.3 A CONNECTION TO THE TRUTH

Given my internalist assumptions, any primitively rational belief-forming process is rational purely in virtue of the nature of the internal mental states involved. Thus, if a connection to the truth is part of what makes the process of taking experiences at face value a rational belief-forming process, this connection to the truth must also hold purely in virtue of the nature of the mental states involved. That is, whatever exactly this connection to the truth may be, it must be *necessary* that, given all the mental states involved in the process, the process has this connection with the truth. So, the process must have this connection to the truth in all possible worlds – even in 'demon worlds' where (because of the machinations of deceiving demons and the like) the process is not a reliable way to pursue the truth.

What sort of connection to the truth could this be? It must obviously be a *weak* connection to the truth, if the process has this connection with the truth even in 'demon worlds', where the process fails to be a reliable guide to the truth.

I propose that every primitively rational process has the following connection with the truth. The essential nature of the mental states that are involved in this process requires that any thinker who is capable of these mental states must have certain *dispositions*; and when all of these mental states arise from the manifestation of these essential dispositions, the process really is a reliable guide to the truth. With the process that I have called 'taking experiences at face value' the mental states involved are (a) an experience, and (b) the belief that one forms on the basis of the experience. So to see whether this process really has this sort of connection with the truth, we must investigate the essential nature of these mental states, to find out what are the dispositions that any thinker capable of these mental states must possess.

Of course, questions about the nature of experience are among the most controversial and disputed questions in the philosophy of mind. But most philosophers would agree that every experience has a *content*. This content is a representation of a possible state of affairs. When this state of affairs actually obtains, the experience will count as a *veridical* experience. For example, my current experience could represent the state of affairs of my

holding my hands in front of my face; if I really am holding my hands in front of my face, then it counts as a veridical experience.

I tentatively suggest that it is an essential feature of sensory experiences that any subject who has experiences at all must have some *disposition* to have experiences that veridically represent certain aspects of her environment. This suggestion should not be read as implying that every such subject has a disposition for *completely* veridical experience – that is, a disposition to have experiences that do not represent *any* aspects of the environment non-veridically (let alone a disposition to have god-like experiences that represent *all* aspects of the environment veridically). Instead, this suggestion should be read as implying only that for every subject who has experiences at all, there are *some* aspects of the environment such that the subject has some disposition to have experiences that veridically represent *those* aspects of the environment. Presumably, one cannot have a disposition to have such veridical experiences by *chance*; so the relevant disposition must be a disposition to have an experience that veridically represents a certain state of affairs precisely in *response* to the fact that that state of affairs really *does* obtain. To fix ideas, we may assume that disposition ascriptions are equivalent to generalizations about what happens in 'normal conditions' – in other words, in conditions in which *cetera* are *paria* or 'other things are equal'.[9] More precisely, then, my suggestion is this: it may be a necessary truth that, for every subject of experience, there is a range of propositions such that for every proposition *p* within that range, in normal conditions, the subject will respond to being in a situation in which *p* is the case by having a sensory experience as of *p*'s being the case.[10]

Why might this suggestion be true? It may be that this disposition to have partly veridical experiences is essential to explaining how experience has the specific *content* that it has. That is, the fact that a mental state is of the kind that would be involved in manifesting this disposition in response

[9] It will obviously be a challenging task to give a detailed account of 'normal' conditions (i.e. conditions in which 'other things are equal'). But it seems that it is a task for metaphysics and the philosophy of science in general; it is not a special problem for the branch of epistemology that I am focusing on here.

[10] How can this disposition be ascribed to a subject of experience who is radically deceived, like the 'brain in a vat'? Perhaps the reason is that all subjects of experiences must have something like a *brain* that was, at least at some point during its existence, connected to a body in such a way that the subject was capable of perceiving and acting in a normal environment, and that brain must continue to function internally the same way as it did when it was connected to such a body. This may be enough to make it the case that the subject has this sort of disposition for veridical experiences (even though its current envatted condition now prevents this disposition from being manifested). But clearly the issue will require much further investigation.

to being in a situation in which *p* is the case may be an essential part of what *makes* this mental state count as a sensory experience as of *p*'s being the case in the first place.[11]

This suggestion does not imply that *all* of our sensory experiences arise from the operation of this disposition for veridical experiences. It only implies that we would not be capable of having experiences at all unless we had this disposition. Even if the manifestation of this disposition is blocked or inhibited by other factors (such as the machinations of an evil demon), the subject may still *have* this disposition. Admittedly, it may be impossible to have this disposition unless the disposition has at least sometimes been manifested. So it may be impossible to have experiences at all unless one has at some time in one's life actually manifested this disposition for veridical experiences. This would be a modest sort of 'anti-individualism' about sensory experiences. However, this sort of anti-individualism is still compatible with 'mid-life envatting': you might become a 'brain in a vat' only *after* your sensory experiences have been appropriately connected to the environment for long enough for them to be genuine conscious experiences with representational contents. So this sort of anti-individualism is still compatible with the possibility that you are now being radically and undetectably deceived.

Nonetheless, this sort of anti-individualism about experience does guarantee a certain sort of connection between the process of taking one's experiences at face value and the truth. Whenever one's experience *does* consist entirely in the manifestation of this essential disposition, then the content of the experience will be true. In short, there is a certain disposition, which is essential to having the capacity for sensory experiences at all, and when the process of taking one's experience at face value involves the manifestation of this disposition in this way, the process is reliable.

This sort of connection to the truth has the two features that we noted at the beginning of this section. It is clearly a *weak* connection to the truth, since a thinker can *have* such a disposition to have veridical experiences, even if the manifestation of this disposition is currently blocked or inhibited – for example, by the machinations of a deceiving demon. Nonetheless, this is still a connection that holds *necessarily*, in virtue of

[11] Compare Peacocke's (2004: 69) idea that experiences are '*instance-individuated* with respect to certain of their contents'. Compare also Burge's (2003: 511) '*perceptual anti-individualism*' – the idea that the 'nature of a perceptual state, as marked by its representational content, is partly determined by relations between the perceptual system . . . and features of the environment that cause instances of its states and that states of the system represent'.

the essential nature of the internal mental states that are involved in the process of taking experiences at face value.[12]

It may be that some such essential connection to the truth is a necessary part of what makes any belief-forming process primitively rational. Let us return to the metaphorical formulation that I used earlier, when I said that every rational belief-forming process must be a 'reasonable way of pursuing the truth'. To explore the significance of this metaphor, let us imagine that you were *literally* 'pursuing' the truth about a range of propositions; that is, for every proposition *p* in this range, you actually intend to realize the goal of believing *p* if and only if *p* is true. Which processes would you choose to engage in as your means for pursuing this goal? In order to model our 'internalist' assumptions within the context of this metaphor, suppose that you had to choose which belief-forming processes to engage in purely on the basis of the 'internal' features of this belief-forming process. In particular, how are you to choose which processes to treat as *basic* – that is, as the processes that you choose to engage in even in advance of any process-independent justification of those processes? It seems that one would have to choose processes that had some connection to the truth that held necessarily, purely in virtue of the nature or essence of the mental states that are involved in that process.

One might wonder whether a thinker who was choosing which processes to treat as basic in this way would in fact always choose a set of processes that is *infallible*, in the sense that I discussed earlier. Why is it not the case that the only processes that it is reasonable for one to treat as basic in one's pursuit of the truth are such infallible processes?

In fact, it seems that it would not be reasonable of the thinker to limit her belief-forming processes to these infallible processes. The goal is not just to *avoid* believing the proposition in question *p* if *p* is *not* true; it is also to *believe p* if *p is* true. To pursue this goal effectively, one needs processes that are not just *reliably error-avoiding*, but also *reliably belief-yielding*.[13] While a set of belief-forming processes that includes only infallible processes could be reliably error-avoiding, it would not score well with respect to being belief-yielding. In particular, these infallible processes only seem capable of giving us beliefs about our own introspectible mental states and their logical consequences. But it seems plausible that we absolutely need to have beliefs

[12] Compare George Bealer's (2000: 7–10) 'modal reliabilism', according to which any 'basic source' of evidence must have a weak but necessary connection with the truth.

[13] For a discussion of this point in slightly more detail, see some of my earlier work (Wedgwood 2002a) – although a more complete treatment of this issue would admittedly require much more investigation.

about the contingent features of the external world if we are to be intelligent agents at all. No set of cognitive processes could score well with respect to being belief-yielding unless it is at least possible for these processes to give us beliefs about the external world. So it seems that it would not in fact be a 'reasonable way for one to pursue the truth' to limit oneself to these infallible belief-forming processes. The only sort of connection to the truth that is *necessary* for all primitively rational belief-forming processes is the *weak* sort of connection that I have described.

Still, it seems doubtful whether any such link to the truth can be a *sufficient* condition for the rationality of a belief-forming process. There seem to be many possible belief-forming processes that have some such necessary connection to the truth but are not intuitively rational. As BonJour (1980) in effect famously pointed out, the reliability of a process will not make the process rational if the configuration of mental states in the thinker's mind does not *justify* the thinker in believing the process to be reliable.

To illustrate this point, consider the belief-forming process that consists in coming to believe a certain necessarily true proposition directly and immediately on the basis of simply considering the proposition. For example, this necessarily true proposition could be a highly non-obvious mathematical theorem; or it could be a necessary truth that can only be known empirically, such as 'Water=$H_2O$'. Intuitively, this process is not rational even though it is necessarily reliable. It is only rational to form a belief in a non-obvious mathematical theorem on the basis of something like a *proof* (or on the basis of what one rationally believes to be reliable testimony); it is only rational to form a belief in the proposition that water is $H_2O$ on the basis of empirical evidence. It would not help to make this process rational if a whimsically manipulative neuroscientist had somehow 'wired you up' in such a way that this process was as psychologically basic for you as (say) the process of immediately coming to believe obvious logical truths like '1=1'. And if being wired up to form beliefs by means of this process by a whimsical neuroscientist cannot help to make it rational, I do not see how being wired up to do so by nature can either.

Thus, even though this connection between the process of taking one's experiences at face value and the truth is a necessary condition of the primitive rationality of the process, it is not a sufficient condition of its rationality. The fundamental problem seems to be that it is not enough if one just *happens* to form beliefs by means of a process that has this sort of connection to the truth. It is also necessary that on the relevant occasions, one forms beliefs by means of this process *precisely because* the process has this sort of connection to the truth. In this sense, it is only processes that

one can engage in precisely *because* of their connection to the truth that can count as 'reasonable ways of pursuing the truth'. But we still do not yet understand how it is possible for one to form beliefs by means of a belief-forming process precisely because of the process's connection to the truth. In the last two sections of this chapter, I shall turn to another essential feature of primitively rational belief-forming processes, which seems to me to be capable of illuminating this idea in more detail.

## 9.4 *A PRIORI* PROCESSES

Philosophers often distinguish between *a priori* justification and empirical justification. It is widely assumed that it is only *beliefs* that can have either *a priori* or empirical justification. However, there is a natural generalization of this notion of *a priori* justification, so that it applies to *belief-forming processes* as well as to beliefs. In this section, I shall first explain what it would mean for the rationality of a belief-forming process to be *a priori*, and then I shall argue that the rationality of the primitively rational processes is *a priori* in this way.

When a belief-forming process is rational, we can ask: What *makes* this process rational? In some cases, the rationality of the thinker's engaging in the process depends on the specific experiences that the thinker has had, or on the thinker's empirical background beliefs. In these cases, the rationality of the process is empirical. But in other cases, the rationality of the thinker's engaging in the process is independent of absolutely all the specific experiences and empirical beliefs that the thinker has had. In those cases, the rationality of the process is *a priori*.

If the rationality of a belief-forming process is *a priori*, and so independent of all of the thinker's specific experiences and empirical beliefs, what does make the process rational? In some special cases (for example, in the case of certain special mathematical proof techniques), it may be that the process is rational because the thinker has some independent *a priori* proof of the process's reliability. But this cannot be the case with the *primitively* rational processes – which we have defined to be processes the rationality of which is *not* due to the availability of any such independent justification. So it seems that if the rationality of a primitively rational process is *a priori*, then the process must already be rational for *all* thinkers who have the capacities that are necessary for engaging in the process. Such a process would be rational purely because the rationality of this process is in some way '*built into*' the structure of those cognitive capacities themselves.

I am relying here on a broadly Kantian conception of the *a priori*. This is how Kant (1787: 1) introduces the central question of the *Critique of Pure Reason*:

> But even though all our cognition begins with experience, it does not follow that it all arises out of experience. For it may well be that even our empirical cognition is composed both out of what we receive through impressions and out of what our own cognitive faculty (sensory impressions serving merely as the occasion) supplies out of itself.

Here Kant seems to be assuming that the *a priori* is what our cognitive faculties or capacities 'supply out of themselves'. Our cognitive capacities 'supply' a belief-forming process 'out of themselves', we may suppose, just in case any thinker who has those capacities either already has the ability to engage rationally in that process, or else is at least in a position to acquire that ability simply through exercising those capacities, without relying on any experiences or any other mental states that are not necessarily available to everyone who possesses those capacities.

This then is what it would be for the rationality of a basic belief-forming processes to be *a priori*: it would have to be a process that our cognitive capacities 'supply out of themselves' in this way. But why should we think that the primitively rational processes are *a priori*?

If we consider our favourite example of a primitively rational process – the process of taking experience at face value – it seems plausible that the rationality of this process is *a priori*. It seems to be rational for *every* thinker to engage in this process – or at least for every thinker who has the capacities that are necessary for engaging in the process – regardless of the specific experiences that they have had. Even if the thinker has the experience as of a demon's speaking to him out of the palm of his hand, mocking him with the claim that all his experiences are illusory, it would not become irrational for the thinker to engage in this process; since the process only involves forming beliefs on the basis of experience when special defeating considerations are absent, and in this case such special defeaters are present, it is simply *impossible* to form beliefs by means of the process in this case. Thus, the rationality of forming beliefs by taking experience at face value seems to be independent of all experiences and empirical beliefs whatsoever, and so to be *a priori*. As we might put it, this process cannot itself be read off from one's experiences, because it is precisely by means of this process that one reads anything off from one's experiences in the first place.

This point can be generalized to all the primitively rational belief-forming processes. If the rationality of a belief-forming process were empirical rather than *a priori*, then the rationality of the process would *depend* on the specific experiences and empirical beliefs that the thinker has had: if the thinker had had sufficiently different experiences and empirical beliefs, the belief-forming process would not have been rational. But what possible collection of experiences and empirical beliefs could undermine the rationality of these basic belief-forming processes in this way? There presumably are some possible collections of experiences and empirical beliefs that would prevent it from being possible to form rational beliefs by means of these processes. But these collections of experiences and empirical beliefs will in effect count as a *defeating condition*, and it seems plausible that each of these primitively rational processes is structured in such a way that it already involves sensitivity to the presence or absence of such defeating conditions: forming beliefs in the presence of these defeating conditions would not be an instance of this process at all. For this reason, it is just not true that when these defeating conditions are present, it becomes irrational to form beliefs by means of these processes. Thus, it seems that the rationality of these basic processes cannot be undermined by any possible collection of experiences and empirical beliefs whatsoever. In short, the rationality of these basic processes seems independent of absolutely all experiences and empirical beliefs – that is, it seems *a priori*.

If the rationality of a belief-forming process is *a priori*, the process must be rational for *all* thinkers who possess the relevant cognitive capacities, regardless of the experiences and empirical beliefs that they have. Presumably, if it is rational for a thinker to engage in a belief-forming process, the thinker must either already possess the ability for rationally engaging in the process, or at least already be in a position to acquire this ability. So if a rational process is *a priori*, all thinkers who possess the relevant cognitive capacities must either already have the ability to engage rationally in the process, or at least be in a position to acquire an ability to engage rationally in the process (simply through exercising these capacities, without relying on any experiences or additional mental states that are not necessarily available to all thinkers who have those capacities).

However, if the process in question is a *primitively* rational process, it seems that all thinkers who have the relevant cognitive capacities must *already* have the ability to engage rationally in the process. The simplest explanation for this would be if the possession of these capacities *essentially requires* having some *disposition* to engage in the process. As I shall now propose, this is why these primitively rational processes count as *a priori*: a

process is primitively rational if and only if there is some cognitive capacity of the relevant kind that essentially requires some disposition to engage in the process.

What is the 'relevant kind' of 'cognitive capacity'? One sort of capacity that is surely relevant is the *possession of a concept*. For example, the possession of the concept 'green' is the capacity to have thoughts and attitudes involving this concept (that is, to have the sorts of thoughts and attitudes that are normally ascribed in English by means of attitude-ascribing phrases in which the term 'green' appears in a hyperintensional context, such as 'thinking that grass is green', 'hoping for green pastures', and the like).[14]

It would, however, be unwarranted to assume that our possession of concepts is the *only* sort of capacity that is relevant here. It is equally important that we are capable of various sorts of *attitudes*, such as belief and intention. Moreover, in drawing up an inventory of the attitudes of which we are capable, it is important not to overlook the full range and complexity of these attitudes. Belief itself comes in various *degrees* of belief; and our thoughts can also be mere suppositions or wonderings rather than full-blown beliefs. Similarly, intentions seem *prima facie* different from other conative states, like aims and wishes and desires. So there are a great many more types of attitudes than just belief and intention.

In general, the capacities that are relevant here include: (i) one's possession of each of the various concepts that one possesses, and (ii) one's capacity for each of the various types of attitudes and mental states of which one is capable. My current proposal, then, is that what makes the *primitively* rational cognitive processes *a priori* is that they are precisely those processes that we must have some disposition to engage in if we are to possess these basic cognitive capacities.[15]

For example, consider the capacity of possessing the concept 'if' – the capacity for thinking conditional propositions. Perhaps possessing this

[14] In this way, my proposal has a lot in common with that of several other philosophers – such as Boghossian (2002, 2003a), Peacocke (2004), and Pollock and Cruz (1999) – who have all maintained that many of the most fundamental belief-forming processes are in some way implicit in our possession of concepts. However, there are several differences between my proposal and theirs: the most important difference is that the possession of concepts is by no means the *only* sort of cognitive capacity that I am appealing to here; the capacity for certain types of *attitude* is at least as important as the possession of concepts.

[15] An additional advantage of this proposal is that it may help to solve the problem of how to *individuate* the rational belief-forming processes. According to this proposal, the primitively rational processes are the processes that are built into the nature of these basic cognitive capacities; so an investigation of the nature of these cognitive capacities should also tell us how to individuate these primitively rational processes.

capacity essentially involves a disposition to engage in the process of making inferences that have the form of *modus ponens*: 'If *p* then *q*, but *p*; hence *q*.' Perhaps, if you had absolutely no disposition of this sort, you could not even possess the concept 'if' at all.[16] In that case, according to the proposals that I have made here, the process of making *modus ponens* inferences would be a primitively rational process; and the way in which the process is built into our possession of the concept 'if' would explain why it counts as an *a priori* cognitive process.

Can this account of why the primitively rational processes count as *a priori* also be applied to the process of taking one's experiences at face value? Suppose that it is the case that if one had no disposition to engage in this process, one would be incapable of forming beliefs in any contingent propositions about one's environment. Without such beliefs, one would not be capable of intelligent agency or practical reasoning at all, since all such practical reasoning has to start from some beliefs in contingent propositions about one's environment. Then there would be a particularly basic cognitive capacity – the capacity for intelligent agency and practical reasoning – that essentially requires some disposition to engage in the process of taking one's experiences at face value. If all this is right, then according to my proposals, the process of taking experiences at face value would also be an *a priori* primitively rational process.

*Is* it the case that unless one has some disposition to form beliefs by taking one's experiences at face value, one would be incapable of forming beliefs in contingent propositions about one's environment? It might seem that it is not the case. It may seem that there are other ways of forming such beliefs. For example, perhaps one could form beliefs *about* one's experiences by means of introspection, and then form beliefs about one's environment by means of an inference to the best explanation of one's experiences?

On reflection, it is not clear that it really would be possible to form beliefs in contingent propositions about one's environment in this alternative way if one had absolutely no disposition to take one's experiences at face value. In particular, it is not clear that one could form introspective beliefs about one's current experiences if one had no disposition to take one's experiences at face value. One clearly cannot form such introspective beliefs unless one can identify the conceptual content (or the 'conceptualized upshot', as I put

[16] The claim that possession of the concept 'if' requires a disposition to accept *modus ponens* inferences has been criticized by Williamson (2007: ch. 4). I have tried to defend this claim against Williamson's objections elsewhere (Wedgwood 2007b). But for present purposes, I do not need actually to endorse this specific claim – only the more general claim that for each concept there is *some* cognitive process that one must have some disposition to engage in if one is to possess that concept.

it in section 9.1) of those experiences. But it is not clear that one's experiences would *have* any conceptual content at all, unless one were at least *disposed* to form beliefs directly on the basis of one's experiences. (Indeed, on some views of the matter, the conceptual content of an experience precisely *consists* in the proposition that the thinker is disposed to come to believe in direct response to that experience.) So we may tentatively conclude that a disposition to engage in the process of taking one's experience at face value is necessary if one is to be capable of having beliefs in contingent propositions about one's environment – and so of being an intelligent agent – at all. So a disposition to engage in this process is a necessary condition of this basic cognitive capacity. If that is right, then according to my proposal, this process would count as an *a priori* primitively rational process.[17]

In this section, I have proposed an explanation of why the primitively rational processes count as *a priori* – they are processes that our basic cognitive capacities 'supply out of themselves' in this way – and I have provided some reasons for thinking that we can give the very same explanation of why the process of taking our sensory experience at face value is *a priori*. But what does this proposal have to do with the problem that we confronted at the end of the previous section? How does this proposal help us to see how it is possible for us to engage in this process precisely *because* of its connection with the truth?

## 9.5 A SOLUTION TO THE PROBLEM

In the previous section, I proposed that the primitively rational cognitive processes are precisely those that we must have some disposition to engage in if we are to possess certain basic cognitive capacities (such as the capacity of possessing a particular concept, or the capacity for a certain type of attitude). However, there is an obvious objection that some philosophers might be tempted to raise against this proposal. Why could it not be the case that one of these cognitive capacities requires a disposition to engage in a thoroughly *irrational* belief-forming process?

In section 9.3, I proposed that the primitively rational belief-forming processes must at least have *some* essential connection to the truth (even if it is only a weak connection). So taken together my proposals imply

[17] A full defence of this proposal would of course have to make it plausible that *all* of the primitively rational processes – including reliance on memory, induction, and inference to the best explanation, among others – are processes that our basic cognitive capacities 'supply out of themselves' in this way. Unfortunately I cannot attempt to carry out this task here (although I hope to do so on another occasion).

that possessing these basic cognitive capacities cannot require having the disposition to engage in any belief-forming process unless that process has at least some such connection to the truth. But why could it not be the case that one of these basic capacities requires a disposition to engage in a *misleading* process that altogether *lacks* any such connection to the truth?

