



OXFORD

Epistemic Value

edited by

ADRIAN HADDOCK,
ALAN MILLAR, AND
DUNCAN PRITCHARD

EPISTEMIC VALUE

This page intentionally left blank

Epistemic Value

Edited by

ADRIAN HADDOCK, ALAN MILLAR,
AND DUNCAN PRITCHARD

OXFORD
UNIVERSITY PRESS

OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

© the several contributors 2009

The moral rights of the authors have been asserted
Database right Oxford University Press (maker)

First published 2009

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover
and you must impose the same condition on any acquirer

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloguing in Publication Data

Epistemic value / edited by Adrian Haddock, Alan Millar, and Duncan Pritchard.
p. cm.

Includes bibliographical references and index.

ISBN 978-0-19-923118-8

1. Knowledge, Theory of. I. Haddock, Adrian. II. Millar, Alan, Ph. D. III. Pritchard, Duncan.
BD161.E59 2009 121—dc22
2009013903

Typeset by Laserwords Private Limited, Chennai, India
Printed in Great Britain
on acid-free paper by
MPG Biddles Standards Ltd, King's Lynn, Norfolk

ISBN 978-0-19-923118-8

1 3 5 7 9 10 8 6 4 2

Contents

<i>List of Contributors</i>	vii
<i>Acknowledgements</i>	viii

Introduction	1
<i>Alan Millar, Adrian Haddock, and Duncan Pritchard</i>	

I. THE VALUE OF KNOWLEDGE

1. Reliabilism and the Value of Knowledge	19
<i>Alvin I. Goldman and Erik J. Olsson</i>	
2. Is There a Value Problem?	42
<i>Jason Baehr</i>	
3. Testimony and the Value of Knowledge	60
<i>Martin Kusch</i>	
4. The Value of Understanding	95
<i>Jonathan Kvanvig</i>	
5. Ugly Analyses and Value	112
<i>Michael R. DePaul</i>	
6. The Goods and the Motivation of Believing	139
<i>Ward E. Jones</i>	
7. Practical Reasoning and the Concept of Knowledge	163
<i>Matthew Weiner</i>	
8. Pragmatic Encroachment and Epistemic Value	183
<i>Pascal Engel</i>	
9. Luck, Knowledge, and Control	204
<i>Wayne D. Riggs</i>	

II. TRUTH AND EPISTEMIC APPRAISAL

10. The Values of Truth and the Truth of Values	225
<i>Michael P. Lynch</i>	
11. Epistemic Normativity	243
<i>Stephen R. Grimm</i>	

12. Curiosity and the Value of Truth	265
<i>Michael S. Brady</i>	
13. The Trivial Argument for Epistemic Value Pluralism, or, How I Learned to Stop Caring about Truth	284
<i>Berit Brogaard</i>	

APPENDIX. SYMPOSIUM ON JONATHAN KVANVIG'S
The Value of Knowledge and the Pursuit of Understanding

A. Precis of <i>The Value of Knowledge and the Pursuit of Understanding</i>	309
<i>Jonathan Kvanvig</i>	
B. The Value Problem	313
<i>John Greco</i>	
C. Is Understanding Factive?	322
<i>Catherine Z. Elgin</i>	
D. Understanding, Knowledge, and the <i>Meno</i> Requirement	331
<i>Wayne D. Riggs</i>	
E. Responses to Critics	339
<i>Jonathan Kvanvig</i>	
<i>Index</i>	353

List of Contributors

Jason Baehr, Loyola Marymount University
Michael S. Brady, University of Glasgow
Berit Brogaard, University of Missouri, St Louis
Michael R. DePaul, University of Notre Dame
Catherine Z. Elgin, Harvard University
Pascal Engel, University of Geneva
Alvin I. Goldman, Rutgers University
John Greco, St Louis University
Stephen R. Grimm, Fordham University
Ward E. Jones, Rhodes University
Martin Kusch, University of Cambridge
Jonathan Kvanvig, Baylor University
Michael P. Lynch, University of Connecticut
Eric J. Olsson, Lund University
Wayne D. Riggs, University of Oklahoma
Matthew Weiner, University of Vermont

Acknowledgements

We would like to thank the authors who contributed to this volume for their articles and for their cooperativeness during the production process. We gratefully acknowledge a Research Grant from the UK's Arts and Humanities Research Council that enabled us to organize, among other events, a workshop on epistemic value in October 2005 and conference on epistemic value in August 2006. Several of the contributions to this volume derive from presentations given at those events. Thanks are also due to the University of Stirling for its support, and to Yvonne McClymont and Nikki Leavitt for administrative and secretarial support. We are also grateful to Nikki for her assistance in producing the typescript, and to Steinvör Thöll Árnadóttir for further assistance with the typescript and for the preparation of the index. Finally, we thank Peter Momtchiloff at Oxford University Press for his encouragement and support throughout, and Tessa Eaton for efficiently supervising the production process.

Introduction

Alan Millar, Adrian Haddock, and Duncan Pritchard

1. VALUE PROBLEMS

In recent years there has been a surge of interest in issues about value in epistemology. Two themes loom large. One is provided by puzzles about the value of knowledge. The starting point, but only the starting point, is provided by the question, 'How can knowledge be more valuable than merely true belief?' The other theme concerns epistemic value in a broad sense. Here the central issue is how best to make sense of epistemic appraisal conceived broadly to include evaluation of beliefs not just with respect to whether they amount to knowledge, but also with respect to whether they are, for instance, justified, or reliably formed, that is, formed through methods or processes that reliably yield true beliefs. A common approach is to think of truth as the end for the sake of which we value a belief's being justified or reliably formed. The themes are related. For instance, we might hope to explain why knowledge is valuable in terms of the value we place on truth. Yet an interest in epistemic appraisal can lead in other directions. For instance, it can prompt enquiry into why we should value truth, how the goal of truth should be conceived, and why, if at all, we should think of truth as *the* goal of enquiry or as the fundamental goal of enquiry. All of these issues are represented in the essays collected here.

2. THE *MENO* PROBLEM

Suppose that you are a tourist in a foreign city. You ask for directions to the cathedral and you are told that this street, which you are facing, will take you there. You believe this and head along the street in question. What difference does it make that you should know that the street leads to the cathedral rather than have a merely true belief to this effect? If you act on your merely true belief

then, it seems, you will reach the cathedral as well as you would if you knew. This is a variant of Plato's road-to-Larissa example from *Meno*. Like Plato's it invites us to think that knowing something is no more valuable than having the corresponding true belief. It does not seem unduly strained to take the value in question to be practical. After all, the worry is posed by the thought that if you acted on the merely true belief you would reach your goal of arriving at the cathedral as well as you would if you knew. One way of reacting to the problem is, obviously, to concede that from a practical point of view knowledge is not any more valuable than having a merely true belief. Having made that concession, two options become available. One is to conclude that, contrary to what many think, knowledge is not any more valuable than merely true belief on the grounds that the only relevant dimension of evaluation is practical. The other is to reject the assumption that the only relevant dimension of evaluation is practical. This line is developed by, for instance, virtue epistemologists, who argue that knowing is valuable for its own sake, and thus irrespective of whether there are practical benefits to knowing. Not all theorists are willing to make the concession, however. The idea that from a practical point of view knowing is no better than having a merely true belief is open to challenge. Socrates' original proposal in *Meno* was to the effect that when we know our beliefs are tethered down. Timothy Williamson (2000: 78–80) defends a view that is very much in keeping with this proposal. He argues that knowledge is less vulnerable than merely true belief to rational undermining by future evidence. If this is right, then there are, at least potentially, practical benefits to knowing that are greater than those accruing to merely true belief. A merely true belief that *p* might be readily undermined. If it were, the person who had it would be left wondering whether or not *p* and thus be in no position for confident action on the assumption that *p*.