In fact, however, it seems, at least to me, to be most doubtful whether these capacities could essentially involve a disposition for such irrational or misleading processes. How can these capacities – our possession of concepts and our capacities for all of the various types of attitude – essentially require a disposition to engage in a belief-forming process that has no such connection with the truth? These capacities are cognitive powers, not cognitive foibles or liabilities.[18] Possessing the concept 'if' may essentially involve a disposition to accept certain truth-preserving inferences, such as *modus ponens*; but it is surely perfectly possible to possess this concept without having any disposition to accept fallacious inferences, such as 'affirming the consequent' or 'denying the antecedent'. In general, it seems plausible that these capacities cannot *essentially* require a disposition to engage in a belief-forming process unless that process has some appropriate connection with the truth.

We can dramatize this point by considering the idea of a perfectly rational being. It seems plausible to me that every concept that we possess could also be possessed by this perfectly rational being. (For example, this perfectly rational being would have to make use of these concepts in order to ascribe mental states to us and to diagnose the errors and confusions that mar our thinking.) I am also inclined to accept the following thesis: for every concept, there is a specific cognitive process that one must have some disposition to engage in if one is to possess the concept. But of course the perfectly rational being has no disposition to engage in any irrational processes. So the processes that are essentially built into these concepts cannot be irrational processes; they must be rational processes instead. Unless a belief-forming process has some essential connection to the truth, it cannot be a rational process. So the processes that are in this way built

[18] This point is in effect the central idea of my argument for the normativity of the intentional (Wedgwood 2007a: 169), which was originally suggested to me by reflecting on Gareth Evans's (1982: 331) insistence that understanding a linguistic expression is a species of knowledge and so cannot be based on false belief: 'Truth is seamless: there can be no truth which it requires acceptance of a falsehood to appreciate.' Some philosophers will object by claiming that there are 'essentially defective concepts' – concepts possession of which essentially requires irrational dispositions of some kind. However, as I have argued elsewhere (Wedgwood 2007b), there are in fact compelling reasons to deny the existence of such 'defective concepts'.

into our possession of concepts must be rational processes that have an essential connection to the truth of this kind.

At the end of section 9.3, I argued that the mere existence of an essential connection between a belief-forming process and the truth was not enough to make the process rational; it is also necessary that one should be able to engage in this process precisely *because* of its connection with the truth. In the case of the *non*-primitively or *derivatively* rational belief-forming processes, one can engage in them because one has *independent justification* for *believing* them to have an appropriate connection with the truth – that is, because there is some *other* rational process that can lead one to the belief that these derivatively rational belief-forming processes have an appropriate connection to the truth. But in the case of the *primitively* rational processes, one can engage in these processes simply because one must have a disposition to do so, purely in virtue of possessing certain concepts or being capable of certain types of attitude. Since it is necessary for any disposition that is built into one's possession of a concept or one's capacity for some type of attitude in this way to have some such connection to the truth, one can also engage in these processes *precisely because* of their connection to the truth – even if one does not engage in these processes because one has any process-independent justification for believing them to have such a connection to the truth.

According to my proposals, then, these processes are built into the very nature of our capacities for the concepts that we possess, or for the types of attitude that we are capable of. This explains why they are *a priori*, as I explained in the previous section; and it also explains why it is possible for us to engage in these processes precisely because of their connection with the truth. In this way, my proposals explain why these processes have these crucial features that any primitively rational process would have to have.[19]

If I am also right that we must have a disposition to engage in the process of taking our experiences at face value if we are to have the capacities of intelligent agency and practical reasoning, then we have all the materials that we need to solve the problem about how this basic belief-forming process can be rational. Just like all other primitively rational processes, this process is built into the very nature of our fundamental cognitive capacities, and is therefore an *a priori* process, which it is possible for us

[19] This is a slightly different account from the one that I proposed in earlier work (Wedgwood 1999). There I claimed, in effect, that the belief-forming processes that are built into our concepts and capacities for attitude-types are primitively rational because they form the 'only non-arbitrary starting point'. It is not that that claim is wrong, but I am no longer confident that I was entitled to make that claim on the basis of the arguments that I offered at that time.

to engage in precisely because of its essential connection with the truth. This should help to dispel any sense of puzzlement that we might have had about how it can be that it is a genuinely rational belief-forming process.[20]

[20] Earlier versions of this chapter were presented at a conference at the University of Edinburgh, at a conference that was organized by New York University in Florence, and (in German) at the University of Göttingen. I am grateful to all those audiences, to my commentators at those conferences – Alan Millar in Edinburgh, and Ram Neta and Stephen Schiffer in Florence – and also to Cian Dorr, Lizzie Fricker, Maria Lasonen-Aarnio, Joshua Schechter, Nicholas Shea, David Wallace, Timothy Williamson, and an anonymous referee, for helpful comments. The first draft of this paper was completed while I was a Visiting Fellow at the Institute for Advanced Study at the Hebrew University of Jerusalem, and the final draft while I held a Research Fellowship from the Leverhulme Trust; I thank both institutions for their generous support.

CHAPTER 10

# *What does it take to "have" a reason?*

*Mark Schroeder*

If I believe, for no good reason, that P and I infer (correctly) from this that Q, I don't think we want to say that I "have" P as evidence for Q. Only things that I believe (or could believe) rationally, or perhaps, with justification, count as part of the evidence that I have. It seems to me that this is a good reason to include an epistemic acceptability constraint on evidence possessed.[1]

It is a truism that adopting an unjustified belief does not put you in a better evidential position with respect to believing its consequences. This truism has led many philosophers to assume that there must, at a minimum, be a justification condition (or perhaps a knowledge condition) on what it takes to count as having evidence. This is the best (or only) possible explanation of the truism, these philosophers have believed. This chapter explores an alternative explanation for the truism. According to the alternative explanation, unjustified beliefs do not put you in a better evidential position with respect to believing their consequences because evidence you have in virtue of having an unjustified belief is guaranteed to be defeated. Since lack of justification for a belief guarantees its defeat, I will suggest, we don't need to postulate a special justification condition (much less a knowledge condition) on what it takes to count as having evidence.

Why is this important? Because the assumption that there must be a justification condition (or a knowledge condition) on what it takes to count as having evidence places a *high bar* on what it takes to have evidence – such a high bar that it is difficult to see how it could be met in the case of basic, perceptually justified beliefs. As a result, the high bar set by this condition plays a fundamental role, I will claim, in central features of a core dialectic from the epistemology of basic perceptual belief. Given the degree to which this dialectic is shaped by the assumption of the high bar on what it takes to have evidence, I believe that it is worth taking seriously

[1] Feldman (1988b: 227).

what an alternative view might look like. My aim in this chapter is not to defend this alternative picture or its virtues in detail; it is simply to make room for the alternative.

## 10.1 CLARIFICATIONS: HAVING EVIDENCE

If Max is smiling, that is evidence that he is happy. So if you see Max, and he appears to be smiling, then you have some evidence that he is happy. Similarly, if you find out by testimony that Max is smiling, you have some evidence that he is happy. But if Max is a thousand miles away from you and you have no access to whether he is smiling or not, then even though the fact that he is smiling is evidence that he is happy, it is not evidence that you have. This difference between evidence that you have and evidence that you don't is crucially important. For example, only evidence that you have affects what it is rational for you to believe. If you see Max, and he appears to you to be smiling, that affects how rational it would be for you to believe that Max is happy. But if Max is thousands of miles from you and you have no access to whether he is smiling, the fact that he is smiling makes no difference to how rational it would be for you to believe that he is happy. Similarly, your beliefs can be *based* only on evidence that you have; if Max is thousands of miles away and you have no access to whether he is smiling, the fact that he is smiling can't be the evidence *for which* or *on the basis of which* you believe that he is happy. Whereas if you see Max, and he visually appears to be smiling, it then becomes possible for this evidence to be that for which you believe that he is happy. Because only evidence that you have affects what it is rational for you to believe, and because you can believe on the basis only of evidence that you have, the notion of evidence that you have is very important in epistemology.

The same considerations arise for the case of reasons for action. If your glass is filled with petrol, then that is a reason for you not to drink from it. So if you realize that your glass is filled with petrol, then you have a reason not to drink from it. But if the petrol in your glass was cleverly disguised to smell like a gin and tonic and it was handed to you by your usual bartender, then even though the fact that your glass is filled with petrol is a reason for you not to drink from it, it isn't a reason that you have. The difference between reasons that you have and reasons that you don't is crucially important – only reasons that you have affect what it is rational for you to do, and only reasons that you have can be the reasons *for which* or *on the basis of which* you act. Reasons that you don't have affect the advisability of actions, but they don't affect their rationality. The best

hypothesis I know for why what looks like a precisely analogous distinction arises both for evidence and for reasons for action, is that evidence is a kind (perhaps the only kind) of reason for *belief*. I will assume so in what follows.

Talk about evidence that you "have" makes it sound like when you have some evidence, what happens is that there is something which is, independently, *evidence*, and which, moreover, you *have* – which is in your possession or grasp in some way. This is the same idea that I have elsewhere called the Factoring Account, because it factors "having evidence" into two parts – "having," and "evidence." However, I don't actually think that the Factoring Account is true (see Schroeder 2008). To begin to get suspicious about whether the Factoring Account should be obvious or not, compare talk about having evidence to talk about having fathers. Todd has a father; his father is Larry. But Larry is not, independently, a father, whom Todd happens to possess. What makes Larry the father Todd has is simply that he stands to Todd in the *father-of* relation. So certainly not all talk about "having" something is talk about possession. Moreover, the Factoring Account has a strong consequence: if Max is not really smiling, then it's not really evidence that he is happy, that he is smiling. That is, talk about what is evidence *simpliciter* is naturally understood as *factive*. So if you can have evidence only if it is really evidence, then you never have evidence, unless what you believe or perceive to be the case is true. This is a big conclusion to pull out of such a little hat.

In opposition to the Factoring Account is the Two Relations Account, which I favor instead. According to the Two Relations Account, there are simply two relations that we need to be careful about – the *objective* evidence relation, and the *subjective* evidence relation. When I started off by noting that if Max is smiling, that is evidence that he is happy, what I said was true about evidence in the objective sense, and when I spoke about what evidence you *have* that Max is happy, what I said was true about evidence in the subjective sense. Objective evidence must be true, but subjective evidence need not be true; subjective evidence must be in some sense yet to be elucidated in the subject's "grasp," but objective evidence need not be in the subject's grasp. Moreover, according to the Two Relations Account, these two relations are not unconnected; on the contrary, they are related by the twin approximate facts that objective evidence that is in the subject's grasp (in the appropriate sense) is subjective evidence, and that subjective evidence that is true is objective evidence.

Unlike the Factoring Account, which treats subjective evidence as a *restriction* on objective evidence, the Two Relations Account allows that

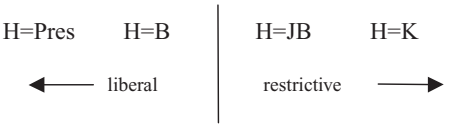
subjective evidence need not be true. But both accounts appeal at a crucial juncture to a kind of grasp or possession relation. The Factoring Account appeals to this possession relation when it claims that subjective evidence is (objective) evidence that you *have*. The Two Relations Account appeals to this possession relation when it claims that subjective evidence must be in the subject's "grasp." It is the nature of this possession relation that I will be interested in, in this chapter. The truism that adopting an unjustified belief does not put you in a better evidential position with respect to believing its consequences leads most philosophers to conclude that bare belief is insufficient for the *having* relation involved in subjective evidence. These philosophers adopt the view that, at least with respect to inferentially justified beliefs, your evidence is restricted to what you *justifiably* believe, or even to what you *know*. It is this conclusion that I want to illustrate how to resist. What I will be trying to show is that since any evidence you have in virtue of your unjustified beliefs is guaranteed to be defeated, it does not matter (at least for the purposes of accommodating the truism) whether we count their contents as part of what you "have" in the relevant sense, or not. Consequently we can take whichever view leads to the best overall theory.

## 10.2 THE SETUP: TWO ARGUMENTS FOR A HIGH BAR, AND TWO READINGS OF THE TRUISM

As I indicated at the beginning, many philosophers assume that we need to place a substantial restriction or "high bar" on this *having* relation. In order to get a sense for how this high bar works, we can introduce some simple notation for thinking about different views about the having relation as ranging from very liberal to very restrictive, as in the diagram below. For ease of notation, I will use "H=K" as an abbreviation for the thesis that the having relation is knowledge.[2] Similarly, I will use "H=JB" as an abbreviation for the thesis that the having relation is justified belief, and "H=B" will denote the thesis that the having relation is simply belief. The thesis for which I will be trying to make room, if not to argue for outright, is H=Pres – that the having relation is that of having a *presentational* attitude toward a proposition. By *presentational* attitudes, I mean attitudes which present their content to their subject as being true, which I understand to include both belief and perceptual experience

[2] Williamson (2000) uses "E=K" as his name for a similar thesis.

(both veridical and non-veridical), but not (for example) desire, wonder, supposition, or assumption.[3]




There are two ways of understanding the motivation for postulating the high bar on the having relation: one direct, and one indirect. The indirect argument begins with our truism: that adopting an unjustified belief does not put you in a better evidential position with respect to its consequences. It then argues, by way of some sort of auxiliary premise, such as that you have evidence for *p* by being in a state of mind S only if being in S puts you in a better evidential position with respect to *p*, that since unjustified beliefs do not put you in a better evidential position with respect to their consequences, they are not ways of having evidence for their consequences (similar points go for propositions which would be supported, but are not outright consequences, of the contents of these beliefs). This argument has one impeccable premise – the truism with which we began. But it is indirect, because it relies on an auxiliary premise connecting having evidence with improving one's evidential position.

But there is also a *direct* argument for the high bar on the having relation. It proceeds by pumping the intuition that someone who has no other reason to believe *p* does not have any evidence for *p*, simply in virtue of unjustifiedly believing something which has *p* as a consequence. This argument skips the auxiliary premise, but the intuition on which it relies is correspondingly more controversial. In section 10.3 I'll respond to the indirect argument by offering an alternative explanation for the truism – an explanation which incidentally sheds some light on why the auxiliary premise is false. Then in section 10.4 I'll take up the direct argument, and argue, as I have elsewhere, that negative intuitions about reasons or evidence, like those it relies on, are systematically misleading under the

[3] For one attempt to elucidate the way in which belief but not supposition represents its content as being true, see Velleman (2000). It is controversial itself whether there even is any such category as Pres, which is something that belief and perceptual experiences have in common, in contradistinction to these other states, much less whether I have adequately characterized this class with the slogan that they represent their contents as being true. Unfortunately, I won't be able to pursue such questions here, but the work I need Pres to do will become clearer as the chapter progresses.

same circumstances which make the indirect argument's auxiliary premise false – and in fact, in a wider range of circumstances. The direct argument, I'll therefore be arguing, is not independent from the indirect argument at all, and my earlier response to the indirect argument therefore suffices to set it aside as well.

Before I get to either of these arguments, however, we require yet one more important clarification as to what, exactly, is a truism, about our truism. Epistemologists importantly distinguish between *having* a justification to believe *p*, and being justified *in* believing *p*. This is sometimes called the distinction between *propositional* and *doxastic* justification. The idea is that *p* is propositionally justified for Maria when she *has* a justification to believe it, even if she does not believe it, or believes it, but not for the reasons involved in that justification at all, but instead (coincidentally) believes it because she read it in her horoscope. Whereas *p* is doxastically justified for Maria only if Maria believes *p*, and does so on the basis of her justification to believe *p*.

Given this important distinction, a belief in *p* can be unjustified either in the sense that it lacks propositional justification, or in the sense that the subject does not believe it on the basis of that justification. And that means that if we assume that a "better evidential position" reflects some improvement in one of our two senses of "justification," our truism is four ways ambiguous:

| | |
|---|---|
| PP | When *p* lacks propositional justification, it does not contribute to the agent's propositional justification for its consequences. |
| PD | When *p* lacks propositional justification, it does not contribute to the agent's doxastic justification for its consequences. |
| DP | When *p* lacks doxastic justification, it does not contribute to the agent's propositional justification for its consequences. |
| DD | When *p* lacks doxastic justification, it does not contribute to the agent's doxastic justification for its consequences. |

Of these four, I believe that only three are truisms. If Maria has an adequate propositional justification to believe *p*, but her belief in *p* lacks doxastic justification only because she believes it for the wrong reasons, I believe that she can still be propositionally justified in believing *p*'s consequences. Her propositional justification for believing *p* will also justify her in believing *p*'s consequences, on this view, though in accordance with DD, she will

not be doxastically justified in believing those consequences on the basis of *p* unless she believes *p* for the right reasons.

Moreover, though I take all three of PP, PD, and DD to be truisms, I think that of these, PP is the most basic and most important. DD can be explained on the basis of the assumption that to be doxastically justified, there must be no gaps in the doxastic justification of the premises on the basis of which it is held. This is an assumption specifically about doxastic justification, and needn't have consequences for the notion of having evidence – in fact, if we held that DD is true because when *p* lacks doxastic justification, the agent does not have *p* as evidence for *q*, we would intuitively predict that DP should also be true – which, as I have just suggested, it is not. Similarly, since doxastic justification implies propositional justification, PP's truth will suffice to explain the truth of PD. Consequently, in section 10.3 I am going to focus on an argument which seeks to predict and explain the truth of PP – the most basic and important of the readings of our truism. It is this reading which I believe plays the central role in the indirect argument for a high bar on the having relation, so it is this reading which I will try to explain.

## 10.3 THE ALTERNATIVE EXPLANATION: LACK OF JUSTIFICATION GUARANTEES DEFEAT

In this section I offer my alternative explanation of PP – the truism that when a belief lacks propositional justification, it does not contribute to the propositional justification for its consequences. Actually, I'll argue for something stronger: what I call *Lack of Justification Guarantees Defeat*:

| | |
|---|---|
| LJGD | When an agent's belief in *p* lacks propositional justification, it does not contribute to her propositional justification to believe *q*. |

Since LJGD is true for all *p* and *q*, it is not just that unjustified beliefs do not contribute to the propositional justification for their consequences; they do not contribute to the justification for anything at all.

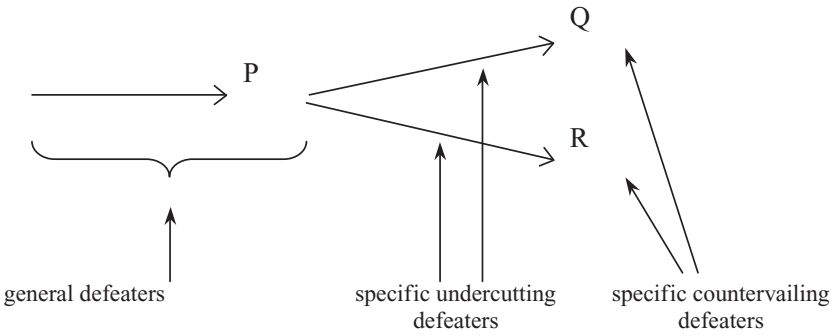
I give this thesis its name – *Lack of Justification Guarantees Defeat* – in order to highlight a feature of its explanatory structure. Rather than explaining our truism by restricting the *having* relation, my proposal is to explain the truism by appeal to general principles about evidence being

*defeated*. Epistemologists importantly distinguish between two kinds of "defeat"; a piece of evidence has a *countervailing* defeater when there is better contrary evidence – for example, though the fact that Max told you that he would come is evidence that he would come, the fact that you can see with your own eyes that he did not come is even better evidence to the contrary. In this example, the fact that you can see with your own eyes that Max did not come is a countervailing defeater for the evidence that he would come. A piece of evidence thus suffers from countervailing defeat when it "loses" to better evidence on the other side.

In contrast, a piece of evidence has an *undercutting* defeater when there is some further consideration which mitigates its force – perhaps so much so, that it does not carry any weight at all. For example, the fact that Max is often deceitful is an undercutting defeater for the fact that he said that he would come, as evidence that he would come. If Max is often deceitful, then what he says is simply less good evidence about what is true, than it otherwise would be. A piece of evidence can be undercut without being countervailed; Max's word may still be all that you have to go by, and if he is often *not* deceitful, then it may still count for something. But when it is undercut, it counts for less, and maybe for nothing at all. (In evaluating whether a piece of evidence is undercut or countervailed, keep in mind that sometimes one and the same consideration does *both* – for example, if Max is a reliable liar, then not only does what he says not count in favor of its truth, what he says, together with the fact that he is a liar, counts *against* its truth. In such a case there is an interaction between undercutting and countervailing.)

Ordinary cases of undercutting defeaters are cases of what I call *specific* defeat. In these cases, $d$ defeats $p$ as a reason to believe $q$, but does not defeat $p$ as a reason to believe other things. For example, that Tweety is a bird is evidence that Tweety flies. It is also evidence that Tweety has feathers. That Tweety nests in Antarctica is an undercutting defeater for the former piece of evidence, but not for the latter. It defeats the evidence provided by the proposition that Tweety is a bird for the conclusion that Tweety flies, but it does not defeat the evidence provided by the proposition that Tweety is a bird for the conclusion that Tweety has feathers. Such defeaters I'll call *specific*, because their defeat is specific to particular conclusions that we might draw on the basis of some particular piece of evidence.

In contrast, the kind of defeat postulated by LJGD is what I'll call *general* defeat. General defeaters don't just defeat one or another particular conclusion that we might draw on the basis of a piece of evidence; they defeat any such conclusion. A picture may help to clarify the distinction:




In the picture, the "open" ended arrows (⟶) represent the direction of reasons to believe P, to believe Q, and to believe R, and the "closed" ended arrows (→) represent different kinds of defeater. As illustrated, defeaters may be *countervailing* or *undercutting*. Countervailing defeaters attack the conclusion directly; they are reasons to *not* believe something – usually, reasons to believe its negation.[4] Undercutting defeaters, in contrast, do not attack the conclusion directly, but rather attack the reasons which support it. Specific undercutting defeaters attack these reasons "downstream" from the proposition which constitutes the reason, and as a result their effects are specific to particular conclusions. Whereas general defeaters, in contrast, attack these reasons "upstream" from the proposition which constitutes the reason, which means that they may undermine everything that P is a reason for.

You might think that general defeaters are too different from specific defeaters to really count as a kind of defeat. You might think that their very generality makes it more natural to describe considerations which have general defeaters as not evidence at all, rather than as evidence that has been defeated. I'm not so much concerned about how it is natural to describe them, however, as with the underlying explanation of the phenomenon. After offering the following argument for the thesis that beliefs which lack propositional justification *are*, in fact, guaranteed to have general defeaters, I'll therefore go back and explain why I think this argument effectively explains these general defeaters by the same general principles that suffice to explain how specific defeaters work, as well.

[4] Although see Schroeder (forthcoming); some reasons not to believe $p$ might be reasons to withhold, rather than reasons to believe $\sim p$.