3. A RELATED VALUE PROBLEM AND 'SWAMPING'

Issues about the value of knowledge are often posed within a framework that assumes that knowledge is analysable in the traditional manner in terms of true belief plus other ingredients. A natural thought is that being justified is a necessary condition for knowing. Notoriously, on a common understanding of what it takes to be justified, having a justified belief that *p* that is true is not sufficient for knowledge that *p*. As Edmund Gettier (1963) famously showed, there are examples in which subjects have a justified true belief but in which it is accidental that the belief is true. These include examples in which a subject justifiably draws a conclusion that happens to be true from a premise that is justifiably believed yet false. On the assumption that knowledge admits of an analysis along traditional lines there will have to be some further condition—an anti-Gettier condition—to screen out Gettier cases.

Against this background Duncan Pritchard (2007: 86–7) distinguishes between the primary value problem, which is posed by the *Meno*, and the secondary value problem: why is knowledge more valuable than any proper subset of its parts? Knowledge might be better than merely true belief because if one knows one is justified in believing, and being justified in believing is better than having a merely true belief. Even if that is so it does not establish that knowing is better than having a justified true belief that falls short of knowledge, as in Gettier cases. The general idea behind this formulation of the secondary problem can be extended to frameworks, such as that of Williamson (2000), in which it is denied that knowledge is analysable in the traditional way. For within such a framework it still makes sense to ask, for instance, ‘Is knowing that p better than having a justified true belief that p that falls short of knowledge that p?’

Addressing the traditional analytical framework, Jonathan Kvanvig (2003) has argued that attempts to show that knowledge is better than any proper subset of its parts are doomed to failure due to the *swamping problem*. Suppose that we attempt to show that knowing that p is better than some lesser state that implicates a true belief that p, in terms of there being some property that one’s belief that p has when one knows that p but lacks when one is in the lesser state. Suppose, further, that the value of the property in question is parasitic on the value of true belief. Kvanvig thinks that no such property can solve the problem since if one has a true belief that p then, with respect to what is valuable, it is neither here nor there whether one’s belief has the property in question. As he puts it, the value of truth swamps any value that might be thought to attach to the proposed property (Kvanvig 2003: 45–8).

Virtue theorists attempt to show that knowledge is valuable for its own sake since there is value to having come by the truth through the exercise of particular virtues or competences (for instance, Sosa 1991; 2007; Riggs 2002; Zagzebski 2003; Greco 2003). As Ernest Sosa views the matter (2007: lecture 4), we rightly think that credit is due to performances, for instance, the shot of an archer, which achieve the aim of the relevant activity not merely by accident or luck, but due to the exercise of competence on the part of the performer. This general approach raises a host of interesting issues about the individuation of cognitive competences and the role of such competences in a general account of knowledge. It also faces a number of challenges. For instance, one might think that an adequate answer to the *Meno* problem should explain why the tourist seeking the way to the cathedral should prefer to know rather than have a merely true, or a merely true and justified, belief on the matter. From this perspective it is not clear that the consideration that knowing would be worthy of credit gives the right kind of answer. We were led by the problem to wonder what good it would be to an enquirer to know, not how good the enquirer would be to know.

4. TRUTH AND EPISTEMIC APPRAISAL

A common idea is that it is for the sake of truth that it matters that we should acquire knowledge or matters that our beliefs should be justified. Philosophical issues quickly arise when we try to probe more deeply. Here are a few:

- (a) It does seem important that we should avoid believing what is false. But we might strive towards that by trying to minimize what we believe, so far as it is in our power to do so. Is it not also important that we should believe what is true? Yet it seems implausible that we should believe *whatever* is true if for no other reason than that, lacking omnipotence, it is impossible for us to do so. The question, then, is how best to formulate a more restrictive requirement and whether a restrictive requirement enables us to make sense of epistemic appraisal.
- (b) Should we think of the truth as being instrumentally valuable or as valuable for its own sake? Either way, is the value of a true belief independent of its content? For some at least it seems powerfully intuitive that the value of a true belief, or of an item of knowledge, is dependent on its content. But now if that is so what determines that true belief or knowledge with this or that content is of value, when true belief or knowledge with different content is not? Is this simply determined by the goals of enquiry that we actually set ourselves or in some other way?
- (c) Is it right to conceive of our cognitive achievements as being directed ultimately at true belief? Might there be other epistemic goals? Kvanvig (2003 and this volume Ch. 4) argues that understanding has a special kind of value that knowledge lacks. A number of the contributions in this volume address this matter.

5. THE CONTRIBUTIONS IN PART I

The essays in Part I are all more or less directly concerned with the value of *knowledge*. Problems concerning the value of knowledge that stem from Plato's *Meno* are often made vivid in the context of criticism of reliabilist theories of knowledge or justification. The swamping problem applied to reliabilism depends on the assumption that if a belief is true its value would be not raised by the fact that it was reliably produced. In their contribution (Ch. 1), Alvin Goldman and Erik Olsson offer two independent responses to the problem.

The first response is to reject the presupposition just noted and develop *the conditional probability solution*. This invokes the idea that the composite state consisting in having a true belief produced by a reliable process has the property

of *making it likely that one's future beliefs of a similar kind will also be true*. Goldman and Olsson argue that a true belief's being formed by a reliable process enhances the probability that the next belief formed by the same process will be true to a greater degree than, all else equal, a belief of the same type's being produced by an unreliable process would enhance the probability that the next belief of the same type formed by that unreliable process would be true. Goldman and Olsson emphasize that the value of being in the state of having a belief produced by a reliable process is not, as they put it, contained in the value of the implicated true belief. They note that the conditional probability solution allows for the possibility that reliabilist knowledge is not always more valuable than true belief but argue that this is no objection to the proposed solution.

Goldman and Olsson's second solution offers an explanation of why we come to accord reliable belief-forming processes an *autonomous value* that is not dependent on the value of the resultant true beliefs. Furthermore, they think that the 'the value imputed to a token process is inherited from the value imputed to its type' and 'isn't a function solely of that token's own consequences' (p. 31). Applying this to the present case they suggest that '[w]hile many [reliable belief-forming] processes are originally regarded as merely instrumentally valuable to true-belief attainment, they are later upgraded to the status of independent value, thereby accommodating the legitimacy of *adding* their value to that of true-belief outcomes' (pp. 34–5). From this perspective the added value accruing to knowledge is such that knowing will always be more valuable than the implicated true belief.

In his contribution (Ch. 2) Jason Baehr asks whether there is a problem about the value of knowledge. There seems to be a problem in the light of 'the guiding intuition' that knowledge is more valuable than true belief. This intuition is thought to supply a constraint on the analysis of knowledge: the analysis should identify among the ingredients of knowledge something that has truth-independent value. Baehr suggests that if it is to perform its role the content of this guiding intuition must be 'entirely *general*, in the sense that it must . . . cover *all* instances of knowledge' (p. 45). This is because an intuition about value that concerned only some cases of knowledge would not supply a constraint on an analysis of knowledge aimed at identifying its essential features. Baehr then observes that the content of the guiding intuition must be entirely formal in that 'it must not provide any indication of *why* or *that in virtue of which* knowledge is more valuable than true belief' (p. 46). This derives from the fact that the constraint imposed by the intuition is supposed to be entirely formal in that 'it does not require anything of an analysis of knowledge *beyond* that one or more of the conditions specified by this analysis have truth-independent value' (p. 44). Baehr argues that it is implausible that there is a widespread guiding intuition that has the required features. For instance, he argues that it is implausible that when we think a case of knowledge more valuable than a corresponding true belief this has nothing to do with specific features of the knowledge. He finds recent ideas of John Greco on the value of knowledge to be in keeping with the

conclusions reached so far since Greco's conception of the value of knowledge turns on a substantive account of what knowledge is. Reviewing the situation, Baehr raises a further question suggested by the guiding intuition: what besides truth might be included within the range of epistemic values?