Finally, we get to the argument for LJGD. It is quite simple:

| | |
|---|---|
| P1 | When an agent's belief in *p* lacks propositional justification, she has insufficient reason to believe *p*. |
| P2 | When an agent has insufficient reason to believe *p*, she has conclusive reason to not believe *p*. |
| P3 | When an agent has conclusive reason not to believe *p*, she has conclusive reason not to take *p* into account in her reasoning about *q*. |
| P4 | When an agent has conclusive reason not to take *p* into account in her reasoning about *q*, it is irrational for her to take *p* into account in her reasoning about *q*. |
| P5 | An agent's belief in *p* contributes to her propositional justification to believe *q* only if it is rational for her to take *p* into account in her reasoning about *q*. |
| LJGD | When an agent's belief in *p* lacks propositional justification, it does not contribute to her propositional justification to believe *q*. |

This argument has five premises. Of these, I take P1 to be a highly compelling principle about propositional justification and P2 to follow from the definitions of "sufficient" and "conclusive." I take P3 to be independently highly compelling; we should not rely in our reasoning on things that we don't believe to be the case. Consequently, we should not rely in our reasoning on things which we *should* not believe to be the case. P4 I hope is close to a platitude, so long as we keep in mind that we are throughout talking about *subjective* reasons – reasons the agent "has" – as opposed to objective reasons, which might be conclusive even though the agent is unaware of them and could be rational in flouting. P4 is certainly consistent with the idea that there is much more to rationality than complying with conclusive subjective reasons. And finally, P5 looks to me like it is at least highly plausible – certainly the normal way in which beliefs contribute to the justification for other beliefs is because it is rational to take their contents into account in reasoning. If the belief that *p* does not contribute to the justification for believing *q* in this way, then it is not clear how it does contribute to it.

So together, P1–P5 provide an argument from highly plausible premises for LJGD, the thesis that lack of propositional justification guarantees defeat. This argument shows how we can rule out unjustified beliefs helping to justify other beliefs, without the need to postulate a special justification condition on the "having" relation involved in having evidence. According to this argument, *even if* you can "have" evidence by having a propositionally unjustified belief, it is guaranteed to be defeated anyway.

It is natural to question whether the argument that I have just provided gives us an independent *explanation* of our truism, or just predicts that our truism is, in fact, true – a fact which we should explain by positing a high bar on the having relation. I believe that the argument I have just given does, in fact, provide an independent explanation of the truism, because I believe that it illustrates that the very same general principles which explain the behavior of undercutting defeaters in general, predict and explain why there would be general undercutting defeaters in these cases. This is because in general, undercutting defeaters work by making it rational to take the proposition which they defeat less into account in reasoning toward their conclusion. To see how this works, let's go back to one of our examples. The fact that Tweety lives in Antarctica is an undercutting defeater for the fact that Tweety is a bird, as evidence that Tweety flies. If you are aware that Tweety lives in Antarctica, this also makes it rational (given other common background knowledge) for you to place very little weight on the fact that Tweety is a bird, in reasoning about whether Tweety flies.

The structure of the argument that I have just given aims to show that the very same thing is going on when you have an unjustified belief – namely, that you have some reason which makes it rational to place very little – even no – weight on the content of your unjustified belief in your reasoning. Which reason is this? It is the sufficient reason not to believe the proposition in the first place – which in most cases is largely constituted by the evidence that you have against it, which made it an unjustified thing to believe in the first place. So the idea of the argument is that the very same principles which explain how specific undercutting defeaters work – by making it rational to pay the evidence they undercut less attention in reasoning – also serve to explain how general undercutting defeaters work – by making it rational to pay the evidence they undercut less attention in reasoning. Consequently, the idea is that our truism is simply a consequence of these general features of the behavior of undercutting defeaters. It is therefore not something of which we need to provide a special explanation, by postulating special high bars on what it takes to have evidence.

## 10.4 Negative intuitions about reasons/evidence and the direct argument

So far, I take myself to have responded to the indirect argument for the high bar on having evidence. I now turn to the direct argument, which trades on eliciting the intuition that someone (who has no *other* evidence for $q$) who believes $p$ only unjustifiedly *has no* reason to believe $q$ – or, put in terms of

evidence, that she *has no* evidence for *q*. My response constitutes an error theory about negative intuitions about reasons and evidence under certain circumstances which are present in this case. In particular, I'll show that when reasons or evidence are undercut or otherwise very low in weight – at least in contrast to the competing reasons on the other side – it is quite easy to elicit the intuition that someone does not have a reason to do something, even when she really does have such a reason, albeit one of low weight. Here I'll make the argument directly in terms of intuitions about reasons, rather than in terms of intuitions about evidence, though I believe that nearly identical considerations apply. (Readers who can anticipate how this argument will go may profitably skip directly to section 10.5.)

The reason why the direct argument does not do any more work than the indirect argument, is that our intuitive judgments about what there is no reason to do or to believe are strongly influenced by the weight of reasons, and by how much of an effect those reasons have the potential to have on what we ought to do, in the context of the other reasons that there are. When there is a reason to believe something, but it is of low enough weight that it obviously makes no difference, given the other reasons in play in the situation, to what it is rational to believe, pragmatic considerations successfully predict that we will incorrectly find it intuitively compelling that there is in fact no reason to believe.[5]

That pragmatic factors should affect our intuitive judgments about what there is no reason to do, is predicted by general Gricean considerations, in conjunction with the plain fact that reasons differ widely in their weight. Think of reasons like the considerations which go into the "pros" and "cons" columns that God might make in an exhaustive evaluation of what you should do. Since God's capacities are unlimited, He may be able to think of an awful lot of aspects of the situation which merit inclusion as "pros" or "cons." We, however, are mere finite creatures – with severe cognitive processing limitations, at that. We don't have time or energy for the kind of exhaustive search of which God is capable. We only have time to look at the top of God's lists. So long as we know that the reasons in favor that we have been able to survey are good *enough* to outweigh any comers, it is not that important to find out what those comers are – for we know in advance that they are insufficient anyway.

If any picture even roughly like this one is correct, and we are mostly interested in reasons in order to figure out what we ought to do or to

[5] Here I am summarizing arguments I have developed elsewhere in greater detail. For further discussion of the points which follow, see chapter 5 of Schroeder (2007).

believe, then reasons that are too far down God's list of pros and cons to be ones that it makes sense for us to bother with will be ones that it is not worth talking about. So this predicts, first, that pointing out or focusing on reasons of negligible or trivial weight will be conversationally unhelpful, at best. In fact, we should predict that even reasons of non-negligible weight will be not worth talking about, provided that they are obviously insufficent. And so this yields a second prediction: that pointing out or focusing on reasons that are obviously insufficient will be conversationally unhelpful, at best.

In fact, we can do more than make these two predictions. We can check to see that they are in fact borne out in cases. To see how, consider the case of Tom Grabit, familiar from the literature on undercutting defeat.[6] In case 1, you see Tom come out of the library, pull a book out from under his shirt, cackle gleefully, and scurry off. You think, "Hey – Tom just stole a book!" Intuitively, in this case you have some evidence for your conclusion – some evidence that Tom just stole a book. But now consider a twist: case 2 is the same as case 1, except that you know that Tom has an identical twin, Tim, from who you cannot visually distinguish him. Intuitively, in this case you don't have a reason for your conclusion after all – you don't have evidence that Tom stole a book. In case you lack this intuition, it is important to note that this is one of the paradigm cases of undercutting defeat in the literature, and undercutting defeaters are often *defined* as cases in which, because of the defeater, you don't have any reason for your conclusion after all.[7] So even if you don't share this intuitive judgment about case 2, it is one that has been widely shared.

But now compare case 3, which is similar, except that now you know that Tom and Tim have a third identical sibling, Tam. I submit that in case 3, you have even less reason to believe that Tom stole a book than in case 2, and hence that you must have had some reason to believe that Tom stole a book in case 2. If this is unobvious, observe that it would take more additional evidence that Tom stole a book to make you justified in drawing the conclusion that Tom stole a book in case 3 than in case 2. If this is right, then the intuitive judgment about that case that has been shared by so many epistemologists over the years is mistaken. And that is precisely what was predicted by our pragmatic considerations: we predicted that reasons of relatively low weight would seem not to be reasons at all.

So the Grabit cases, because they are subject to an independent argument that our intuitive judgments are wrong, provide evidence that the

[6] See Lehrer and Paxson (1969). [7] Compare, for example, Pollock and Cruz (1999: 195–197).

pragmatic effects that we earlier predicted based on abstract considerations are genuinely in play, and really do affect the way that we think about philosophical examples. The idea that our intuitive judgments about what someone has reason to do might be affected by pragmatic factors is therefore not just an abstract possibility. It is a very real phenomenon that we need to be careful about in evaluating intuitive judgments about reasons or evidence. Consequently, it safer to go by our indirect evidence. If unjustified beliefs do not improve our evidential position for their consequences, then even if we do have reasons to believe their consequences, those reasons must be of little or no weight, and consequently will be ones about which we will have misleading negative intuitions. And this is precisely what the argument of the last section predicts. It predicts that any reasons we have by having unjustified beliefs will be utterly defeated – totally reduced in weight.

## 10.5 THE PAYOFF FOR EPISTEMOLOGY?

So far in this chapter, I have been concentrating my energies on rebutting two arguments for an apparently abstruse conclusion: that there must be a relatively high bar on the "having" relation that is involved in having evidence. The point of all of this hard work, I claim, is that it expands the options in epistemology in a fruitful way, which makes it easier to understand how basic perceptual beliefs could be based on evidence. This is because the assumption of a high bar on what it takes to have evidence – despite its overwhelming initial plausibility – has been complicating the dialectic about basic perceptual epistemology, and consequently the dialectic about foundationalism and coherentism, internalism and externalism, and rationalism and empiricism, for a very long time.

To see why, think about our high bar as at bottom a high bar on what it takes to have evidence in cases of *inferential* justification – cases in which one belief is based on evidence you have in virtue of having another belief. Inferentially justified beliefs can meet this bar, provided that they are based on other beliefs that are justified. So inferentially justified beliefs can be based on evidence. But the problem is whether *non*-inferentially justified, *basic* perceptual beliefs can also be based on evidence, on this view. And the problem essentially comes in two parts. The first part of the problem is that the high bar on having evidence for inferentially justified beliefs looks like it is too high to be met, in the case of basic perceptual beliefs, if those are justified only on the basis of perceptual experiences. So that means that if basic perceptual beliefs are to count as based on evidence at

all, then in order to keep the high bar on what it takes to have evidence for inferentially justified beliefs, we need to adopt a *mixed* view, with mixed standards about what it takes to have evidence – sometimes requiring a justification condition, and sometimes not. But the second part of the problem is that this kind of view looks *ad hoc*, unstable, and explanatorily unsatisfactory.

In the remainder of this section I'll rehearse this kind of dialectic in more detail; my goal will not be to definitively set aside any of the options that come up along the way, but only to emphasize that dropping the high bar on what it takes to have evidence relaxes the constraints on perceptual epistemology in an important way. It makes it much easier to see how to tell a unified explanatory story about how beliefs can be based on evidence, in perceptual and non-perceptual cases alike. Perhaps this view will suffer from its own problems, as well, but that should be a matter for theorizing to decide, not to be stipulated in advance by ruling it out of court.

To take the first side of the problem first, let's look at whether basic perceptual beliefs could be based on evidence, if the high bar has to be satisfied. So to address that, take first the thesis that H=JB, which was one kind of view which upheld the high bar. We are assuming that a belief is based on evidence only if it is based on evidence that you have. So if having evidence requires justifiedly believing it, it follows from this view that beliefs can be based on evidence only if they are based on other beliefs. That is, non-inferentially justified beliefs can't be based on evidence at all. But this is a highly counterintuitive result! It is one thing to say that some beliefs are justified without being justified on the basis of other beliefs. Any foundationalist accepts this. It is another thing to say that basic perceptual beliefs are not based on any evidence at all. This is a bizarre claim – beliefs formed directly on the basis of perceptual experiences are intuitively a *paradigm* of beliefs that are based on evidence, because perceptual experience is a paradigm method of gaining evidence about the world. Now, it is easy to see why a theory might lead some philosophers to endorse the conclusion that this is misleading, and basic perceptual beliefs are not, after all, based on any evidence at all. But it smells to me like the sort that we should shy away from without some clear motivation. So it's worth looking at how we can avoid this conclusion.

So far, the problem arises from the *B* part of H=JB, rather than the *J* part. But the high bar focuses on the *J* part – the idea that there must be – at least – a justification condition on having evidence. So one response *might* be to generalize. Maybe perceptual experiences, like beliefs, can be ways of having evidence – provided that *they* are justified. This leads to the

identification, H=JPres,[8] which preserves the justification as a minimal bar for having evidence, but is liberal about the kind of *attitude* that having a reason can consist in – it can consist either in belief or in perceptual experience.

The problem with this view is that perceptual experiences aren't the kind of thing that *can* be justified. Beliefs can be justified or unjustified because they fall under the reach of rational criticism, and they fall under the reach of rational criticism because they are states that we hold for reasons. Perceptual states are not things we enter into for reasons. True, we can choose whether to look in the direction of the stick in the water or to look away. But once we look in its direction, its looking bent to us is not something that is responsive to our reasons, it is not something that we do for reasons, it is not subject to rational criticism, and it is not subject to justification. It is "outside the space of reasons." So H=JPres is a failed venture.

One way of trying to get the benefits of H=JPres without its costs is to double down on the high bar by endorsing H=K. Maybe perceptual states do not themselves admit of justification, this view goes, but they *do* admit of constituting *knowledge*. Take any of our wide range of factive experiential verbs, like "sees that," "observed that," and "smelled that." Williamson (2000) tells us that each of these implies "knows that." If some perceptual states actually constitute knowledge, then by H=K, these perceptual states will constitute ways of having evidence on which beliefs can be based. Hence, non-inferentially justified beliefs can be based on evidence – not evidence provided by other beliefs, but evidence provided by perceptual states. On this view, all justified beliefs are based on evidence; some are based on evidence provided by other beliefs which constitute knowledge, and that is the inferential case, whereas others are based on evidence provided by perceptual states which constitute knowledge, and that is the basic case. But in both cases, on this view, all justified beliefs are based on evidence.

That all sounds very good (though there are important and not-to-be-glossed-over complications in the details[9]). But it has one very striking

[8] Here as before, think of "Pres" as just a placeholder for some genus of attitude which includes both belief and perceptual experience in contrast to wondering, assuming, supposing, and imagining.

[9] For example, according to Williamson (2000), knowledge entails belief. So on that view, the perceptual states that constitute knowledge would also constitute belief. Since they would not be based on further evidence, they would therefore be knowledge that was not based on any evidence. Williamson avoids the question of whether all knowledge is based on evidence, by observing that on his view, all knowledge is *supported* by the evidence – because if you know that $p$, then $p$ is among your evidence, and that supports $p$. This simply evades the question of whether someone's knowledge that $p$ can be *based* on evidence.

problematic feature. And that is that on this picture, there is a very important difference between veridical and non-veridical perceptual experiences' abilities to play a role in justification. This is because even if some perceptual experiences constitute knowledge, non-veridical perceptual experiences certainly don't – for knowledge requires truth. That means that this view predicts that internal duplicates with duplicate pasts in very similar environments can nevertheless differ not just in whether they know, but in whether their belief is justified, if for the very first time one sees something green in front of her while the other experiences a vivid hallucination as of something green in front of her, and both form the perceptual belief that there is something green in front of them. This is extremely unintuitive.[10] Like the thesis that some beliefs are justified without being based on any evidence itself, it is the kind of thesis one can imagine talking oneself into, if one has the right kind of theory. But again, it smells like the kind of conclusion that we should shy away from, if we can.

So unfortunately the high bar on having evidence is not a bar that can be met in the full range of basic perceptual cases. At most it can be met in the case of veridical perceptual experiences, but it would be best to avoid needing to distinguish between the justification provided by veridical and by non-veridical perceptual experiences. So if we want to also both avoid the conclusion that some beliefs are justified on the basis of no evidence whatsoever, and hang onto the high bar in the cases of inferentially justified belief appealed to by the truism, there is just one way to go: *mixed views* – views which propose high standards on having evidence in the cases of inferential justification, and low standards in the case of basic perceptual justification.

This is what a variety of epistemologists try to do. On views like these, there are *two ways* to "have" evidence. In order to have some kinds of evidence, you have to justifiedly believe it. For example, to have Max's smiling as evidence that Max is happy, you have to justifiedly believe that Max is smiling. But according to this view, for other kinds of evidence,

[10] Note that what I am claiming is unintuitive here is not just a feature of *internal duplicates*. For example, internal duplicates might have different pasts, or be differentially reliable, or be in different environments, one more conducive to truth belief formation than the other, and many philosophers believe that these conditions can affect whether they are justified or not. What I am pointing out here, is that two people might be internal duplicates, with identical pasts, equally reliable, in environments which are identical up to the very thing that each is looking at right now, which one veridically sees to be green and the other experiences a perceptual illusion as of being green. Surely in *this* case one is justified in her perceptual belief formed on the basis of this perceptual experience if the other is. But the H=K thesis only admits, at best, veridical perceptual experiences as among their evidence.

only something much weaker is required. In fact, according to many of these views, you can have some kinds of evidence just by virtue of its being true.[11] For example, to have your perceptual experience as of something green in front of you as evidence that there is something green in front of you, you need only to have a perceptual experience as of something green in front of you – you don't need to have any belief about this perceptual experience at all – let alone a justified one.

Now, I ultimately want to agree that *when* you have such a perceptual experience you have evidence for the belief that there is something green in front of you – even if you have no prior beliefs about your perceptual experience, let alone justified ones. So I think there is something descriptively accurate about this view – in contrast to views which hold that since "having" evidence requires having a prior belief, your belief that there is something green in front of you can be justified only if you first believe *that you perceive* as of something green being in front of you.[12] But there are nevertheless at least three serious problems with mixed views.

The first problem is that it is easy to provide an error theory for what could lead proponents of mixed views to hold that you have certain sorts of evidence simply in virtue of its being true. Certainly, given that your perceptual experiences are even moderately reliable, the fact that you are having a perceptual experience as of something green in front of you is evidence in the *objective* sense that there is something green in front of you. It is objective evidence even for people who are not you and do not know about your perceptual experience. This makes the move of calling this perceptual experience part of your evidence even though you don't satisfy the epistemic access ("having") condition for subjective evidence very suspicious – after all, it is *true* that it is part of your evidence in *one* sense of "evidence." It is part of the objective evidence. What is at stake is

[11] Compare, for example, Pryor (2005) and Conee and Feldman (2004).

[12] The idea that having evidence involves a *having* relation makes it easy to see why early twentieth-century empiricists were so wedded to the idea that the epistemologically basic beliefs were existential beliefs about perceptual experiences, and hence that you need to have beliefs *about* your perceptual experiences, in order for them to play a role in justifying your other beliefs. For example, compare Lewis (1946) and Chisholm (1957). But not only was this view descriptively inaccurate, incorrectly classifying subjects who lacked beliefs about their own perceptual experiences as being unable to have beliefs justified on the basis of perception, they ran into troubles both explaining how beliefs about perceptual experiences could be justified in the first place, and how beliefs about perceptual experiences could provide as good evidence about the external world as perceptual experience seems to provide. Compare Pollock (1974).

whether it is part of your *subjective* evidence. But in all of the non-basic cases, even this view admits that some further condition is required.

The second problem for mixed views like this one is larger. If the only evidence you have that there is something green in front of you is that it perceptually appears to you as if there is something green in front of you, then there is a serious question of why this evidence is so good. It is a big jump, after all, to get from "it perceptually appears to me as if *p*" to "*p*." In fact, it is precisely this jump that plays a central role in some important kinds of skeptical argument, which point out that it is consistent with all of your evidence that you are a brain in a vat. Whatever the correct response to other kinds of skeptical arguments, whether this particular skeptical argument gets going depends on whether your evidence includes only facts about your own perceptual state, or includes information about the world. Even putting such skeptical arguments aside, it is hard to see what makes perception an especially good source of evidence about the external world, if the only evidence it gives us directly is about our own mind. Even if you can have the fact that it perceptually seems to you that *p* as evidence just by virtue of its being true – and thus without needing to stand in any further positive epistemic relation to it, such as believing or perceiving it to be the case – it is at least not obvious why this is particularly good evidence to have.

I don't point this out in order to argue that this is an unsolvable problem for views according to which your basic perceptual evidence about the external world consists in facts about your own perceptual psychology. Certainly many philosophers have tried to explain why this is nevertheless good evidence about the external world. My point is rather that it is a problem that has a particular kind of *source*, which is illustrated by how easily the contrasting views avoid this pitfall. According to the contrasting picture, in its perceptually seeming to you as if *p*, you have *p* as evidence that *p* (and count as "having" it because you stand in a positive epistemic relation to it, to boot). On this sort of picture, it is easy to see why that is a good situation to be in, and hence why perception is a privileged source of evidence about the external world. This is because in its perceptually seeming to you that *p*, you come to have *p* as a reason to believe *p* – and no other evidence for *p* could be better.

Finally, the third problem for this kind of view is I think quite serious. If there are two such different ways of "having" evidence – such that one way requires that you justifiably believe the evidence, and the other requires that the evidence be true – then there will be two very different ways of

*basing* a belief on evidence. But that calls into question what, exactly, this sort of theorist has gained by saying that all justified beliefs are based on evidence. They are based on evidence in such different ways that it stretches the imagination to understand how there is even a unified sense of "based on the evidence" in which both are based on the evidence.

Of course, as theorists we can *use* the words "based on the evidence" however we like. We can even stipulate that no matter how beliefs are justified, we will say that they are "based on the evidence." But that falls far short of it turning out that there is such a thing as being based on the evidence which happens with inferentially justified beliefs, and which also happens – the very same thing – with basic perceptual beliefs. Again, the idea that there is no single relation of being based on the evidence, such that, holding that sense of "based on evidence" fixed, all justified beliefs are based on evidence, is the kind of thing that a theory can lead one to accept. But my sense of smell is getting to me again, and it tells me that this is the kind of thesis that we should avoid, if there is any easy way to do so.

It is crucially important not to underestimate the centrality of this problem in epistemology. Repeatedly in the history of twentieth-century epistemology, philosophers have returned to the idea that there are simply different things going on in the case of basic perceptual justification and in the case of inferential justification – that one philosophical story is required for one, and another quite different story for the other. Mixed views fall squarely within this tradition. And the prevalence of mixed views like this has been one of the strongest original motivations for both coherentism and externalism. Coherentists can offer a unified account of justification, by applying what goes for the inferential case to what are apparently basic perceptual cases.[13] And since the beginning of externalism in epistemology, externalists have argued that foundationalists have to accept externalist explanations of basic perceptual justification anyway, and that once you take that on board, they are merely extending what foundationalists accept anyway to the inferential cases.[14] It's hard to complain about causal or pure reliabilist theories of knowledge, when your own story about basic perceptual justification is distinguishable from them only by being less unified and less explicit.[15] So for these reasons I take it to be quite a serious charge that the view we've been considering does not provide a unified picture of having evidence and basing beliefs on evidence. It would be nice to be able to do better.

[13] See especially BonJour (1985). [14] See especially Armstrong (1973).
[15] On this instability of mixed views, compare McDowell (1994b).

## 10.6 WRAP-UP

Just to rehearse: in the last two sections we have encountered a familiar dialectic. The arguments I've sought to rebut in this chapter would have us place a high bar on what it takes to have evidence for inferentially justified beliefs. That high bar cannot be met by basic perceptual beliefs, which are based on perceptual experiences, rather than on beliefs. Since perceptual experiences don't admit of justification, it doesn't help to liberalize H=JB to H=JPres. And since knowledge requires truth, the strategy of doubling down on H=K at best nets cases of veridical perceptual experiences.

So if the high bar applies across the board, then we are led to the conclusion that basic perceptual beliefs cannot be based on any evidence. In order to retain the high bar on having evidence for inferentially justified beliefs, but make it possible for basic perceptual beliefs to be based on evidence, some have advocated mixed views. But mixed views appear to conflate objective and subjective evidence, make it puzzling why the evidence provided by perceptual experiences is particularly good and open at least one door to the skeptic, and present a problematically disunified picture of justification that is vulnerable to attack from both sides.

In the context of this dialectic, I *do* think it is worth taking a closer look at the motivation for raising the bar on having evidence so high in the inferential case. By lowering that bar, we can endorse H=Pres, and thereby offer a single, unified account of what it is to have evidence that is satisfiable in both inferential and basic cases, without excluding non-veridical perceptual states. This unified account is very simple: it says that a belief is propositionally justified just in case it is supported by the balance of the undefeated evidence the agent has, and it understands having evidence as everywhere the same: a matter of bearing some presentational attitude (either belief or a perceptual state) toward a proposition.

Moreover, because it is a unified account,[16] we can better fend off the challenges from coherentists, on the right, and from pure externalists, on the left, who seek to generalize either our inferential case or our basic case to include all of justification. I also believe that such a unified account promises to be more deeply explanatory, and even to tell us not just *when* beliefs are justified, but what justification *is*. I happen to find this picture attractive because it avoids all of our earlier pitfalls. But all I have been

[16] Whether this is, in fact, a unified account depends on whether there is anything unified about the category, Pres, which includes both beliefs and perceptual states. I am optimistic on this score, but I haven't discharged this obligation here.

trying to argue for here, is that it is a serious alternative, rather than one that is ruled out of court by a truism that we all need to recognize. If something is wrong with it, that is for the theoretical balancing of costs and benefits to decide, not a snap judgment that it doesn't merit consideration, due to the truism.[17]

[17] Special thanks to Jake Ross, to conversations with Jim Pryor and Timothy Williamson, to an anonymous referee, and to Andrew Reisner and Asbjørn Steglich-Petersen.

CHAPTER 11

# *Knowledge and reasons for belief*

*Alan Millar*

## 11.1 THE PROBLEM

There are epistemological puzzles that take the form: how can we gain so much from what seems so little? Knowledge from testimony nicely illustrates this. In certain cases, which I shall call straightforward cases, we gain knowledge that something is so from a person's telling us that it is so, despite the fact that we have not, or not obviously, engaged in any reasoning bearing on the credibility of what we have been told. We accept what we have been told straight away on the say-so of the informant and thereby gain knowledge. What makes this problematic is that, given the rather special standing we take knowledge to be, it is puzzling how a person's say-so can be the means of acquiring knowledge. Arguably, part of what makes knowledge special is that it implicates justified belief. On the natural assumption that being justified in believing that *p* is a matter of having an adequate reason to believe that *p*, it can easily seem obscure that someone's telling us that *p* can ever be an adequate reason to believe that *p*.

Accepting that there is this difficulty does not commit us to accepting that knowledge admits of a reductive conceptual analysis in terms of justified belief and other conditions. If knowledge does not admit of such an analysis, as has been argued by Timothy Williamson (2000), it still might be that necessarily, one who knows that *p* is, in a way that is tied up with what it is to know that *p*, justified in believing that *p*. This latter view is also disputable. One might reject it on the grounds that while being justified seems to implicate reflective capacities – those involved in thinking about and weighing up reasons – animals that lack such reflective capacities have knowledge. For present purposes I think we can sidestep such matters. I am concerned with human knowledge and thus with the knowledge of agents who have reflective capacities and are capable of thinking about why they believe as they do. I take it as a working assumption that when such agents

know that something is so they are, at least normally, justified in believing it to be so on account of having an adequate reason to believe that it is so – a reason to which they will normally have access. I seek to make sense of the idea that, for such agents, knowledge at least normally goes along with being justified in believing.

A quick response to the challenge posed by testimony would be to say that the mere fact that a person tells us that *p* never is an adequate reason to believe that *p*. So it is right that being apprised of this fact is 'too little' to give us knowledge that *p*, but wrong to suppose that this poses a problem. On the contrary, following this line of thought, we can be justified in believing that *p* in view of other things we believe which, together with the fact that we have been told that *p*, provide us with adequate reason to believe that *p*. Non-testimonial cases provide the model. A plausible view with roots in Hume is that, for instance, we can come to know that the grass making up the lawn lacks nutrients on the basis of the fact that it is yellow. We can do so with the help of a generalization to the effect that when grass is yellow it lacks (or is very likely to lack) nutrients. Applied to testimonial cases this suggests that we need a generalization relating to the veracity of the testimonial utterances with which we are confronted. The needed generalization would be to the effect that when somebody of a certain sort tells us something and various further conditions are satisfied then what they tell us will be true or very likely to be true. The reason for which we are looking would be constituted by such a generalization along with a corresponding statement to the effect that this speaker is of the relevant sort and the further conditions are satisfied.[1] An account along these lines would be reductionist in that it assimilates the epistemology of knowledge or justified belief from testimony to a general epistemology of empirical evidence. Attractive though it seems, this approach is not free of difficulties. It is not a trivial matter to specify a generalization covering the informant's utterance on which we could plausibly be taken to rely. It is not that we are powerless to pick out factors that are relevant: type of person, manner of speaking, content of utterance, the immediate context of utterance (the situation in which it is made), the wider context of the utterance incorporating any shared histories of informant and recipient, and the culture or cultures to which they belong. The problem is to work determinate forms of these factors into a generalization that could do the business on a specific occasion.[2] There are really two problems here. One

[1] For a recent version, see Fumerton (2006).

[2] Fumerton (2006: 80) works with a schema ('says that *p* in conditions C') that does not bring out the range of factors and the elusiveness of their determinate forms.

is that it is far from clear that we routinely have a suitable generalization at our disposal. The other is that it is far from clear that even if we had we would routinely be justified in accepting it when we gain knowledge from testimony. No doubt we shall have had much experience that affects the level of confidence that we have in what we have been told. But it is a further matter whether this experience will even have led us to believe a determinate generalization that will do the job, far less have furnished us with adequate evidence for accepting it.

This is the kind of problem I wish to explore in this discussion. In the next section I consider perceptual knowledge. In section 11.3 I give an indication of how one might develop an account of testimony building on a non-standard account of knowledge from indicator phenomena. An instance of the latter would be knowing that it has rained because the streets and other outside surfaces are wet. From these discussions emerges a conception of the importance of recognitional abilities which in section 11.4 is developed to yield an account of reasons for belief in the cases under consideration and a general view of the connection between knowledge, justified belief and reasons. Section 11.5 explores what I call detached standing knowledge. I do not claim to establish my general view on these matters within the space available here, only to have made it plausible enough to merit further investigation.

## 11.2 PERCEPTUAL KNOWLEDGE

A problem structure analogous to that described in relation to testimony arises in cases of perceptual knowledge. Here too it can seem that we gain a lot from a little, and it is hard to see how the little can supply us with the required adequate reason. This is how things are liable to seem from a certain theoretical stance that has been widely accepted. On this stance, seeing that Bill has arrived is a matter of being in a certain psychological state – having an experience such that it looks to one as if Bill has arrived – plus the satisfaction of further conditions. The conditions are, roughly, that Bill has indeed arrived and one's experience is explicable in a way characteristic of normal perception by Bill's having arrived. The question arises as to how the experience can furnish us with a reason to believe that Bill has arrived. The issue concerns how we should think of the reason we have to believe that Bill has arrived in the situation in which we see that, and thereby know that, he has. Paradigm cases of believing for a reason are cases in which we believe one thing in view of something else, which we also believe. Thus I might believe that my wife has returned from work in

view of the fact that her car is in the driveway. This fact is at least part of my reason for believing that she has returned from work. But, on the face of it, my knowing that Bill has arrived, when I see him arriving, is not a matter of my believing that he has arrived in view of something else that I believe. It simply strikes me on seeing him that he has arrived. I do not appear to believe this because I believe something else that I treat as a reason for me to believe. This creates a pressure to suppose that at some stage we need the idea that experiences themselves can be justifiers.[3] The idea is not that *the fact* that one has certain experiences can provide a reason to think that something is so. That would still treat one's justification as provided by propositions one believes. Rather, the point is that one can be justified in believing certain things in virtue of having appropriate experiences, in the absence of any reason not to believe. On this account the experiences themselves are supposed to be justifiers.

It is one thing to feel compelled to accept this view and another to see how it can be true. We would need some account of how experiences can justify beliefs. An apparently promising account is available via a combination of two ideas: (i) that being justified in believing something is to be explained in terms of forming the belief in a competent way – specifically, in a way that manifests competence in deploying the relevant concepts; (ii) that basing beliefs on experiences of an appropriate type can be a manifestation of conceptual competence, analogous to basing a belief on other things one believes.[4] Being justified in believing that *p* on the basis of other things one believes requires that in forming the belief that *p* on the basis of those other beliefs one exploits one's mastery of the various concepts involved. Obviously it matters that the other beliefs should have an appropriate standing. (What this must be is debatable.) The analogy for believing that *p* when seeing that *p* is that the experiences should be of an appropriate sort and that in forming one's belief one exploits one's mastery of concepts implicated by believing that *p*. This approach has the merit of yielding a way of developing a familiar way of thinking about Dretske's (1970) zebra case and cases like it. Consider a case – the good case – in which I am looking at a zebra in a zoo enclosure, and, in the absence of any countervailing reasons, correctly believe it to be a zebra, and another case – the bad case – in which, in the same setting, I am looking at a mule cleverly disguised as a zebra and, in the absence of any

[3] See, for instance, Pollock (1986) and Pryor (2000). Davidson (1983) objected to any such view. I mounted a defence in Millar (1991).

[4] This is the line I took in Millar (1991).

countervailing reasons, incorrectly believe it to be a zebra. On the familiar way of thinking that I have in mind we need to accommodate the idea that these cases are on a par with respect to the justification of the belief. This will seem to be a natural requirement if one accepts the usual understanding of Gettier cases. On that understanding it is assumed that, in cases in which what is believed is concluded from other things one believes, one can be justified even if the other things believed include a falsehood. With respect to perceptual Gettier cases it is assumed that a belief based on misleading experience can be justified. A standard example of the latter is correctly believing that someone one knows is at a certain location in the space in front of one, though one is looking at a life-size photograph obscuring the person one takes oneself to be looking at. An account of justified belief based on conceptual competence, conceived along the lines sketched, gives an explanation of the supposed parity of justification in good/bad pairs: in both cases the belief is competently formed. Looking at the disguised mule I make a mistake in believing it to be a zebra but there is a sense in which it is quite reasonable that I should believe as I do. This reflects the fact that the belief is competently formed, as that is being understood. After all, the animal I am looking at is indistinguishable from a zebra from a point from which I can normally tell that an animal is a zebra from the way it looks, and I have not judged carelessly. It is not as if I am under some misapprehension as to what zebras are or what they look like. The mistake is due to the unusual character of the situation. There is then a case for trying to capture the sense in which my belief in the bad case is reasonable in terms of this way of thinking. Its reasonableness reflects the fact the belief is competently formed. This might encourage one to think that the suggested line of thought provides the right account of how experiences can be justifiers. But even if it is granted that there is a sense in which competence is exercised no less in bad cases than in corresponding good cases, there remain problems for the view that experiences can be justifiers.

One obvious problem is that justified belief seems to be a matter of having adequate reasons and experiences seem to be in the wrong category to constitute any kind of reasons. The most natural way to think of reasons is as being constituted by facts or true propositions. (For the purposes of the present discussion I need make no distinction between these.) It is not easy to spell out in general terms what it takes for a truth to constitute an adequate reason to believe something, but we have some sense of what governs our judgements on this matter. We want a truth that constitutes a reason to believe that *p* to be something that clinches it that *p* or is at least such that it is not likely that the truth should be a truth and it be false

that *p*. There is no similarly natural mode of evaluating beliefs that involves treating experiences as reasons under the operative understanding of experiences. Indeed, the operative understanding is a philosopher's construction. Philosophers talk of having a visual experience such that it looks to one just as if Bill has arrived. We can fairly readily acquire an understanding of what this amounts to by coming to understand why it might be thought philosophically illuminating to characterize the experiences in a way that is non-committal with respect to what if anything the subject sees.[5] The point is to introduce a convenient way of expressing the possibility of more or less radical differences between how it appears to us that things are and how things actually are. But (a) the non-committal conception of experiences does not routinely figure in our thinking about knowledge or beliefs and (b) the order of understanding is from more familiar conceptions of, for instance, seeing X to a conception of experience as non-committally characterized. What do figure routinely in our thinking about knowledge and belief are notions of the various modes of perceiving. If we know what it is to have an experience such that it looks as if Bill is arriving, it is in terms of what it is to see Bill arriving.[6] This should make us wonder whether it is right to work *from* the non-committal characterizations towards a philosophical account of perception and perceptual knowledge.

A radical response to the foregoing would be to sever the connection between being justified and possessing an adequate reason.[7] It might be argued that experiences must be justifiers and that there must be some basic *a priori* principle of justification such that if it perceptually seems to one just as if *p*, then, in the absence of countervailing reasons, one is justified in believing that *p*. If there are such principles they are not self-explanatory and, in any case, the problems are not solved by denying that being justified implicates having reasons. There remains the very basic problem that experiences, conceived non-committally, seem to provide too little to account for the special standing in which knowledge consists. The problem might not seem evident. After all, it is routinely accepted that being justified in believing that *p* is compatible with its being false that *p*.[8] But even if that is so we still lack an explanation of how perception can settle it that something is so: Gettier-style justification settles nothing.

[5] The drill described by P. F. Strawson (1979: 43–44) is intended to initiate us into such an understanding.

[6] The point about order of understanding is congenial to those who favour disjunctivist accounts of experience. See Child (1994: 143–146).

[7] James Pryor (2000) treats experiences as justifiers but does not represent them as reasons.

[8] For a contrary view, see Sutton (2007).

That matters since it is natural to suppose that evidence adequate for knowledge that *p* should settle it that *p*.[9] Further, even if it is granted that false beliefs are sometimes justified, and we help ourselves to the account of justification in terms of competence sketched earlier, there remains the problem posed by the fact that our evaluations of beliefs do not routinely deal with the operative understanding of experiences. Here I shall stick with the idea that justified belief depends on having reasons, and that reasons are constituted by propositions, and see whether we can end up with a view that reflects our evaluative practices.

## 11.3 KNOWLEDGE FROM INDICATOR PHENOMENA AND KNOWLEDGE THROUGH BEING TOLD

As we saw, gaining knowledge from testimony is problematic in straightforward cases in which we accept the testimony straight off on the say-so of the informant, because it seems to involve gaining a lot from a little. The issue is how it can supply an adequate reason. In section 11.1 I outlined a quick response to the problem that involved conceding that the fact that one has been told by someone that *p* does not provide an adequate reason for one to believe that *p* but which suggests that an adequate reason is available when this fact is combined with a suitable covering generalization. One attraction of this approach is that it applies to testimony what looks like a plausible way of dealing with empirical evidence. By contrast with perceptual knowledge, the problem here is not that what is thought to supply the reason is not of the right category to constitute a reason. It is that the general approach posits reasons that it is far from obvious that we have in the cases under consideration. It is not evident that the needed generalizations are routinely available to us. More specifically, it is not evident that we can so much as specify a determinate generalization that would do the trick; nor is it evident that if there are such generalizations we have adequate evidence for thinking them true.

Against this background it is understandable that some epistemologists dealing with testimony should have been attracted by defaultist conceptions according to which we are justified in accepting what people tell us, and tend to do so, in the absence of reasons to do otherwise. This way of thinking has the merit of doing better justice to the phenomenology – in

[9] I take for granted that when knowledge that *p* is based on evidence that *p* the evidence that *p* must clinch it that *p* in a sense that entails that there would be this evidence only if *p*. I see this view as being of a type instances of which have been held by Fred Dretske (1971), John McDowell (1982), and Charles Travis (2005). See further my contribution to Pritchard, Millar, and Haddock (2010).

particular, the straight-off acceptance of what we are told and apparent absence of reliance on supplementary premises to bolster an inference from being told something to the conclusion that what we are told is true. It also makes provision for our being discriminating with respect to believing what we are told. If it is to be made good though it requires a wider picture to make sense of why acceptance is the appropriate default stance towards testimony. Wider pictures designed to do this job are available.[10] I shall not discuss them here but simply record that I am sceptical that any highly general considerations about the social role of testimony or about human rationality will yield an adequate account of how we can gain knowledge from testimony in particular cases. For knowledge that *p* from being told that *p* in the straightforward cases we need to make sense of the idea that the testimony clinches it that *p* and I do not see how a defaultist view can do this. (Recall footnote 10.) A (perhaps) less controversial objection is that it distorts the phenomenology to suppose that in practice, by and large, we trust testimony unless there is reason to do otherwise. When nothing is at stake much of what people tell us washes over us without our adopting any stance towards it. When we do take ourselves to have learned something from being told it, we take the informant to be trustworthy on the matter in hand, not simply in virtue of being an informant in general, whom we have no particular reason to distrust, but as being *this* informant in *this* circumstance. The question is what entitles us to do this if we reject the reductionist picture that assimilates the epistemology of testimony to the epistemology of empirical evidence, using the model I described earlier.

The problem with the reductionist picture of testimony is not just that it assimilates the epistemology of testimony to the epistemology of empirical evidence in general. There is reason to think that the implicated account of empirical evidence is itself problematic. It is instructive in the present context to see why this is so.

Of particular interest is what we may call *knowledge from indicator phenomena* because this provides a model for reductionism about testimony. We gain such knowledge when, for instance, we tell from a reading on a thermometer that someone has a temperature of 38°, or tell from a blood smear on a surface that someone has touched it with bloodied hands, or tell from the frost on the grass that it has been freezing overnight. Note that in many cases in which judgements are made from indicator phenomena, the person making the judgement need not be in possession of evidence

[10] See, for instance, Burge (1993). Some of the considerations advanced in Coady (1992) could be adapted to this end. For some critical discussion of such approaches, see Moran (2006).

that directly confirms and establishes a suitable covering generalization. A person trained in the use of a thermometer can tell by applying the technique for its use, and thus know, what temperature a patient has. Such a person is committed to thinking that, for any n, readings of n° from properly applied thermometers indicate that the patient has a temperature of n°. It does not follow that the person need have evidence and on that account know that the generalization is true. Indeed, there are cases in which the relevant generalization might well be false. Imagine a community in which a certain style of house indicates that the resident has some official role in the community, say, that of being a police officer. (If no police officer resides in the house its style is altered.) Suppose that unknown to the people in this community there happen to be other communities in which there are houses in the same style but which are not occupied by police officers.[11] If I am a member of the community in question I might believe that any house in the given style is occupied by a police officer. If so I would believe falsely, yet by ordinary standards for attributions of knowledge I would not on that account be precluded from being able to tell, and thus come to know, from the style of a house in this community that it is the residence of a police officer.[12] Examples such as this, when taken along with the problem already mentioned of identifying suitable covering generalizations for indicator cases for which we have adequate evidential support, strongly suggest that acquiring knowledge from indicators has more do to with mastering a technique than it has to do with having adequate evidence for covering generalizations.

The conclusion just reached can be reinforced by reflecting on how we identify a phenomenon as having some indicative significance. Think of something as straightforward as telling from the wetness of surfaces outdoors (streets, pavements, etc.) that it has been raining. This is a simple ability yet it requires one to discriminate between rain-indicating wetness and wetness that has different causes, like people washing their cars, street cleaning vehicles, hydrants spraying water, and so forth. Learning to do this is a matter of learning to recognize rain-indicating wetness. The kind of procedure here is not in any fundamental way different from the procedure in perceptual cases, like learning to recognize goldfinches or thermometers by sight. It is a matter of being attuned to just the right kinds of visual

[11] The example is of the familiar fake-barn type (Goldman 1976, and much discussed subsequently).

[12] While the domain of quantification of universal generalizations is often determined in part by contextual factors, I can see no reason to assume that the members of the envisaged community would not understand the generalization in question as being true with respect to dwellings in general rather than simply to the dwellings in the area with which they are familiar.

appearance. It is not a matter of acquiring a list of features that collectively are distinctive, learning to judge when they are present, and on particular occasions basing a judgement that the thing in question is a goldfinch, or whatever. Indeed, one might have the recognitional ability yet lack the conceptual resources to form judgements to the effect that such-and-such a feature is present. Of course, in judging something by sight to be of some kind one is responding to its features, but the response is a recognitional response to the distinctive *Gestalt* that the features make up, not an inferential response to considerations pertaining to the features. The same applies to indicator cases. One learns that the right kind of wetness indicates that it has rained. The ability thus acquired deserves to be regarded as recognitional. It is a matter of being able to recognize that occurrences of an observed phenomenon – wetness of the right kind – have a certain significance: they indicate that it has rained. The work of learning is towards making the right kind of discriminations. It would not be false that one has this ability if there happened to be some place where, because of the peculiar water-spraying practices of the inhabitants, you could not tell that it has rained recently from the wetness of the surfaces in view. If there were such a place it could be false that when (or even most times when) there is the wetness that we associate with rain it has rained. Yet this would be compatible with our being able, around here, to tell that it has rained by exercising the ability.

It is of some interest to note in passing that the abilities of which I am speaking amount to a kind of competence in the deployment of concepts. But competence so understood is very different from that discussed in section II.2 in connection with the supposed parity of justification in pairs of good and bad cases. It is built into perceptual-recognitional abilities that they are abilities with respect to favourable environments and that their exercise *is* the acquisition of knowledge. In a suitably strange environment I might not be able to recognize something to be a zebra from the way it looks, because too easily could something look that way and not be a zebra. The nature of the environment is crucial for whether one has the ability to recognize things in that environment to be of some kind.[13]

Our reflections on indicator cases support the view that the standard way of thinking about indicator cases is inadequate. I have suggested that

[13] There are theories (e.g. that of Ernest Sosa 2007) on which a subject is represented as having a true belief that is competently formed when luckily looking at a barn. The understanding of competence is not the same as any understanding on which the relevant competences are recognitional abilities. I explore the difference in Millar (2009).

knowledge from indicator phenomena is best understood as a kind of recognitional knowledge made possible by suitable recognitional abilities. As I have already observed some such abilities are exercised in the acquisition of perceptual knowledge.[14] If I can tell whether a bird is a goldfinch from the way it looks, and I exercise that ability when looking at a goldfinch, I thereby acquire knowledge that a goldfinch is there. The knowledge is recognitional rather than being a matter of drawing a conclusion from other things I believe. I have sought to make it plausible that recognitional abilities are also exercised in recognizing occurrences of a phenomenon as having a certain indicative significance. In these cases the state of affairs indicated need not fall within the scope of what we currently perceive. In telling that it has been raining from the wetness that is around I need not see the rain; in telling that a patient has a temperature of 38° I do not see the temperature. It is striking though that in many of these indicator cases the natural description of the observed phenomenon entails that it has the significance in question. Tracks on a path, recognized as deer tracks, are naturally described as, of course, deer tracks, entailing that they have been caused by the passage of deer. That marks on a road are skid marks entails that they have been caused by a skidding vehicle. The fact that these descriptions entail something about the cause of the phenomenon described is not at odds with the recognitional character of the knowledge that a suitably equipped person can acquire by discerning the phenomenon. Whether or not our natural way of describing the indicator phenomenon entails anything about what it indicates, it is very often, perhaps always, natural to talk in perceptual terms of our knowledge of what is indicated.[15]

The role of recognition is of first importance in connection with an overarching theme of this chapter, which concerns the problem posed by kinds of knowledge with respect to which it seems puzzling that we gain so much from what apparently is so little. It contributes towards appreciating that in the problem cases we have more to go on than might initially appear. More specifically, (a) in virtue of our recognitional abilities we can take in facts that are rich in the sense that they concern not just superficial features of objects or situations, and (b) in exercising recognitional abilities whereby we take in rich facts we are responsive to phenomena that have

[14] See further Millar (2007, 2008a, 2008b, 2009), and my contribution to Pritchard, Millar, and Haddock (2010).

[15] This is how Dretske (1969) conceives of such cases.

a high degree of informational richness. A good example of both of these points is the recognition of people. We can recognize some people as named individuals. I can see that it's Bill, a person I know, entering the room, not just something with certain superficial features. In doing so I am responsive to a host of features that make up the visual appearance of Bill entering. While I tell that it's Bill from his visual appearance I could not give a specific description of everything that goes to make up that appearance. In this kind of case we respond to a *Gestalt* that is informationally rich, being the resultant of determinate forms of a number of dimensions of variation. The lesson to draw, as much from indicator cases as from perceptual cases like that just described, is that after all we have a lot to go on. The real puzzle is not that we gain a lot from a little, since there is a great deal to which we respond when we gain knowledge of the kinds under discussion. The problem is to make sense of how the many features to which we respond can impinge on our thinking. It is in connection with this that it matters that we acknowledge the role of recognition and recognitional abilities. We court failure if we try to do justice to the relevance of the features by supposing that we register their presence at the level of judgement and then apply a suitable covering generalization. The problem for that view is not just whether routinely we have evidence for the required generalization, nor even that some of the generalizations in play in the cases in question might be false (as suggested by analogues of fake-barn cases). It is that in practice we need not take in the relevant features at the level of judgement. To recognize people, animals, birds, plants, or, for that matter, smoke, skid marks, deer tracks, and so forth you need to 'get your eye in'.

Analogous considerations apply to the epistemology of testimony. For gaining knowledge from testimony in straightforward cases it matters that the informant is trustworthy with respect to the thing told and that the recipient of the testimony has identified the informant as one who may be trusted on that matter. The reductionist approach can be viewed as aiming to explain how recipients identify trustworthiness in terms of features of the occasion of utterance that make it such that utterances with those features are always, or highly likely to be, true. That a host of such features is relevant to the case is not in dispute. If we reject the reductionist approach then we need to account for how they bear on our thinking and make sense of how the recipient can end up with knowledge. On the approach that I commend, instead of supposing that we register the features at the level of judgement, we should think rather in terms of our having acquired a certain sensibility that enables us to pick out acts of telling that may be

relied upon.[16] I envisage this sensibility as being shaped by our experience of individual people and of the workings of the human world and our physical environment. The shaping is to be conceived as a matter of honing our discriminative capacities rather than as supplying us with information on the basis of which we accept generalizations that we then apply to particular cases. It is not that acquiring new information is irrelevant. Suppose that I have learned visually to identify roses from examples, or illustrations of examples of garden-variety roses. When walking by a hedgerow along a country road I fail to recognize a rambling rose as a rose. When it is pointed out to me that this rambling plant is a rose, though its flowers seem so unlike the full blooms of garden roses, I acquire new information about how roses can look. My ability to tell by sight when a flower is a rose is refined so that I avoid false denials that I would have made previously and become generally more circumspect about denying that something is a rose when it is unlike the flowers I originally learned to call roses. It might well be that all this is accompanied by some adjustment to rough generalizations I have previously accepted. But it is the refinement of the recognitional ability that is explanatorily central. Any generalizations that informed my thinking about roses would be unspecific with respect to features. I thought of roses as having full blooms like *these* (pointing, say, to examples in a garden flowerbed). I then realize that roses can look like the rambling roses one sometimes finds in country hedgerows, so I no longer think that roses always have the flowers of the typical garden varieties. Closely analogous considerations apply to straightforward cases of testimony. The features of an act of telling that lead us, rightly, to accept what we are told are elusive. In particular circumstances people come over to us as telling the truth or as not telling the truth or leave us in doubt or indifferent as to whether they are telling the truth. As I remarked previously, we are not at a loss to point to some of the variable factors that influence our response. It is implausible to suppose that we latch on to determinate forms of these at the level of judgement, but entirely realistic to suppose that discriminative capacities come into play, capacities that not only explain our responses to what we are told but which can amount to abilities to tell that the informant is telling the truth. If the departmental secretary phones me at home to tell me that a student wishes to see me urgently, I thereby know that there is a student who wishes to see me urgently. That requires both that the secretary should be highly reliable on such matters and that I am attuned

[16] Miranda Fricker (2003) deploys the notion of a testimonial sensibility in the context of a discussion of cases in which reliable informants are not taken seriously by others.

to this and to the character of her communication. My point has been that this attunement is better seen as involving a sensibility acquired through experience than explained in terms of the reductionist model of reasoning from matters of fact.

I have only gestured towards an account that accords a central role to sensibility in the epistemology of testimony. A fuller account should, I believe, accord an equally central role to the idea that there is a social practice surrounding the speech act of telling. I mean by this that telling is a move in a cluster of essentially rule-governed activities into which we are socially initiated. The practice accounts for the ease with which we can identify utterances as acts of telling whereby speakers give others to understand that they are informing and thus speaking from knowledge.[17] It also accounts for widely shared understanding that telling is *supposed* to impart knowledge. I do not, however, think that an adequate account of the epistemology testimony drops out of the theory of the speech act of telling any more than an account of what justifies us in trusting promises made to us falls out of the theory of the speech act of promising. In both cases the rules of the practice can be readily flouted. We need in addition an account of the powers of discrimination whereby we can, sometimes, trust the informant or the person making the promise.

Even if what I have said thus far about perceptual knowledge, knowledge from indicators, and knowledge from testimony seems promising it has probably not escaped notice that I have not as yet explained how reasons fit into the picture. It is to this that I now turn.

## 11.4 KNOWING AND HAVING REASONS TO BELIEVE

I see my neighbour in her garden and thereby come to know that she is in her garden. By the operative assumptions I am justified in believing this since I have an adequate reason to believe it. Common sense says that my reason is that I see her in her garden, or something along these lines. Of course, it could be true of me that I see her but do not know that it is she. I need to recognize the person I am looking at as my neighbour. In other words, I need to *see that* it is she. Strictly, it is the fact that I see that she is in her garden that is my reason for believing that she is there. All this is unexceptionable from the perspective of common sense, but it

[17] The operative conception of a practice is outlined in Millar (2004) and applied to testimony in Millar (2010). I take for granted that not every act that counts as an assertion is an act of telling. One can assert things in giving advice, or expressing opinions, for instance, without being in the business of telling and thus giving it to be understood that one is informing.

can all too easily seem unsatisfying to the epistemologist. Though there are notable exceptions, epistemologists have not in general given factive states, like seeing that something is so, a central role in their accounts of the justification that is linked with our perceptual knowledge.[18] It is not difficult to see why. The reason I am supposed to have in the case in question is that I see that my neighbour is in her garden. Call the proposition that constitutes this reason R. It seems that if R is to be available to me to serve as a reason to believe that my neighbour is in her garden then I must have reason to believe it. But now the worry is that since R is true only if my neighbour is in her garden it looks as if I shall need to have reason to think that she is in her garden, if I am to have reason to think that I see that she is in her garden. But we were assuming that the former reason – the reason to believe that she is in her garden – was constituted by R. So it looks as if our account gets us nowhere. In explaining why I have a reason to believe that my neighbour is in her garden I have invoked a reason, constituted by R, which it seems would be available to me only if I have a reason to believe that my neighbour is in her garden – a reason that, on pain of circularity, cannot be given by R. And if I have some other reason to believe that my neighbour is in her garden then the original proposal is at best incomplete. Before diagnosing what goes wrong with this line of thought I shall say how, in my view, we should think of the situation.

First, we need to note that I recognize my neighbour because I have a specific recognitional ability that serves this purpose – in suitable conditions I am able to tell by looking at a person whether or not he or she is my neighbour. On this occasion I have exercised this ability and have thereby come to know that my neighbour is in her garden. There is no doubt that the natural way to spell out my reason for taking my neighbour to be there is in terms of the fact that I see that she is there. So next we need to explain how that reason is available to me – how I know that I see that she is there. This is not so very difficult. For my knowledge here is also recognitional, arising as it does in direct response to my current visual experience, via the exercise of a certain recognitional ability, specifically, an ability to tell by looking at an F/something G/X that I see it to be an F/something G/X. In the case in hand, I recognize my neighbour to be someone I see to be my neighbour in her garden. That amounts to knowing that I see that she is in her garden. It would be at least misleading to say that I introspect my experience for my attention is directed outwards towards my neighbour.

[18] Notable exceptions include McDowell (1994a), Williamson (1995, 2000), Stroud (2004), and Cassam (2007).

Just as I exercise a recognitional ability when I apply the concept of *my neighbour being in her garden* to my neighbour, in direct response to the visual experiences I have as I look over at her in the garden, so I exercise a distinct recognitional ability when I apply the concept *seeing my neighbour to be in her garden* to me and my neighbour, respectively, in direct response to the very same experiences. By exercising these abilities I know both that my neighbour is in her garden and that I see that she is. The fact made available to me by knowing that I see that she is there serves as a reason to believe that she is, and for some time thereafter serves as a reason to believe that she was then in her garden.

A significant feature of this account is that we explain how I know that my neighbour is in her garden without adverting to the reason for me to believe that she is that is constituted by the fact that I see that she is there. Rather than account for the knowledge in question in terms of believing for a reason, we account for my possession of a reason for believing in terms of my possession of two items of knowledge – first-order knowledge that she is there, and second-order knowledge that I see that she is there – along with my understanding of the relevance of my seeing that she is there to whether she is there. The second-order knowledge is my access to a reason to believe what I know in the first-order way. The problem noted in the paragraph before last was how to account for the availability to me of the reason to believe that my neighbour is in her garden that is constituted by the fact that I see that she is. It seemed that this fact would be available to me only if I have reason to believe that my neighbour is in her garden. *Ex hypothesi* the reason for so believing is constituted by the fact that I see that my neighbour is in her garden, which takes us nowhere in explaining how the fact in question is available to me. Under the present account we explain the availability of this fact in terms of the exercise of a second-order recognitional ability that works in tandem with the first-order recognitional ability whereby I come to know that my neighbour is in her garden. The problem is thus avoided. The reason we are looking for is what pre-theoretically we all know it to be: I see that my neighbour is in her garden. Of course, if I were trying to give a reductive conceptual-analytical account of what it is to know that my neighbour is in her garden, this account would be hopeless, since the very idea of a recognitional ability implicates the concept of knowledge. But the account is still informative since I might have come to know that my neighbour is in her garden by some other means than through seeing her and recognizing her.

The account reverses the traditional philosophical order of explanation as between knowledge and justification in cases of perceptual knowledge.

Possession of justification in these cases arises out of what we know about our environment and about our mode of perceptual access to it. This removes any incentive to address the problem of justification for perceptual knowledge by looking for justification in experiences, conceived in non-committal terms, yet it accords perception an explanatory role in relation both to the acquisition of knowledge and the possession of reasons. Provided that we have suitable recognitional abilities, it is what we perceive that enables us both to know that something is so through perceiving that it is so *and* to have access to a reason to believe that it is so.

In indicator cases the natural explanation of how we come to know is that we have discerned, often directly through perception, that the indicator phenomenon has occurred, and are able to recognize the significance of the occurrence of that phenomenon. Recognizing the significance of the wetness of surfaces outside just is seeing that this wetness indicates, and thus provides a clinching reason to think, that it has been raining. The same applies to straightforward testimonial cases. Someone's telling me that *p* can be a reason to think that *p*, and as in other indicator cases, it can be a reason to think that *p* that clinches it that *p*. But one has to be suitably equipped to recognize this reason for what it is.

Earlier I challenged the idea that we can explain knowledge from indicator phenomena in terms of having justification deriving from evidence for believing a suitable covering generalization. The challenge leaves me with a challenge in turn. Consider the case of telling that a vehicle has skidded from marks on the road. If I do this, I do not do it blindly. That is to say, I do not simply find myself judging that a vehicle has skidded with no understanding of what led me so to judge. I am aware that I have judged that a car has skidded on the basis of the presence of the marks and it is part of my understanding of what must have happened, and of my ability to tell what must have happened, that marks of this sort indicate that a vehicle has skidded. So it is natural to ask what entitles me to suppose that this is so. The right answer might simply be (i) that, at least in the environment in which I make judgements to the effect that a vehicle has skidded on the basis of such marks, such marks always or very nearly always do indicate that a vehicle has skidded, and (ii) that I have learned that marks of this sort indicate that a vehicle has skidded. What I have thus learned is a generalization but we need to recall the discussion about generalizations in section 11.2. I need not have adequate direct evidence is support of the generalization. Much of what we learn is learned from others; I need not have exercised the discriminative capacities required to gain knowledge through testimony any more than very young children at

school, who imbibe information from teachers, need have exercised such discriminative capacities. (I shall say more about this in section 11.5.) I might simply have learned a technique – that this is a way of telling that a vehicle has skidded – and as part of that learning I have acquired an understanding that there is this technique and that I have it.

It is against this background that I wish to account for the acquisition of knowledge through testimony. We need, I think, to turn our attention away from covering generalizations and towards the sensibilities that enable us to discriminate truthful acts of telling. These sensibilities are applied not simply to discriminate between people, but to pick out from among a person's utterances those that are to be believed on the say-so of that person. People who are knowledgeable on some topics are not on others. Those who are sincere on some topics can be insincere on others. It is only through experience that we become attuned to picking out what is to be believed. However, the role of experience is not best conceived as that of supplying evidence on the basis of which we hold people trustworthy on this or that topic. Its role is that of inculcating and shaping a sensibility.

## 11.5 DETACHED STANDING KNOWLEDGE

There is a type of knowledge that has not been extensively discussed though much of our knowledge is of this type. It is knowledge of factual matters that we have picked up and retained though we have lost touch with the relevant sources of information. I call it *detached standing knowledge*. This is of interest both because it poses further problems about reasons for belief and because reflection on this type of knowledge will serve to reinforce the claim that the epistemological role of experience is not confined to that of supplying evidence on which we base our beliefs. I suspect that the reason why detached standing knowledge has not received much discussion is that it is assumed that it is straightforwardly based on evidence. The matter is more complicated than this suggests.

I know that kangaroos are marsupials. I have learned this from numerous sources of information including books and documentary films, but I do not recall any specific source. Is my current knowledge evidence-based? One might say that knowledge based on evidence is simply knowledge that was acquired through encounters with evidence and that has been retained. In that rather weak sense this knowledge would be knowledge based on evidence. Yet the fact that the evidence explains how I came to acquire the knowledge does not account for the status of my knowledge now. For one thing, I have lost touch with the evidence: no memory of it is playing any

role in sustaining my current knowledge. For another, my past encounters with the relevant sources of information do not explain why I count as having retained *the knowledge* rather than just a willingness to avow that I know.

We can, I think, make sense of the idea that we routinely have knowledge based on evidence in a stronger sense. I might know that deer were in the vicinity yesterday because I recall seeing fresh deer tracks. In that case the manner in which I now know that there were deer around yesterday is tied to recollection of the evidence provided by the tracks. It is based on the evidence recalled, not just in that my having seen the tracks caused me to know in the first place but in the sense that the recalled evidence plays a role in both sustaining my knowledge and explaining why I now know. My current knowledge that kangaroos are marsupials is not like that: it is not evidence-based in that strong sense. Even if it is evidence-based in the weaker sense, we still need an explanation of why it maintains its status as knowledge. I want to sketch a picture of detached standing knowledge that makes abilities central once again.

The key idea is that detached standing knowledge is, roughly speaking, an ability to recall a fact. For convenience I shall take recall to include cases like that in which, faced with the question, 'What is the capital of X?' one can correctly answer 'Y', as well as cases in which, faced with the question 'Is Y the capital of X' one can correctly answer 'Yes'. One might be able to recall that Y is the capital of X in the second way, yet not be able to recall it in the first way. The abilities in which our detached standing knowledge consists have been acquired either through direct confrontations with the relevant facts or through reliable sources of information that record or report the relevant facts. Some of our knowledge was originally inculcated by teachers. We soaked up the facts, quite reasonably trusting our teachers. We might have learned some geographical facts through news reports of areas of conflict, again quite reasonably trusting the sources on such matters. But we should not think of what is happening here on the model of gaining knowledge from testimony if only because much of what we have imbibed has not brought into play the discriminatory capacities that are necessary for the acquisition of knowledge on the say-so of another. On the model I suggest, we acquire the ability to recollect a fact through encounters – usually repeated encounters – with reliable sources of information. The role of experience is that of honing the ability rather than furnishing us with evidence-based knowledge in the sense explained above – the strong sense that requires that recollection of the evidence plays a role in sustaining the knowledge. It is important not to take the view to imply that *any*

disposition to acknowledge that it is a fact that such-and-such counts as knowledge of that fact. It matters that the facts recalled are available, and became available to us, though reliable sources.

A strength of the proposed account is that it accommodates the possibility that much of our standing knowledge derives from having soaked up information before we had the discriminatory capacities for weighing up testimony and other sources of information, which is to say, before we were in a position to acquire evidenced-based knowledge. Indeed, since this seems to be a very live possibility, a weakness of the view that all standing knowledge is evidence-based is that it can account for early learning only by finessing on what it takes to have evidence-based knowledge. There is no need to go in for such finessing if we acknowledge that experience can have an ability-inculcating role that is distinct from its role in providing us with evidence.

Adopting this view, can we respect the idea that, even with respect to detached standing knowledge, reflective agents routinely know how they know? The first point to note is that we can discriminate between, on the one hand, what we know, clearly recalling it to be so, and, on the other hand, what we think might well be so but do not know. If we are sufficiently discriminating in these matters, our confidence that we recall some fact can be an indicator that what is recalled is a fact – an indicator from which we tell that it is a fact. These discriminatory capacities are possessed and exercised in the context of an understanding of the workings and practices of the human world and of our own experience of sources of information. If we do not know specifically how we came to know the facts in question, we know about the kind of sources from which we gleaned our information and we know that we make use of these sources. The upshot is that we can rightly be confident that we remember some fact if we have exercised a capacity for discriminating between what we do and do not remember. Such a capacity will be informed by an understanding of the general ways in which the information would have been acquired. In these cases we do not know specifically how we came to know, but there is a sense in which we still know how we now know. We know in that we remember.

How can detached standing knowledge that *p*, as I have conceived it, yield justified belief that *p*? My reason for thinking that kangaroos are marsupials is simply that I remember that kangaroos are marsupials. Ask me if Tasmanian Devils are marsupials and I could not say without checking. But that kangaroos are marsupials I know because I remember and the fact that I remember can serve as a reason for believing this to be so. So long as

I remember, and am good at telling what I remember, this knowledge and this justification will be sustained.

## 11.6 conclusion

The perspective I have outlined explains how with respect to various kinds of knowledge we gain a lot from what initially seemed to be so little. It strengthens the case for thinking that we gain fresh insights into how best to make sense of various kinds of knowledge if we think of knowledge as prior in the order of explanation to being justified. The task for a substantive epistemology is to account for how we acquire knowledge of the various kinds, why what we acquire counts as knowledge, and how the knowledge can yield justified belief. I have addressed these matters by according a central role to recognitional abilities in relation to perceptual knowledge, and knowledge from indicators. I have given some indication of the direction that an account of knowledge from testimony should take. In relation to these kinds of knowledge I have linked the availability of a reason for belief to an understanding of how we know what we know. Finally, I explored a way of dealing with detached standing knowledge and the related problems that it poses. Much more needs to be investigated but I hope to have conveyed that enquiry in this direction is worthwhile.[19]

[19] The ideas outlined in this chapter derive from work on a project on the Value of Knowledge funded by the UK's Arts and Humanities Research Council. I am grateful to the Council for its support and to my fellow participants, Adrian Haddock and Duncan Pritchard, for numerous helpful discussions. Special thanks are due to Adrian and to an anonymous reader for written comments that I believe led to improvements.

CHAPTER 12

# *What is the swamping problem?*

*Duncan Pritchard*

## 12.1 TRUTH, BELIEF AND EPISTEMIC VALUE T-MONISM

Most will agree that there is a constitutive relationship between belief and truth. Often, this relationship is expressed in terms of the slogan that belief in some sense 'aims' at the truth; that the *telos* of belief is truth. I think this slogan is essentially correct, though we need to be careful how we are understanding it. After all, since most beliefs are involuntarily formed, we do not want to be reading this slogan in such a way that it commits us to supposing that the agent forming the belief has some sort of occurent intention to form true beliefs whenever a belief is formed. Instead, I take it that what is meant by this claim is rather that what makes a belief the kind of propositional attitude that it is is that it is sensitive to considerations which indicate that the belief is either true or false. This is why one cannot self-consciously regard oneself as having overwhelming reason to think that a proposition is false and yet believe it nonetheless, for this would not then be a case of belief, but rather wishful thinking or something like that.[1]

Given that belief has this aim, a primary way of evaluating beliefs is in terms of whether they achieve this aim: whether they are true. Relatedly, we will assess beliefs in terms of whether they are based on reasons which suggest the truth of the proposition believed. Of course, there are other ways of assessing beliefs where truth does not enter the picture at all but, since they do not concern the aim of belief, they will not be *primary* ways, in the sense that I am using the term here.

[1] There may be some special cases in this regard. Perhaps, for example, some beliefs simply cannot be given up, even when the countervailing evidence is incredibly strong. Interestingly, though, once one discovered that this propositional attitude was involuntary in this strong fashion, and so not susceptible to evidential considerations, it ceases to be so clear that it should continue to be regarded as a belief. Thus, even the tricky special cases ultimately speak in favour of this claim. For further discussion of the idea that truth is the aim of belief, see Wedgwood (2002a), Shah (2003), Shah and Velleman (2005), and Steglich-Petersen (2009).

One might, for example, evaluate beliefs in terms of how useful they are. Believing that one is very good-looking may be very useful, for instance, in that it gives you lots of confidence in social encounters, even if it isn't true. In one sense, then, one has a reason to believe the target proposition even though one lacks a reason for thinking that this proposition is true. But that there are these other types of belief appraisal does not undermine the fact that a primary form of belief appraisal is in terms of its relationship to the truth. This kind of belief appraisal is paradigmatically *epistemic*.

Given the foregoing, it becomes very tempting to suppose that this form of epistemic evaluation is not just *a* primary form of belief appraisal, but is rather *the* primary form of belief appraisal. That is, one might be tempted to endorse a view that I am going to call *epistemic value T-monism*, a position which I think is widely accepted within epistemology:[2]

*Epistemic value T-monism*

True belief is the sole ultimate fundamental epistemic good.[3]

What do I mean here by 'fundamental'? Well, a good is fundamental, according to how I am using the term, if its value is not instrumental value relative to further goods of the same type. In terms of the value of true belief, the claim is thus that the value of true belief is not instrumental value relative to further *epistemic* goods. Notice that this way of describing the matter leaves it open as to whether true belief is non-instrumentally (i.e. finally) valuable. Perhaps it is, or perhaps it isn't and ultimately the value of true belief is just instrumental value relative to further non-epistemic goods, such as the achievement of our practical goals. Either way, true belief counts as a fundamental epistemic good on this view.

The reason why granting that the primary way of evaluating beliefs is in terms of their truth seems to lead to this position is that it strongly suggests both that there could be no further form of *epistemic* appraisal of a belief (hence, true belief is a *fundamental* epistemic good), and also that all other epistemic goods ultimately derive their epistemic value from the epistemic value of true belief (hence, true belief is the sole *ultimate* epistemic good).

[2] For recent discussion of this view, see David (2001) and the exchange between David (2005) and Kvanvig (2005).

[3] Note that sometimes this view is expressed in terms of the thesis that it is *truth* (rather than true belief) which is the only ultimate fundamental epistemic good (see, for example, Kvanvig 2003: ch. 2). The difference between these two views is complex and raises some subtle questions. However, since nothing is lost for our purposes by focusing on the version of epistemic value T-monism set out here, I will ignore this complication in what follows.

This second claim might not seem so obvious. Aren't there lots of other epistemic goods beside true belief, such as knowledge, justification, rationality etc.? Well, yes, but the thought is that what makes these goods epistemic goods, on this view, is that they are a means to true belief. Thus, the epistemic value of, say, justification, is due to the fact that justified beliefs tend to be true beliefs. Hence, the epistemic value of this good is parasitic on the epistemic value of true belief in the sense that the epistemic value of this good is instrumental epistemic value relative to the fundamental epistemic good of true belief.

I think that many are attracted to epistemic value T-monism, and attracted to it, at least in substantial part, because of the intuition that we started with: that belief in the relevant sense aims at truth (and all that this implies). Relatedly, this conception of epistemic value accords with a very natural conception of what properly constitutes a reason for belief – *viz.*, that reasons for belief are considerations which speak in favour of the truth of the believed proposition. As we will see, however, there is a fundamental difficulty facing this position, although few have recognized this fact.

## 12.2 THE SWAMPING PROBLEM

This difficulty is the so-called 'swamping problem', as defended most prominently by Jonathan Kvanvig (e.g. 2003), but also put forward in various forms by Ward Jones (1997), Richard Swinburne (1999, 2000), Wayne Riggs (2002b), Linda Zagzebski (2003), and John Greco (2009). The standard way of expressing this problem is as posing a difficulty not for a general epistemological thesis like epistemic value T-monism but rather for particular epistemological proposals, such as reliabilism.[4] With the argument so directed, here is how it goes.[5]

Imagine two great cups of coffee identical in every relevant respect – they look the same, taste the same, smell the same, are of the same quantity, and so on. Clearly, we value great cups of coffee. Moreover, given that we value great cups of coffee, it follows that we also value reliable coffee-making machines – i.e. machines which regularly produce good coffee. Notice, however, that once we've got the great coffee, we don't then care whether it was produced by a reliable coffee-making machine. That is, that the great coffee was produced by a reliable coffee-making machine doesn't

[4] Though see Riggs (2002b).

[5] The following discussion closely follows that offered by Zagzebski (2003).

contribute any additional value to it. In order to see this, note that if one were told that only one of the great identical cups of coffee before one had been produced by a reliable coffee-making machine, this would have no bearing on which cup one preferred; one would still be indifferent on this score. In short, whatever value is conferred on a cup of coffee through being produced by a reliable coffee-making machine, this value is 'swamped' by the value conferred on that coffee in virtue of it being a great cup of coffee.

The supposed import of this example to reliabilist theories of knowledge – theories which hold that knowledge is reliably formed true belief – is that it follows, by analogy, that if we are faced with two identical true beliefs, one of them reliably formed and one not, it shouldn't make any difference to us which one we have. After all, we only value reliable belief-forming processes as a means to true belief, just as we only value reliable coffee-making machines as a means to great coffee, and so once we have the good in question – true belief or great coffee – then it shouldn't matter to us whether that good was in addition acquired in a reliable fashion. Intuitively, though, it *is* better to have reliable true belief rather than just mere true belief. Moreover, if, as the reliabilist maintains, reliable true belief is knowledge, then this intuition is stronger still. For surely many of us have the intuition that it would be better to have knowledge rather than mere true belief.

More generally, the claim implicit in the coffee cup case is that if a property (like being reliably formed, when it comes to beliefs, or being reliably produced, when it comes to coffee) is only instrumentally valuable relative to some further good (e.g. true belief or great coffee), then in cases in which the further good in question is already present, no further value is conferred by the presence of the instrumentally valuable property. This is, of course, a general thesis about value. Moreover, as the coffee cup example illustrates, there seems every reason to think that it is true.

Interestingly, however, once we recognize that what is driving the swamping argument is this general thesis about value, then it ceases to become plausible that it should only affect reliabilist views. After all, the general value thesis just described will have an impact on *any* epistemological proposal which has the same relevant features as reliabilism – *viz.*, which treats the epistemic standing in question as instrumentally valuable only relative to the good of true belief.

For example, suppose that we only value justification because it is a means to true belief – i.e. we only value it instrumentally relative to the good of true belief. If that's right, then we could just as well run the swamping argument for justification as we can for reliability, since if

the swamping argument works at all then any value that is conferred on a belief in virtue of its being justified will be swamped by the value conferred by the belief's being true. The problem, then, is not specific to reliabilism.[6]

Indeed, once we recognize the role that this entirely general thesis about value plays in the swamping argument, then I think it starts to become clear that the real challenge posed by the swamping problem is not to any specific epistemological proposal but rather to epistemic value T-monism. The reason for this is that epistemic value T-monism entails that any epistemic value that is contributed to a belief in virtue of it enjoying an epistemic standing like being justified or reliably formed is necessarily instrumental epistemic value relative to the fundamental epistemic good of true belief. But if that's right, then the swamping argument, if effective, demonstrates that there cannot be anything more epistemically valuable than mere true belief. *A fortiori*, it demonstrates that knowledge cannot be more epistemically valuable than mere true belief.

More formally, we can express the swamping argument in terms of an inconsistent triad of claims. The first claim is just the thesis of epistemic value T-monism, which we will express as (1):

(1) The epistemic value conferred on a belief by that belief having an epistemic property is instrumental epistemic value relative to the further epistemic good of true belief.

The second claim is the general thesis about value that we noted above:

(2) If the value of a property possessed by an item is only instrumental value relative to a further good and that good is already present in that item, then it can confer no additional value.

As we have just seen, however, (1) and (2) entail that there cannot be an epistemic standing which is epistemically more valuable than mere true belief. After all, whatever epistemic value an epistemic standing contributes could only, given (1), be instrumental epistemic value relative to the epistemic good of true belief. But (2) makes clear that this value gets swamped by the value of true belief. Thus, there can never be a true belief which, in virtue of possessing an epistemic property, is epistemically more valuable than a corresponding mere true belief.

In order to get our inconsistent triad, then, all we require is a thesis to the effect that there is an epistemic standing that is at least sometimes epistemically more valuable than mere true belief. For the sake of simplicity,

[6] Indeed, the problem isn't specific to externalist epistemological proposals more generally either. The argument that was just run as regards justification would work with equal force whether the notion of justification in question were cashed out in an externalist or internalist fashion.

we will here focus on the epistemic standing of knowledge, since I take it that many have the strong intuition that knowledge is better, from an epistemic point of view, than mere true belief. We thus get (3):

(3) Knowledge that *p* is sometimes more epistemically valuable than mere true belief that *p*.

I will comment further on how we should understand (3), and why we should find it so compelling, in a moment.

In short, then, (1), (2), and (3) are jointly inconsistent because it straightforwardly follows from (1) and (2) that (3) is false. Thus, if we accept (1) and (2), we are thereby committed to the denial of (3). Given that this is an inconsistent triad, it follows that one must reject one of the claims that make up this triad. One option, of course, is just to reject the general claim about value expressed as (2). Unfortunately, I can see no way of objecting to this claim, nor am I aware of any good objections to this thesis in the literature (typically, responses to the swamping argument focus on other elements of the argument than this claim). Accordingly, in what follows I am going to take this thesis for granted and focus instead on the status of the other two members of this triad, since this is where I think the real promise of responding to this problem lies.

## 12.3 DENYING THE GREATER VALUE OF KNOWLEDGE

Let me start with the option of denying (3). In particular, it is important to emphasize that (i) it is specifically *epistemic* value that is at issue here, and (ii) the thesis is only that knowledge is *sometimes* epistemically more valuable than mere true belief. This is important because if (3) were the claim that knowledge is *always* and *all things considered* more valuable than mere true belief then it would be obviously false. After all, there may be certain propositions which, if you knew them, would cause you great pain, pain that is not caused by you merely truly believing them. In such a case, it would clearly be better, all things considered, to have true belief rather than knowledge. Clearly, however, the example just given is not a counter-example to (3). After all, the greater value of true belief over knowledge in this case is an all-things-considered value, rather than a specifically epistemic value. Moreover, (3) doesn't say that knowledge must always be more valuable than mere true belief anyway, even when the value is restricted to specifically epistemic value.

Still, once it is clarified that it is specifically epistemic value that is at issue in (3), I think that many might well be tempted to argue that the right way to respond to this inconsistent triad is to concede that, from

an epistemic point of view, there is nothing better about knowledge (or, indeed, any other epistemic standing) compared with mere true belief.[7] On this proposal, what makes it the case that knowledge is sometimes a better thing than mere true belief (assuming that it is sometimes better, all things considered) is purely a function of its additional *practical* (i.e. non-epistemic) value. That is, the greater value of knowledge compared with mere true belief is indeed an additional instrumental value, but this value is not swamped by the value of true belief because it is not epistemic value that is at issue but practical value. Moreover, a further advantage of this proposal is that it can explain away our intuition that knowledge is more valuable than mere true belief in terms of a failure to distinguish between all-things-considered value and specifically epistemic value.

Interestingly, it seems that Plato might well have been inclined to respond to our inconsistent triad in this way. Recall that in the *Meno* Plato raises the question of what makes knowledge more valuable than mere true belief, and conjures up the famous story of the road to Larissa to make this worry vivid. Adapted slightly, the worry he poses is this: given that a true belief about the correct way to Larissa and knowledge of the correct way to Larissa generate no practical differences (either way one gets what one wants, to get to Larissa), why should we suppose that knowledge is more valuable than mere true belief?

It is important to note that the problem that Plato is expressing here is very different to the problem posed by the swamping problem. In particular, what Plato wants is an explanation of why knowledge is of greater *practical* value than mere true belief, with the Larissa story meant to be a reason for thinking that this might not be so. Indeed, notice that the solution that Plato sketches to this problem in the *Meno* – which appeals to the stability of knowledge as opposed to mere true belief – is explicitly meant to be a story which accounts for the practical value of knowledge. On this view, the reason why knowledge is better than mere true belief if this 'stability' story is correct has nothing to do with our epistemic goals – it is not guaranteed that stable true belief is better from an epistemic point of view than mere true belief – but everything to do with our practical goals. The thought is explicitly that we can do a lot more, from a practical point of view, with stable true belief than we can with mere true belief.

Plato's problem, then, is not our problem. Nevertheless, the fact that Plato is only concerned with the greater practical value of knowledge might

[7] This view has been expressed to me in conversation by both Michael Lynch and Mike Ridge (which is not to say, of course, that they would endorse such a view).

make us wonder whether it ought to be important to us to defend the specific claim that knowledge is better than mere true belief from a purely epistemic point of view. After all, one might argue, in a broadly Platonic fashion, that knowledge is generally of greater all-things-considered value than mere true belief whilst simultaneously granting that from a purely epistemic point of view there is no additional value to be had. Moreover, notice that one key advantage of this response to the problem is that one is able to retain epistemic value T-monism (and hence (1)), and hence respect the intuitions about belief and truth which, as we saw above, motivate that thesis.

If (3) were the strong claim that knowledge is *always* epistemically more valuable than mere true belief, then I think this sort of line would have a lot of plausibility. Given that we have seen that (3) is in fact rather the very weak claim that only *sometimes* is knowledge epistemically more valuable than mere true belief, however, it starts to look implausible. After all, consider the following scenario. Suppose that someone comes to you and says that in a moment one of two scenarios will obtain: either one will have a true belief that *p* or one will have knowledge that *p* (where one does not know which proposition is at issue). Furthermore, it is stipulated that all the practical consequences are kept fixed in both scenarios, so there will be no practical benefit to choosing the one option over the other. Nevertheless, shouldn't one choose knowledge rather than mere true belief? If knowledge is even sometimes better than mere true belief from an epistemic point of view, then this would account for this intuition. In contrast, those who defend the claim that (3) is simply false seem unable to account for this intuition at all.[8]

## 12.4 Denying epistemic value T-monism

Assuming that one accepts (2), however, which we are taking for granted here – and that one thinks that the argument is valid – then the only other premise that can be denied is (1). Interestingly, there are two very different ways of denying this premise. In particular, while any response to the swamping problem which denied (1) would be committed to denying epistemic value T-monism, it does not follow that in denying this thesis one is thereby endorsing epistemic pluralism.

[8] Kvanvig (2003) could be read as responding to the swamping problem by denying (3). As I note below, however, I think that ultimately he is best read as rejecting (1).

The quick way to see this is to note that there are different ways of being an epistemic value monist than being a T-monist. Consider, for example, someone like Timothy Williamson (e.g. 2000), who has explicitly argued for what he calls a 'knowledge-first' epistemology. Although (as far as I am aware), he has not endorsed such a view in print, presumably he would be very attracted, given his wider epistemological views, to an epistemic value monism of the following form:

*Epistemic value K-monism:*

Knowledge is the sole ultimate fundamental epistemic good.

According to this view, only knowledge is the ultimate fundamental epistemic good, which means that the value of all other epistemic goods is to be understood along instrumental lines relative to this fundamental good.

On this view, there is no reason to assent to (1), and thus the argument cannot even get started. Moreover, notice that one cannot straightforwardly run a swamping argument against this view. For what motivated the swamping argument described above was that we were comparing mere true belief with a more elevated epistemic standing, and asking the question of why the latter epistemic standing, from an epistemic point of view, should be more valuable. With epistemic value K-monism in play, however, the question of why knowledge is more valuable than mere true belief has a straightforward answer: because knowledge, unlike true belief, is a fundamental epistemic value.

Of course, in order to motivate such a position one would need to give it the further theoretical support that it requires. More specifically, one would also need to revisit the intuition about the relationship between belief and truth that we began with (and potentially also the conception of reasons for belief that goes with it). On this alternative view, presumably, the aim, in the relevant sense, of belief is not truth at all but knowledge. But, of course, Williamson, for one, would be quite happy with *that* claim.

But suppose that one does not wish to go in this direction. Indeed, suppose one is tempted to conclude, on the basis of this argument, that there is more than one fundamental epistemic value and hence denies (1) by endorsing epistemic value pluralism. What, now, are one's options?

To begin with, notice that simply opting for a form of epistemic value pluralism might not be enough by itself to rescue our intuition that knowledge is at least sometimes epistemically more valuable than mere true belief. Consider, for example, a form of epistemic value pluralism which said that

both true belief and a further epistemic property – let's call it 'X' – were fundamental epistemic goods. Simply arguing for such a position would not, however, be enough to ensure the truth of (3). After all, suppose that it were the case that, necessarily, when one knows a proposition, one's belief cannot exhibit the target epistemic property, X. Although this may be unlikely, this doesn't seem to be a possibility that we can rule out *a priori*, since it depends upon one's wider epistemological commitments (in particular, what theory of knowledge one endorses). If this were the case, however, then clearly that X is an epistemic property which, in addition to true belief, is fundamentally epistemically valuable would not suffice to ensure that knowledge is sometimes epistemically more valuable than mere true belief, and hence (3) would still be in question. The whole point of denying (1) in the face of this inconsistent triad, however, is surely so that one is able to consistently endorse (2) and (3).

Accordingly, what is required of this strategy is some way of denying (1) so that (3) gets the support that it requires. In particular, what is needed is a form of epistemic value pluralism which can account for the fact that knowledge is sometimes of greater epistemic value than mere true belief.

Consider, for example, the kind of epistemic value pluralism endorsed by Kvanvig (e.g. 2003). Kvanvig argues that there are several fundamental epistemic goods other than true belief, the chief example that he offers in this respect being understanding. Interestingly, however, Kvanvig does not offer his epistemic value pluralism as a means of rescuing the intuition that knowledge is epistemically more valuable than mere true belief, since he instead uses the swamping argument to demonstrate that we are mistaken in thinking that knowledge is a particularly valuable epistemic standing, as opposed to understanding. Accordingly, it seems that Kvanvig would respond to our inconsistent triad by rejecting *both* (1) and (3).[9]

I don't think that this can be the full story, however, for while Kvanvig argues that understanding is not a species of knowledge, he does not argue for the stronger claim that it is never the case that when one knows a proposition one's true belief also has the relevant epistemic property of understanding. Indeed, although the details of his view are too involved to be usefully expounded here, his conception of understanding and its relationship to knowledge in fact strongly suggests that it is quite often the case that when one has what he calls 'propositional' understanding,

[9] Of course, evaluations of this sort are inevitably moot given that Kvanvig wasn't responding to the swamping problem in the specific form in which it is presented here. What follows should be read with this caveat in mind.

one will also have knowledge of the corresponding proposition.[10] If this is right, though, then Kvanvig ought to be willing to accept (3) after all, and thereby accept that knowledge is at least sometimes epistemically more valuable than mere true belief, since in those cases where knowledge and understanding coincide, the known true belief in question will have an epistemic value which is greater than the epistemic value of the corresponding mere true belief. The ultimate problem posed by the swamping argument on Kvanvig's view, then, is thus not for (3) at all, but rather for epistemic value T-monism (i.e. (1)).

Nevertheless, I think that Kvanvig is on to something here when he seems to in effect treat the swamping argument as ultimately undermining both (1) and (3). For suppose that one argued for the negation of (1) by appeal to a form of epistemic value pluralism which treats knowledge as sometimes epistemically more valuable than mere true belief because of the fundamental epistemic value of an epistemic property which is only sometimes present when one has knowledge (such as understanding). Indeed, just to make this point particularly vivid, suppose that one argued for the negation of (1) by appeal to a form of epistemic value pluralism which treats knowledge as *always* epistemically more valuable than mere true belief because of the fundamental epistemic value of an epistemic property which is *always* present when one has knowledge, even though it is not sufficient, with true belief, for knowledge (i.e. the epistemic property in question is merely necessary for knowledge). Would we regard such a stance as demonstrating that knowledge has the kind of epistemic value that we typically suppose it to have?

I think not. In order to see this, notice that part of what is at issue in the wider debate about epistemic value is the central role that knowledge plays in epistemological enquiry. If it were to turn out, however, that knowledge is only of greater epistemic value than mere true belief because of the greater epistemic value of a necessary component of knowledge (still less, a non-necessary component), then that would surely threaten the central role that knowledge plays in epistemological theorizing almost as much as the claim that knowledge is *never* of greater epistemic value than mere true belief. After all, why should we now care whether we have knowledge, specifically,

[10] For example, consider one's true belief that one's house burned down because of faulty wiring. It seems entirely plausible that it could be the case that this belief both has the epistemic property of knowledge and the epistemic property of understanding (i.e. that one both knows and understands that one's house burned down because of faulty wiring). For more on Kvanvig's conception of understanding, see Brogaard (2005), Grimm (2006), Elgin (2009; cf. Kvanvig 2009a), Kvanvig (2009a, 2009b), Riggs (2009; cf. Kvanvig 2009a) and Pritchard (2009a; cf. Haddock, Millar, and Pritchard 2010: ch. 4). For more on the notion of understanding more generally, see Zagzebski (2001) and Riggs (2003).

rather than just true belief plus the extra fundamentally valuable epistemic property?[11]

Ideally, then, what we want is a defence of (3) which is able to support our wider intuition that knowledge is worthy of the central focus that it enjoys in epistemological theorizing. The most straightforward way of doing this within the model of epistemic pluralism would simply be to argue that knowledge is itself a fundamental epistemic value. The trouble with this suggestion, however, that it is hard to see why knowledge should be a *fundamental* epistemic value, rather than being an epistemic standing which is instrumentally epistemically valuable relative to the epistemic good of true belief.

## 12.5 THE RELEVANCE OF FINAL VALUE

Perhaps, though, there is another way to think about this problem. As I noted back in section 12.1 above, an epistemic value can be 'fundamental', as I am using the term, and yet not be finally (i.e. non-instrumentally) valuable. All that is required for it to qualify as fundamental is that it not be instrumentally valuable relative to any further epistemic good. Similarly, one could offer a form of epistemic value pluralism which was not committed to any final epistemic values. For example, perhaps the reason why both true belief and understanding are fundamental epistemic values is because they are instrumentally valuable relative to some further non-epistemic good. Relatedly, one could defend the thesis that knowledge is fundamentally epistemically valuable without being thereby committed to thinking that knowledge is finally valuable.

Equally, however, one might opt for a form of epistemic value pluralism which *did* appeal to final value. Perhaps, for example, understanding is the sort of thing that is good for its own sake, regardless of any further instrumental value that it might generate?[12] Similarly, perhaps the reason why knowledge is a fundamental epistemic value is precisely *because* it

[11] This problem is what I have referred to elsewhere as the 'secondary' value problem. While the primary value problem is the problem of showing that knowledge is more valuable than mere true belief, the secondary value problem is the problem of showing that knowledge is more valuable than any epistemic standing that falls short of knowledge. The reason why this is a genuine problem is essentially that just given in the main text. For if we simply claim that knowledge is epistemically more valuable than mere true belief because of its relationship to some further epistemic property, like understanding, then that leaves the question of why as epistemologists we ought to be focusing on knowledge completely unanswered. For more on the secondary value problem, see Pritchard 2007a, 2007b, 2008; cf. Haddock, Millar, and Pritchard 2010: ch. 1). As I note below – see footnote 13 – there is also, plausibly, a *tertiary* value problem as well.

[12] As it happens, I defend this very thesis in Pritchard (2009a; cf. Haddock, Millar, and Pritchard 2010: ch. 4).

is finally valuable? Indeed, those who wish to resist the swamping argument by appeal to epistemic value pluralism tend to precisely make a move of this sort. That is, while they grant that true belief is a (mere) fundamental epistemic good in the sense that I am using that term, they argue that knowledge is not just a fundamental epistemic good but also, more specifically, something which is finally valuable. Moreover, a further advantage of this strategy is that it can fully validate our intuition that knowledge should be at the centre of epistemological theorizing: if knowledge is the kind of thing which, unlike other lesser epistemic goods, is finally valuable, then it is no surprise that we have tended to focus on this notion in our epistemological enquiries to the exclusion of other epistemic standings.[13]

It is certainly not implausible to argue that knowledge has certain properties which ensure that it is epistemically valuable for its own sake. Indeed, the thought experiment described above – where we asked whether, keeping all other factors fixed (and in a state of ignorance about the relevant '*p*' under discussion), we should prefer knowledge that *p* over mere true belief that *p* – suggests just that, since it implies that even if we kept all the practical ramifications fixed, and even if we assume that the instrumental epistemic value is kept fixed, there is still something preferable about knowing a proposition rather than merely truly believing it. If knowledge were the kind of thing which, in certain cases at least, is finally valuable, then this would account for this intuition.

The best way that I know of to motivate the thesis that knowledge is finally valuable is by appeal to the idea that achievements are, at least sometimes if not typically or universally, finally valuable, and that knowledge, properly understood, is a kind of achievement. This is the kind of line pursued, in various different ways, by some virtue epistemologists, since on their view knowledge seems to have a structure such that one could reasonably argue that it falls under the more general class of achievements – i.e., that it is a specifically *cognitive* achievement.[14] After all, one plausible

[13] I noted above, in footnote 11, that one can distinguish between a 'primary' and a 'secondary' version of the problem of the value of knowledge. One can also delineate a 'tertiary' version of this problem: the problem of explaining why knowledge is more valuable than that which falls short of knowledge not just as a matter of degree but also as a matter of kind. It should be clear that the thesis just described regarding the final value of knowledge can potentially meet all three of these problems. For further discussion of the tertiary value problem, see Pritchard 2007a, 2007b, 2008; cf. Haddock, Millar, and Pritchard 2010: ch. 1).

[14] Sosa (1988) was, I believe, the first to offer a virtue-theoretic proposal along these lines. See also Sosa (1991, 2007). Versions of this sort of virtue-theoretic proposal can also be found in Zagzebski (e.g. 1996, 1999) and Greco (2002, 2007, 2008, 2009).

account of achievements is that they are successes that are because of one's ability. On standard virtue-theoretical accounts of knowledge, however, knowledge tends to be analysed, roughly, as true belief (i.e. cognitive success) that is because of cognitive ability, and so knowledge turns out to be just a cognitive type of achievement. Moreover, achievements do seem to be the kind of thing which are sometimes, if not typically or universally, of final value. Just as, all other things being equal, we would prefer knowledge over mere true belief, wouldn't we similarly prefer a success which constituted an achievement over a success which didn't constitute an achievement (which was merely lucky, say)? A natural explanation of both claims is that achievements and knowledge are finally valuable. On the virtue-theoretic account of knowledge this is no surprise, since knowledge is simply a type of achievement.[15]

Anyway, if one could show that knowledge has final value in this way, then that would be another way of dealing with the swamping problem, since one would then have grounds to deny (1). My personal view is that this strategy is not going to work for the simple reason that knowledge is not a type of achievement. In particular, I've argued elsewhere that there are cases of 'easy' knowledge in which one knows but where this does not exhibit a cognitive achievement in the relevant sense, and also cases in which one exhibits a cognitive achievement but does not know because cognitive achievements, unlike knowledge, are compatible with a certain kind of epistemic luck. Here is not the place to explore these objections further.[16] Note, however, that even if this form of epistemic value pluralism were rejected, this would not mean that all ways of responding to the swamping problem by appeal to epistemic value pluralism should be rejected, and still less would it mean that all ways of responding to the swamping problem by appeal to a version of epistemic value pluralism which made appeal to final value should be rejected.

## 12.6 Concluding remarks

It has not been my goal here to definitively argue for or against any particular line on the swamping problem but rather to provide a clear

[15] An embryonic version of the general idea that appeal to final value might be a way of responding to the swamping problem can be found in Percival (2003). For further discussion of this way of dealing with the swamping problem, see Brogaard (2006) and Pritchard (2007a, 2007b, 2008, 2009a, 2009b; cf. Haddock, Millar, and Pritchard 2010: ch. 1).

[16] For the main places in which I pursue this line, see Pritchard (2007a, 2007b, 2008, 2009a, 2009b; cf. Haddock, Millar, and Pritchard 2010, ch. 1).

picture of what the problem is and thus what dialectical options are available to us. In particular, I have argued that the swamping problem is best read as an inconsistent triad involving epistemic value T-monism and two further claims, one a general claim about value and another a more specific claim about epistemic value. With the swamping problem so understood, there are several avenues of escape available to us. In particular, given that we grant the general thesis about value formulated in (2), then either we simply deny the intuition that knowledge is sometimes epistemically more valuable than mere true belief or else we opt for a form of epistemic value pluralism. As regards the latter option, we have also seen that there are a number of versions of this style of response, some of them more compelling than others.

For my own part, although I haven't argued for this here, my view is that while we can resist the swamping argument by opting for a form of epistemic value pluralism – like Kvanvig (2003), I hold, for example, that understanding is a fundamental epistemic value, albeit on different grounds – it is an inevitable consequence of this style of response that knowledge, while sometimes epistemically more valuable than mere true belief, is not more valuable in such a way as to justify the preoccupation with this notion in contemporary epistemological theorizing. Thus, we evade the immediate problem posed by the swamping problem, and yet are still left with an epistemological revisionism of sorts, since it is now incumbent upon us to water down our love affair with knowledge and turn our attentions towards other epistemic standings.[17]

However we eventually opt to respond to the swamping problem, it should be clear from the foregoing that recognizing that this is a genuine problem, and moreover a problem which is specifically focused on the thesis of epistemic valule T-monism, will inevitably have an impact on one's conception of what properly constitutes a reason for belief (even if ultimately it leaves that conception intact). For as we noted above, a familiar view about what properly constitutes one's reasons for belief shares a common source with epistemic T-monism – *viz.*, the broad intuition that truth is the *telos* of belief. Any re-evaluation of epistemic value T-monism thus calls into question the underlying thinking which generated support for this thesis in the first place, and in doing so it also poses questions

[17] For the clearest statement of my position in this respect, see Haddock, Millar, and Pritchard (2010: chs. 1–4).

for the common, and closely related, account of what properly constitutes a reason for belief.[18]

[18] Thanks to Brit Brogaard, Campbell Brown, Matthew Chrisman, Pascal Engel, Alvin Goldman, Emma Gordon, Jon Kvanvig, Mike Lynch, Erik Olsson, Christian Piller, Wlodek Rabinowicz, Andrew Reisner, Asbjørn Steglich-Petersen, and Eusebio Waweru for discussion on topics related to this chapter. Special thanks to Adam Carter, Chris Kelp, Georgi Gardiner, Alan Millar, Mike Ridge, Wayne Riggs, and Ernie Sosa who offered extensive feedback on an earlier version of this piece, and to an anonymous referee for Cambridge University Press. My research on this topic was conducted while in receipt of a Philip Leverhulme Prize.

# *References*


Adler, J. 2002. *Belief's Own Ethics*. Cambridge, MA: MIT Press.

Alston, W. 1988. The Deontological Conception of Epistemic Justification. *Philosophical Perspectives* 2: 257–299.

2005. *Beyond Justification: Dimensions of Epistemic Evaluation*. Ithaca, NY: Cornell University Press.

Armstrong, D. 1973. *Belief, Truth, and Knowledge*. Cambridge University Press.

Audi, R. 1997. Moral Judgment and Reasons for Action. In *Ethics and Practical Reason* (ed. G. Cullity and B. Gaut). Cambridge University Press, 125–160.

2001. An Internalist Theory of Normative Grounds. *Philosophical Topics* 29: 19–46.

Bealer, G. 2000. A Theory of the A Priori. *Pacific Philosophical Quarterly* 81: 1–30.

Bird, A. 2007. Justified Judging. *Philosophy and Phenomenological Research* 74: 81–110.

Blackburn, S. 1998. *Ruling Passions*. New York: Oxford University Press.

Boghossian, P. 2002. How Are Objective Epistemic Reasons Possible? In *Reason and Nature: Essays in the Theory of Rationality* (ed. J. L. Bermúdez and A. Millar). Oxford University Press, 15–48.

2003a. Blind Reasoning. *Proceedings of the Aristotelian Society*, Supp. Vol. 77: 225–248.

2003b. The Normativity of Content. *Philosophical Issues* 13: 31–45.

BonJour, L. 1980. Externalist Theories of Empirical Knowledge. *Midwest Studies in Philosophy* 5: 53–74.

1985. *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.

Brady, M. S. and Pritchard, D. H. (eds.) 2003. *Moral and Epistemic Virtues*. Oxford: Blackwell.

Bratman, M. 1999. *Intentions, Plans, and Practical Reason*. Chicago: CSLI Publications.

Brewer, B. 1999. *Perception and Reason*. Oxford University Press.

2004. Realism and the Nature of Perceptual Experience. *Philosophical Issues* 14: 61–77.

2006. Perception and Content. *European Journal of Philosophy* 14(2): 165–181.

2008. How to Account for Illusion. In *Disjunctivism: Perception, Action, Knowledge* (ed. A. Haddock and F. Macpherson). Oxford University Press, 168–180.
Brogaard, B. 2005. I Know. Therefore, I Understand (MS).
2006. Can Virtue Reliabilism Explain the Value of Knowledge? *Canadian Journal of Philosophy* 36: 335–354.
Broome, J. 2000. Normative Requirements. *Ratio* 12: 398–419.
2001. Normative Practical Reasoning. *Proceedings of the Aristotelian Society*, Supp. Vol. 75: 175–193.
2004. Reasons. In *Reason and Value: Themes from the Moral Philosophy of Joseph Raz* (ed. R. J. Wallace, P. Pettit, S. Scheffler, and M. Smith). Oxford: Clarendon Press, 28–55.
2005. Does Rationality Give Us Reasons? *Philosophical Issues* 15: 321–337.
2007a. Does Rationality Consist in Correctly Responding to Reasons? *Journal of Moral Philosophy* 4: 349–374.
2007b. Is Rationality Normative? *Disputatio* 27: 161–178.
2007c. Wide or Narrow Scope? *Mind* 116: 359–370.
Burge, T. 1993. Content Preservation. *Philosophical Review* 102: 457–488.
2003. Perceptual Entitlement. *Philosophy and Phenomenological Research* 67: 503–548.
Bykvist, K. and Hattiangadi, A. 2007. Does Thought Imply Ought? *Analysis* 67: 277–285.
Byrne, A. 2005. Perception and Conceptual Content. In *Contemporary Debates in Epistemology* (ed. M. Steup and E. Sosa). Oxford: Blackwell, 231–250.
Campbell, J. 2002. *Reference and Consciousness*. Oxford University Press.
Cassam, Q. 2007. Ways of Knowing. *Proceedings of the Aristotelian Society* 107: 339–358.
Chen, C. 2006. Empirical Content and Rational Constraint. *Inquiry* 49(3): 242–264.
Child, W. 1994. *Causality, Interpretation and the Mind*. Oxford: Clarendon Press.
Chisholm, R. 1957. *Perceiving: A Philosophical Study*. Ithaca, NY: Cornell University Press.
Coady, C. A. J. 1992. *Testimony: A Philosophical Study*. Oxford: Clarendon Press.
Cohen, S. 1984. Justification and Truth. *Philosophical Studies* 46: 279–295.
Conee, E. and Feldman, R. 2001. Internalism Defended. *American Philosophical Quarterly* 38: 1–18.
2004. *Evidentialism*. Oxford University Press.
Cottingham, J. 2002. Descartes and the Voluntariness of Belief. *The Monist* 85(3): 343–360.
Cullity, G. and Gaut, B. (eds.) 1997. *Ethics and Practical Reason*. Oxford: Clarendon Press.
Cuneo, T. 2007. *The Normative Web*. Oxford University Press.
Dancy, J. 1993. *Moral Reasons*. Oxford: Blackwell.
2000. *Practical Reality*. Oxford University Press.

2004. Enticing Reasons. In *Reason and Value: Themes from the Moral Philosophy of Joseph Raz* (ed. R. J. Wallace, P. Pettit, S. Scheffler, and M. Smith). Oxford: Clarendon Press, 91–118.

Danielsson, S. and Olson, J. 2007. Brentano and the Buck-Passers. *Mind* 116: 511–522.

David, M. 2001. Truth as the Epistemic Goal. In *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Virtue, and Responsibility* (ed. M. Steup). Oxford University Press, 151–169.

2005. Truth as the Primary Epistemic Goal: A Working Hypothesis. In *Contemporary Debates in Epistemology* (ed. M. Steup and E. Sosa). Oxford: Blackwell, 296–312.

Davidson, D. 1983. A Coherence Theory of Truth and Knowledge. In *Kant oder Hegel* (ed. D. Henrich). Stuttgart: Klett-Cotta. Reprinted in *Truth and Interpretation*: *Perspectives on the Philosophy of Donald Davidson* (ed. E. LePore). Oxford: Blackwell, 307–319.

Dreier, J. 2002. The Expressivist Circle: Invoking Norms in the Explanation of Normative Judgement. *Philosophy and Phenomenological Research* 65: 136–143.

Dretske, F. 1969. *Seeing and Knowing*. London: Routledge & Kegan Paul.

1970. Epistemic Operators. *Journal of Philosophy* 67: 1007–1023. Reprinted in Dretske, 2000: 30–47.

1971. Conclusive Reasons. *Australasian Journal of Philosophy* 49: 1–22. Reprinted in Dretske, 2000: 3–29.

2000. *Perception, Knowledge and Belief: Selected Essays*. Cambridge University Press.

Elgin, C. 2009. Is Understanding Factive? In *Epistemic Value* (ed. A. Haddock, A. Millar, and D. H. Pritchard). Oxford University Press, 322–330.

Evans, G. 1982. *The Varieties of Reference*. Oxford: Clarendon Press.

Fantl, J. and McGrath, M. 2002. Evidence, Pragmatics, and Justification. *Philosophical Review* 111: 67–94.

Feldman, R. 1988a. Epistemic Obligations. *Philosophical Perspectives* 2: 235–256.

1988b. Having Evidence. Reprinted in *Evidentialism* (E. Conee and R. Feldman). Oxford University Press, 2004, 219–241.

1988c. Subjective and Objective Justification in Ethics and Epistemology. *The Monist* 71: 405–419.

Finlay, S. 2008. The Error in the Error Theory. *Australasian Journal of Philosophy* 86: 347–369.

Fletcher, G. 2009. Uneasy Companions. *Ratio* 22: 359–368.

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

Foot, P. 1972. Morality as a System of Hypothetical Imperatives. *Philosophical Review* 91: 305–317.

Fricker, M. 2003. Epistemic Justice and a Role for Virtue in the Politics of Knowing. *Metaphilosophy* 34: 154–173.

Fumerton, R. 1995. *Metaepistemology and Skepticism*. Lanham, MA: Rowman and Littlefield.

2006. The Epistemic Role of Testimony: Internalist and Externalist Perspectives. In *The Epistemology of Testimony* (ed. J. Lackey and E. Sosa). Oxford: Clarendon Press, 77–92.
Gardner, J. 2005. Wrongs and Faults. *The Review of Metaphysics* 59: 95–132.
2007. *Offences and Defences*. New York: Oxford University Press.
Gettier, E. 1963. Is Justified True Belief Knowledge? *Analysis* 23: 121–123.
Gibbard, A. 1990. *Wise Choices, Apt Feelings*. Oxford: Clarendon Press.
Gibbons, J. Forthcoming. Things That Make Things Reasonable. *Philosophy and Phenomenological Research*.
Ginsborg, H. 2006a. Aesthetic Judgment and Perceptual Normativity. *Inquiry* 49(5): 403–437.
2006b. Empirical Concepts and the Content of Experience. *European Journal of Philosophy* 14(3): 349–372.
2006c. Reasons for Belief. *Philosophy and Phenomenological Research* 72(2): 286–318.
Forthcoming. Primitive Normativity and Skepticism about Rules. *Journal of Philosophy.*
Gluer, K. and Wikforss, Å. 2009. Against Content Normativity. *Mind* 118: 31–70.
Goldman, A. 1976. Discrimination and Perceptual Knowledge. *Journal of Philosophy* 73: 771–791.
1979. What is Justified Belief? In *Justification and Knowledge* (ed. George Pappas). Dordrecht: Reidel, 1–23.
1999a. Internalism Exposed. *Journal of Philosophy* 96: 271–293.
1999b. *Knowledge in a Social World*. New York: Oxford University Press.
Greco, J. 2002. Knowledge as Credit for True Belief. In *Intellectual Virtue: Perspectives from Ethics and Epistemology* (ed. M. DePaul and L. Zagzebski). Oxford University Press, 111–134.
2007. The Nature of Ability and the Purpose of Knowledge. *Philosophical Issues* 17: 57–69.
2008. What's Wrong With Contextualism? *Philosophical Quarterly* 58: 416–436.
2009. *The Value Problem.* In *Epistemic Value* (ed. A. Haddock, A. Millar, and D. H. Pritchard). Oxford University Press, 313–321.
Grimm, S. 2006. Is Understanding a Species of Knowledge? *British Journal for the Philosophy of Science* 57: 515–536.
2009. Epistemic Normativity. In *Epistemic Value* (ed. A. Haddock, A. Millar, and D. H. Pritchard). Oxford University Press, 243–264.
Gupta, A. 2006. Experience and Knowledge. In *Perceptual Experience* (ed. T. S. Gendler and J. Hawthorne). Oxford University Press, 181–204.
Haddock, A., Millar, A., and Pritchard, D. H. 2010. *The Nature and Value of Knowledge: Three Investigations*. Oxford University Press.
Hampton, J. 1998. *The Authority of Reason.* Cambridge University Press.
Harman, G. 1990. The Intrinsic Quality of Experience. *Philosophical Perspectives* 4: 31–52.
1999. *Reasoning, Meaning, and Mind.* Oxford University Press.
Hattiangadi, A. 2007. *Oughts and Thoughts.* Oxford University Press.

Hawthorne, J. and Stanley, J. 2008. Knowledge and Action. *Journal of Philosophy*, 105: 571–590.
Heathwood, C. 2009. Moral and Epistemic Open Question Arguments. *Philosophical Books* 50: 83–98.
Heck, R. 2000. Nonconceptual Content and the "Space of Reasons." *Philosophical Review* 109(4): 483–523.
Herman, B. 1994. *The Practice of Moral Judgment*. Cambridge, MA: Harvard University Press.
Hieronymi, P. 2005. The Wrong Kind of Reasons. *Journal of Philosophy* 102: 37–57.
2009. Two Kinds of Agency. In *Mental Actions* (ed. L. O'Brien and M. Soteriou). Oxford University Press, 138–162.
Horder, J. 2004. *Excusing Crime*. New York: Oxford University Press.
Huemer, M. 2007. Moore's Paradox and the Norm of Belief. In *Themes from G. E. Moore: New Essays in Epistemology and Ethics* (ed. S. Nuccetelli and G. Seay). New York: Oxford University Press, 142–157.
Husak, D. 1999. Conflicts of Justification. *Law and Philosophy* 18: 41–68.
Jackson, F. 1999. Non-Cognitivism, Normativity, Belief. *Ratio* 12(4): 420–435.
Johnston, M. 2006. Better than Mere Knowledge? The Function of Sensory Awareness. In *Perceptual Experience* (ed. T. S. Gendler and J. Hawthorne). Oxford University Press, 260–290.
Jones, W. 1997. Why Do We Value Knowledge? *American Philosophical Quarterly* 34: 423–440.
Joyce, R. 2001. *The Myth of Morality*. Cambridge University Press.
Forthcoming. The Error in "The Error in the Error Theory." *Australasian Journal of Philosophy*.
Kant, I. 1787. *Kritik der reinen Vernunft*, 2nd (B) edition. Riga: Hartknoch.
Karlander, K. 2008. *The Normativity of Thought and Meaning*. Doctoral dissertation. Stockholm University.
Kasser, J. and Shah, N. 2006. The Metaethics of Belief: An Expressivist Reading of "The Will to Believe." *Social Epistemology* 20(1): 1–17.
Kavka, G. 1983. The Toxin Puzzle. *Analysis* 43: 33–36.
Kelly, T. 2003. Epistemic Rationality as Instrumental Rationality: A Critique. *Philosophy and Phenomenological Research* 66: 612–640.
2006. Evidence. *The Stanford Encyclopedia of Philosophy* (ed. E. N. Zalta). http://plato.stanford.edu/entries/evidence/#EviGuiTruEviSigSymMar
2007. Evidence and Normativity: Reply to Leite. *Philosophy and Phenomenological Research* 75: 465–474.
Kim, J. 1988. What is Naturalized Epistemology? *Philosophical Perspectives* 2: 381–405.
Kolodny, N. 2005. Why Be Rational? *Mind* 114(455): 509–563.
2007. State or Process Requirements? *Mind* 116(462): 371–385.
Kornblith, H. 1993. Epistemic Normativity. *Synthese* 94: 357–378.
Kvanvig, J. 2003. *The Value of Knowledge and the Pursuit of Understanding*. Cambridge: Cambridge University Press.

2005. Truth is Not the Primary Epistemic Goal. In *Contemporary Debates in Epistemology* (ed. M. Steup and E. Sosa). Oxford: Blackwell, 285–296.
2009a. Responses to Critics. In *Epistemic Value* (ed. A. Haddock, A. Millar, and D. H. Pritchard). Oxford University Press, 339–351.
2009b. The Value of Understanding. In *Epistemic Value* (ed. A. Haddock, A. Millar, and D. H. Pritchard). Oxford University Press, 95–111.
Lackey, J. 2007. The Norms of Assertion. *Noûs* 41: 594–626.
Lackey, J. and Sosa, E. (eds.) 2006. *The Epistemology of Testimony*. Oxford: Clarendon Press.
Langsam, H. 2008. Rationality, Justification, and the Internalism/Externalism Debate. *Erkenntnis* 68: 79–101.
Lehrer, K. 1990. *Theory of Knowledge*. Boulder, CO: Westview Press.
Lehrer, K. and Paxson, T. 1969. Knowledge: Undefeated Justified True Belief. *Journal of Philosophy* 66(4): 225–237.
Leite, A. 2005. A Localist Solution to the Regress of Epistemic Justification. *Australasian Journal of Philosophy* 83: 395–421.
2007. Epistemic Instrumentalism and Reasons for Belief: A Reply to Tom Kelly's "Epistemic Rationality as Instrumental Rationality: A Critique." *Philosophy and Phenomenological Research* 75: 456–464.
2008. Believing One's Reasons Are Good. *Synthese* 161(3): 419–441.
Lewis, C. I. 1946. *An Analysis of Knowledge and Valuation*. LaSalle, IL: Open Court.
Lillehammer, H. 2007. *Companions in Guilt*. New York: Palgrave Macmillan.
Littlejohn, C. The Myth of the False, Justified Belief (MS).
2009a. The Externalist's Demon. *Canadian Journal of Philosophy* 39: 399–435.
2009b. Must We Act Only On What We Know? *Journal of Philosophy* 106: 463–474.
2010. Moore's Paradox and Epistemic Norms. *Australasian Journal of Philosophy* 88: 79–100.
Forthcoming. Evidence and Armchair Access. *Synthese*.
Lynch, M. 2004. *True to Life: Why Truth Matters*. Cambridge, MA: MIT Press.
Mackie, J. L. 1977. *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin Books.
Martin, M. G. F. 2002. The Transparency of Experience. *Mind and Language* 17(4): 376–425.
2004. The Limits of Self-Awareness. *Philosophical Studies* 120(1–3): 37–89.
McDowell, J. 1982. Criteria, Defeasibility, and Knowledge. *Proceedings of the British Academy* 68: 455–479. Reprinted in McDowell, 1998.
1994a. Knowledge by Hearsay. In *Knowing from Words: Western and Indian Philosophical Analyses of Understanding and Testimony* (ed. B. K. Matilal and A. Chakrabarti). Dordrecht: Kluwer, 195–224. Reprinted in McDowell, 1998.
1994b. *Mind and World*. Oxford University Press.
1998. *Meaning, Knowledge, and Reality*. Cambridge, MA: Harvard University Press.

Millar, A. 1991. *Reasons and Experience*. Oxford: Clarendon Press.
2004. *Understanding People: Normativity and Rationalizing Explanation*. Oxford: Clarendon Press.
2007. What the Disjunctivist is Right About. *Philosophy and Phenomenological Research* 74: 176–198.
2008a. Disjunctivism and Skepticism. In *The Oxford Handbook of Skepticism* (ed. J. Greco). New York: Oxford University Press, 581–604.
2008b. Perceptual-Recognitional Abilities and Perceptual Knowledge. In *Disjunctivism: Perception, Action, Knowledge* (ed. A. Haddock and F. E. Macpherson). Oxford: Clarendon Press, 330–347.
2009. What is it that Cognitive Abilities are Abilities to Do? *Acta Analytica* 24: 223–236.
2010. Knowing from Being Told. In *Social Epistemology* (ed. A. Haddock, A. Millar, and D. Pritchard). Oxford University Press, 175–193.
Mitova, V. 2009. A Quasi-Pragmatist Explanation of Our Ethics of Belief. *Teorema* 28(3): 113–130.
Moran, R. 2006. Getting Told and Being Believed. In *The Epistemology of Testimony* (ed. J. Lackey and E. Sosa). Oxford: Clarendon Press, 272–306.
Moser, P. 1989. *Knowledge and Evidence*. New York: Cambridge University Press.
Nagel, T. 1979. *Mortal Questions*. New York: Cambridge University Press.
Nozick, R. 1974. *Anarchy, State, and Utopia*. Cambridge, MA: Harvard University Press.
Oddie, G. 2005. *Value, Reality, and Desire*. Oxford University Press.
Olson, J. 2009. Reasons and the New Non-Naturalism. In *Spheres of Reason* (ed. S. Robertson). Oxford University Press, 164–182.
2010. In Defence of Moral Error Theory. In *New Waves in Metaethics* (ed. M. Brady). Basingstoke: Palgrave Macmillan, 62–84.
Owens, D. 2000. *Reason Without Freedom*. New York: Routledge.
Parfit, D. 2001. Reasons and Rationality. In *Exploring Practical Rationality* (ed. D. Egonsson, J. Josefsson, B. Petersson, and T. Rønnow-Rasmussen). Aldershot: Ashgate, 17–39.
Peacocke, C. 2001. Does Perception Have a Nonconceptual Content? *Journal of Philosophy* 98: 239–264.
2004. *The Realm of Reason*. Oxford: Clarendon Press.
Percival, P. 2003. The Pursuit of Epistemic Good. *Metaphilosophy* 34: 29–47. Reprinted in *Moral and Epistemic Virtues* (ed. M. S. Brady and D. H. Pritchard). Oxford: Blackwell, 2003, 29–46.
Pollock, J. 1974. *Knowledge and Justification*. Princeton University Press.
1986. *Contemporary Theories of Knowledge*. Lanham, MD: Rowman and Littlefield.
Pollock, J. and Cruz, J. 1999. *Contemporary Theories of Knowledge*, 2nd edition. New York: Rowman and Littlefield.
Price, H. H. 1932. *Perception*. London: Methuen.
Pritchard, D. H. 2007a. Knowledge, the Value of. *Stanford Encyclopedia of Philosophy* (ed. E. Zalta). http://plato.stanford.edu/entries/knowledge-value/

2007b. Recent Work on Epistemic Value. *American Philosophical Quarterly*, 44: 85–110.
2008. Greco on Knowledge: Virtues, Contexts, Achievements. *Philosophical Quarterly* 58: 437–447.
2009a. Knowledge, Understanding and Epistemic Value. In *Epistemology (Royal Institute of Philosophy Lectures)* (ed. A. O'Hear). Cambridge University Press, 19–44.
2009b. The Value of Knowledge. *Harvard Review of Philosophy* 16: 2–19.
Pritchard, D., Millar, A., and Haddock, A. 2010. *The Nature and Value of Knowledge: Three Investigations*. Oxford University Press.
Pryor, J. 2000. The Skeptic and the Dogmatist. *Noûs* 34: 517–549.
2004. What's Wrong with Moore's Argument? *Philosophical Issues* 14: 349–378.
2005. There is Immediate Justification. In *Contemporary Debates in Epistemology* (ed. M. Steup and E. Sosa). Oxford: Blackwell, 181–202.
Forthcoming. When Warrant Transmits. In *Wittgenstein, Epistemology and Mind: Themes from the Philosophy of Crispin Wright* (ed. A. Coliva). Oxford University Press.
Putnam, H. 1990. Beyond the Fact/Value Dichotomy. In Putnam, *Realism with a Human Face* (ed. J. Conant). Cambridge University Press, 135–141.
Rabinowicz, W. and Rønnow-Rasmussen, T. 2004. The Strike of the Demon: On Fitting Pro-Attitudes and Value. *Ethics* 114: 391–423.
Railton, P. 1997. On the Hypothetical and Non-Hypothetical in Reasoning about Belief and Action. In *Ethics and Practical Reason* (ed. G. Cullity and B. Gaut). Cambridge University Press, 53–80.
Raz, J. 1975. *Practical Reason and Norms*. London: Hutchinson.
1986. *The Morality of Freedom*. Oxford University Press.
1990. *Practical Reason and Norms*. New York: Oxford University Press.
1999. *Engaging Reason*. New York: Oxford University Press.
2005. The Myth of Instrumental Rationality. *Journal of Ethics and Social Philosophy* 1(1): 2–28.
Reisner, A. 2004. Conflicts of Normativity. D.Phil thesis, Oxford University.
2008. Weighing Pragmatic and Evidential Reasons for Belief. *Philosophical Studies* 138: 17–27.
2009a. The Possibility of Pragmatic Reasons for Belief and the Wrong Kind of Reason Problem. *Philosophical Studies* 145(2): 257–272.
2009b. Unifying the Requirements of Rationality. *Philosophical Explorations* 12(3): 243–260.
Riggs, W. 2002a. Beyond Truth and Falsehood: The *Real* Value of Knowing that *p*. *Philosophical Studies* 107: 87–108.
2002b. Reliability and the Value of Knowledge. *Philosophy and Phenomenological Research* 64: 79–96.
2003. Understanding "Virtue" and the Virtue of Understanding. In *Intellectual Virtue: Perspectives from Ethics and Epistemology* (ed. L. Zagzebski and M. DePaul). Oxford University Press, 203–226.

2007. The Value Turn in Epistemology. In *New Waves in Epistemology* (ed. V. Hendricks and D. H. Pritchard). Basingstoke: Palgrave Macmillan, 300–323.
2009. Understanding, Knowledge, and the Meno Requirement. In *Epistemic Value* (ed. A. Haddock, A. Millar, and D. H. Pritchard). Oxford University Press, 331–338.
Robinson, P. 1996. Competing Theories of Justification: Deeds v. Reasons. In *Harm and Culpability* (ed. A. Simester and A. Smith). New York: Oxford University Press, 45–70.
Scanlon, T.M. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
Schroeder, M. 2004. The Scope of Instrumental Reason. *Philosophical Perspectives* 18: 337–364.
2007. *Slaves of the Passions*. Oxford University Press.
2008. Having Reasons. *Philosophical Studies* 139(1): 57–71.
Forthcoming. Stakes, Withholding, and the Subject-Sensitivity of Knowledge. *Philosophical Studies*.
Shah, N. 2003. How Truth Governs Belief. *Philosophical Review* 112: 447–483.
2006. A New Argument for Evidentialism. *Philosophical Quarterly* 56: 481–498.
2008. How Action Governs Intention. *Philosophers' Imprint* 8: 1–19.
Shah, N. and Velleman, D. 2005. Doxastic Deliberation. *Philosophical Review* 114: 497–534.
Smith, M. 1994. *The Moral Problem*. Oxford: Blackwell.
2001. Some Not-Much-Discussed Problems for Non-Cognitivism in Ethics. *Ratio* 14(2): 93–115.
Sobel, D. and Wall, S. (eds.) 2009. *Reasons for Action*. Cambridge University Press.
Sobel, J. H. Good and Gold: A Judgmental History of Metaethics from G. E. Moore through J. L. Mackie (MS). www.utsc.utoronto.ca/~sobel/Gd_Gld/
Sosa, E. 1988. Beyond Skepticism, to the Best of Our Knowledge. *Mind* 97: 153–189.
1991. *Knowledge in Perspective: Selected Essays in Epistemology*. Cambridge University Press.
2007. *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Volume I*. Oxford University Press.
Speaks, J. 2005. Is There a Problem about Nonconceptual Content? *Philosophical Review* 114(3): 359–398.
Steglich-Petersen, A. 2006a. No Norm Needed: On the Aim of Belief. *Philosophical Quarterly* 56: 499–516.
2006b. Voluntarism and Transparent Deliberation. *South African Journal of Philosophy* 25: 171–176.
2008a. Against Essential Normativity of the Mental. *Philosophical Studies* 140: 263–283.
2008b. Does Doxastic Transparency Support Evidentialism? *Dialectica* 62: 541–547.
2009. Weighing the Aim of Belief. *Philosophical Studies* 145: 395–405.
Steup, M. 1999. A Defense of Internalism. In *The Theory of Knowledge: Classical and Contemporary Readings* (ed. L. Pojman). Belmont, CA: Wadsworth, 373–384.

2001. Epistemic Duty, Evidence, and Internality. In *Knowledge, Truth, and Duty* (ed. M. Steup). New York: Oxford University Press, 134–148.
2008. Doxastic Freedom. *Synthese* 161: 375–392.
Stich, S. 1990. *The Fragmentation of Reason: Preface to a Pragmatic Theory of Cognitive Evaluation*. Cambridge, MA: MIT Press.
Stratton-Lake, P. 2002. Introduction. In D. Ross, *The Right and the Good* (ed. P. Stratton-Lake). Oxford: Clarendon Press, ix–xlix.
Strawson, P. F. 1962. Freedom and Resentment. *Proceedings of the British Academy* 48: 1–25.
1979. Perception and Its Objects. In *Perception and Identity: Essays presented to A. J. Ayer with His Replies to Them* (ed. G. F. Macdonald). London: Macmillan, 41–60.
Stroud, B. 2004. Perceptual Knowledge and Epistemological Satisfaction. In *Ernest Sosa and His Critics* (ed. J. Greco). Oxford: Blackwell, 164–173.
Sutton, J. 2005. Stick to What You Know. *Noûs* 39: 359–396.
2007. *Without Justification*. Cambridge, MA: MIT Press.
Sverdlik, S. 1996. Motive and Rightness. *Ethics* 106: 327–349.
Swinburne, R. 1999. *Providence and the Problem of Evil*. Oxford University Press.
2000. *Epistemic Justification*. Oxford University Press.
Textor, M. 2009. Are Particulars or States of Affairs Given in Perception? In *States of Affairs* (ed. M. Reicher). Frankfurt/Main: Ontos, 129–150.
Travis, C. 2004. The Silence of the Senses. *Mind* 113: 57–94.
2005. A Sense of Occasion. *Philosophical Quarterly* 55: 286–314.
Vahid, H. 2006. Aiming at Truth: Doxastic vs. Epistemic Goals. *Philosophical Studies* 131: 303–335.
Velleman, D. 2000. *The Possibility of Practical Reason*. New York: Oxford University Press.
Wallace, J. 2001. Normativity, Commitment, and Instrumental Reason. *Philosophers' Imprint* 1: 1–26.
Wallace, R. J., Pettit, P., Scheffler, S., and Smith, M. (eds.) 2004. *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*. Oxford: Clarendon Press.
Wedgwood, R. 1999. The A Priori Rules of Rationality. *Philosophy and Phenomenological Research* 59: 113–131.
2002a. The Aim of Belief. *Philosophical Perspectives* 16: 267–297.
2002b. Internalism Explained. *Philosophy and Phenomenological Research* 65: 349–369.
2003. Choosing Rationally and Choosing Correctly. In *Weakness of Will and Practical Irrationality* (ed. S. Stroud and C. Tappolet). Oxford University Press, 201–229.
2007a. *The Nature of Normativity*. Oxford: Clarendon Press.
2007b. Normativism Defended. In *Contemporary Debates in Philosophy of Mind* (ed. B. McLaughlin and J. Cohen). Oxford: Blackwell, 85–101.
Williams, B. 1973. *Problems of the Self*. New York: Cambridge University Press.
1980. Internal and External Reasons. *Reprinted in Moral Luck: Philosophical Papers, 1973–1980*. Cambridge University Press, 1981, 101–113.

1995. Internal Reasons and the Obscurity of Blame. In *Making Sense of Humanity*. Cambridge University Press, 35–45.
Williamson, T. 1995. Is Knowledge a State of Mind? *Mind* 104: 533–565.
2000. *Knowledge and its Limits*. Oxford University Press.
2007. *The Philosophy of Philosophy*. Oxford: Blackwell.
Wright, C. 2001. On Basic Logical Knowledge. *Philosophical Studies* 106: 41–85.
2002. (Anti-)Sceptics Simple and Subtle: Moore and McDowell. *Philosophy and Phenomenological Research* 65: 330–348.
2003. Some Reflections on the Acquisition of Warrant by Inference. In *New Essays on Semantic Externalism and Self-Knowledge* (ed. S. Nuccetelli). Cambridge, MA: MIT Press, 57–77.
2004. Warrant for Nothing (and Foundations for Free)? *Aristotelian Society Supplement* 78: 167–212.
2007. The Perils of Dogmatism. In *Themes from G. E. Moore: New Essays in Epistemology and Ethics* (ed. S. Nuccetelli and G. Seay). Oxford University Press, 25–48.
Zagzebski, L. 1996. *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge University Press.
1999. What is Knowledge? In *The Blackwell Guide to Epistemology* (ed. J. Greco and E. Sosa). Oxford: Blackwell, 92–116.
2001. Recovering Understanding. In *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Virtue, and Responsibility* (ed. M. Steup). Oxford University Press, 235–252.
2003. The Search for the Source of Epistemic Good. *Metaphilosophy* 34: 12–28. Reprinted in *Moral and Epistemic Virtues* (ed. M. S. Brady and D. H. Pritchard). Oxford: Blackwell, 13–28.
Zimmerman, M. 2002. Taking Luck Seriously. *Journal of Philosophy* 99: 553–576.

# *Index*


*a priori*, 192, 193, 194, 197, 199–200
anti-Humeanism, 56, 58, 67–69, 70–73

basing, 166, 167, 173, 179, 220, 226, 232
Bayesian belief revision, 87
belief
  aim of, 123, 244
  ascription of, 6, 91–92, 100, 102
  correctness of, 6, 95, 100, 181
  degree of, 26, 182
  dispositional, 165, 167
  involuntariness of, 57, 244
  norm of, 121, 126, 127, 128, 129
  normativism, 91, 92
  occurrent, 50, 165, 179
belief, formation of, 8, 15, 20–22, 23–27, 29–33, 58, 66, 95, 136, 141, 143–145, 180–186, 190–192, 196–199, 226, 244
blocked ascent argument, 44
BonJour, L., 191
bootstrapping problem, 38
Bratman, M., 38
Brewer, B., 132, 133, 134, 138, 149
Broome, J., 16, 34

Campbell, J., 134
cognitive achievement, 256, 257
cognitivism, 5, 54–55, 56, 58, 72, 73, 99
coherentism, 9, 143, 149, 214, 220, 221
companions in guilt, 75
compliance, 113, 116
conative state, 56, 58, 59, 70, 99
conformity, 6, 112, 116
connative states, 6
constructivism, 105–106
content
  conceptual, 7, 131, 135, 139, 196
  experiential, 133
  non-conceptual, 7, 131, 139, 182
context sensitivity, 2
credence, 69, 99
Cuneo, T., 79, 82
curiosity, 21, 59, 60, 65, 66

Dancy, J., 22
Davidson, D., 65, 141
disenabling condition, 163
dogmatism, 158, 170, 183
doxastic deliberation, 95, 103, 104, 121, 122, 158, 161, 165
duty, 21, 114

empiricism, 214
enabling condition, 163
enkrasia, 51
enquiry, 20
epistemic responsibility, 7–8, 90, 164, 173, 176
epistemology
  normative, viii
error theory, 5–6, 54, 75, 77, 81–86, 87–93, 96–103, 105–107, 218
evidence, 2, 7, 13, 18, 22–24, 30–33, 34, 47–48, 50, 52, 68–69, 78–79, 85–88, 91–92, 94–99, 101, 106, 121, 122–126, 162, 167, 201–203, 205, 207–208, 211, 214–221, 229–230, 239–240
  factoring account, 203
  two relations account, 203
evidentialism, 25, 98, 112, 121–123, 124–126, 128
exclusivity problem, 13, 25
excuses, 111, 115, 125
experience, 7, 9, 131–133, 134–135, 136–138, 142–143, 145, 149–152, 154, 156, 158, 163, 170–171, 181–185, 186–189, 192–193, 215–218, 225–228, 235, 236, 237, 241
externalism, 9, 35, 55–56, 58–59, 67, 69–70, 73, 214, 220, 221

Foot, P., 79
foundationalism, 214, 215, 220
functionalism, 159

Gettier cases, 127, 227, 228
Goldman, A., 18
grounding, 136

Humeanism, 56, 58, 67–69

indicator evidence, 87
indicator phenomena, 9, 225, 230
inference to the best explanation, 87, 102, 196
inferential internalism, 8
intention, 38, 59, 62, 64, 65, 123, 195
internalism, 55–56, 58–59, 61–62, 66–67, 69–70, 72–73, 162, 169, 185–187, 190, 214

justification
  deontological theory of, 113
  doxastic, 168–169, 171, 172–174, 175–177, 179, 206–207
  epistemic, 2–3, 6–7

Kant, I., 102–103, 193
Kelly, T., 18, 30, 78, 86, 87
knowledge, 2, 7, 9–10, 121, 126–128, 201, 211, 216, 220, 223–225, 228, 230–234, 238–243, 246, 247–252, 253–256
knowledge norm, 126
Kvanvig, J., 246, 253

Lynch, M., 19

Mackie, J.L., 81
McDowell, J., 133, 135, 139, 143
metaepistemology, 83, 84
metaethics, viii, 4, 54–55, 57, 59, 67, 72, 74, 84
motivation
  epistemic, 4, 55, 58, 59, 69, 71, 72
  moral, 4–5

naturalism, 75–76, 88
non-cognitivism, 5, 56, 58, 67, 73, 99
normative force, 78, 123
normative judgement, 54–55
normative judgment, 6
normative requirement, 34, 40–43, 45–46, 53
normativity, 34–38, 42, 47, 49, 51–53, 152, 155
  epistemic, 54, 57, 61, 79–80, 125
  theoretical, 49
norms
  immanent, 76, 80–81, 84, 90, 92, 93
  institutional, 78
  transcendent, 76, 80–81, 82, 84, 87, 89–90, 92, 93

oughts, 17, 34–35, 38, 39, 67
  defeasible, 43, 46
  narrow-scope, 36–38
  wide-scope, 28, 37–38, 46

Parfit, D., 35
Peacocke, C., 139
perception, 7, 9–10, 72, 133–149, 152, 153–156, 214–221, 225, 228, 239
perceptual knowledge, 9, 10, 225, 228, 229, 233, 236, 237, 238, 239, 243
practical deliberation, 125
practical philosophy, viii, 1–2
pragmatism, 94, 96–97, 100
presentational attitude, 204, 221
propositional attitude, 1, 8, 41, 43, 51, 52, 92, 244
Pryor, J., 158, 171

Railton, P., 50, 78
rationalism, 214
rationality, 1, 8, 34–36, 39, 41–42, 50, 53, 184–185, 192–194
  descriptive theories of, 39
  dispositional account of, 43
  norms of, 35
  practical, 22, 36, 38, 49, 51–53
  requirements of, 34–35, 37, 39, 41
  theoretical, 22, 36, 38, 46–47, 49–53
Raz, J., 69
reasoning, 1, 7–8, 122, 158–159, 161, 165–167, 168–169, 171–174, 177–179, 180, 210–211, 223, 236
reasons
  activity-dependent, 83, 84
  all things considered, 15, 16, 23, 24, 25, 26, 28, 29, 31, 32, 33, 115, 163, 249, 250, 251
  attributions of, 96, 101, 107
  categorical, 76, 79, 80, 86
  epistemic, 3–4, 14–15, 17–20, 22–27, 29–33, 46–48, 49–50, 54, 75–79, 81–85, 88, 120, 161–162, 166
  explanatory, 1, 116, 119
  for action, viii, 1–2, 17, 180, 202
  for belief, viii–ix, 2–4, 6, 7, 9–10, 13–14, 17–18, 20–21, 23, 25–26, 75, 82, 91, 93, 96–102, 146, 180, 240
  hypothetical, 83, 84

instrumental, 4, 15, 27–29, 31, 33
intersubjective, 30
moral, 4, 75–76, 81–82, 88
normative, viii, 1–4, 5, 6–7, 34, 35, 40, 48, 50, 119, 161
object-given, 40, 43, 44, 46, 52
practical, 18, 21–22, 122
pro tanto, 16, 17, 26
role-dependent, 83
state-given, 40, 41, 43, 44, 45, 46
to believe, 15, 22, 23–25, 57, 58, 69, 83–86, 90–92, 97, 121–122, 123, 223–224, 236
reliabilism, 159, 161, 169, 185, 220, 246–247
representational content, 7, 131–132, 135, 139–140, 144, 146, 149, 151, 154–155, 181–183, 189

Scanlon, T., 13, 141
scepticism, 7–8, 82, 158, 159, 170, 171, 175
Stratton-Lake, P., 82
subjective probability, 99
swamping problem, 10, 246, 248, 250, 251, 253, 257, 258

teleology, 3, 13–15, 17–20, 27, 33, 70–72
testimony, 223–224, 229–230, 234, 236, 239, 241
transparency, 123
truth-aptness, 5
truth norm, 96, 121–125, 128–129

value, 1, 3–4, 10, 13–15, 17–22, 30, 54, 245–246, 247–258
epistemic, 58
intrinsic, 3, 13, 14
value problem, 14–15, 33
Velleman, D., 70
virtue, 54, 58, 67, 256

warrant, viii, 2, 7–8, 19, 158–179
immediate, 7, 158, 160, 169–171, 172–176, 177
warrant transmission, 177–178
well-being, 17
Williams, B., 180
Williamson, T., 216, 223, 252
Wright, C., 165, 177

Zagzebski, L., 246