Martin Kusch (Ch. 3) responds to Kvanvig's scepticism about the value of knowledge by outlining 'a *communitarian* form of value-driven epistemology' which 'seeks to understand the values of various cognitive states in relation to the needs and actions of human beings in social interaction with one another' (p. 60). He adopts a genealogical method from Edward Craig (1990), which is designed to bring out why knowledge might matter to us and thereby shed light on what knowledge is. The starting point is the idea that our ancestors, wondering whether *p*, would have had an interest in what Kusch calls *protoknowers*—roughly speaking, those who are detectably likely to be right as to whether *p*. From this perspective the question arises as to why having protoknowledge should be more valuable than having merely true belief. It is at this point that Kusch introduces into the genealogical story a modification inspired by Bernard Williams (2002) to the effect that the institution of testimony is a collective good that exists only if underpinned by a network of intrinsic values. Kusch applies this to, among other things, the swamping problem. If the institution of testimony is a collective good *supported by intrinsic values*, like accuracy and sincerity, then it is not of merely instrumental value; it is of central intrinsic value to our social existence. These values are maintained through interactions, like talking, sanctioning, honouring, and shaming, through which our social existence is shaped. Indeed, attributions of protoknowledge themselves play a key role in sustaining the values and the institution of testimony that depends upon them. Seen in this light, having protoknowledge—at its core true belief deriving from a good informant—is not just a good thing because it involves acquiring a true belief by a good method. It is something that depends essentially on an institution that is a collective good and that contributes to shaping our existence as human beings.

In his piece Jonathan Kvanvig develops his conception of the special epistemic value of objectual understanding—the kind of understanding in play when we understand quantum mechanics or Republican ideology. Understanding on this conception is a matter of grasping explanatory, probabilistic, and logical connections between pieces of information or ideas. Whereas the luckily true beliefs highlighted by Gettier cases cannot constitute knowledge, understanding is compatible with the implicated beliefs being luckily true. What matters for understanding, according to Kvanvig, is that one has the abilities that are constitutive of understanding such as, for instance, the ability to answer questions correctly. The differences between knowledge and understanding have a bearing on the value problem. Attempts to respond to the Gettier problem, Kvanvig thinks, are bound to end up with accounts of knowledge that are 'ad hoc and gerrymandered' (p. 103), making it difficult to see why knowledge should be of value. But because understanding tolerates luck, it is not subject to the same problem.

Could there not be cases in which understanding is undermined because of a failure to know why something is so? Kvanvig addresses an example presented by Michael DePaul and Stephen Grimm (2007) that is designed to show just this. The example is one in which an unreliable source invents an explanation that by chance is a correct account of why the winning goal at a game of soccer was scored by the US team. The account has it that the US team scored the winning goal because the opposition's goalkeeper slipped in the mud. DePaul and Grimm claim that the ingredients for understanding on Kvanvig's conception are in place but the account does not give us understanding of why the winning goal was scored. In defence Kvanvig notes that the use of the example by DePaul and Grimm turns on the idea that we do not understand because we do not *know* why the winning goal was scored. He concedes that there is a sense in which one does not understand why the winning goal was scored, if that entails knowing the correct explanation. But that, he suggests, leaves open the possibility that understanding in some sense—objectual understanding—is present. En route Kvanvig considers methodological issues concerning the role of linguistic intuitions in philosophical enquiry.

In his contribution to the volume (Ch. 5), DePaul assesses a line of reasoning that he discerns in Kvanvig (2003), and which, as we have just observed, also occurs in Kvanvig's contribution to the present collection. This line of reasoning takes as its starting point some 'ugly' analysis of knowledge—one that seems complicated, ad hoc, messy, gerrymandered—and argues to the conclusion that knowledge is not valuable. DePaul discusses a similar argument that focuses not so much on ugliness, but on our inability to see why satisfying the conditions given by a proposed analysis of knowledge would be valuable. He seeks to put pressure on the transition to the conclusion that knowledge so conceived is not valuable. Why, he asks, should we suppose that if *A* is an adequate analysis of knowledge, and knowledge is valuable, then it should be apparent to us after due consideration that satisfying the conditions given by *A* is valuable? In support of this he asks why we should expect an adequate analysis of something to operate at the same level as our ordinary understanding of that thing. Why, in particular, must an adequate analysis of knowledge be spelled out in terms of concepts we ordinarily employ? If an adequate analysis of knowledge need not meet this requirement then, he argues, it is far from clear why we should expect to recognize on due consideration that satisfying the analysis is valuable. Further, DePaul suggests, analyses may have a metaphysical focus: we should not assume that by revealing the nature of knowledge we thereby bring into focus what makes knowledge valuable.

Ward Jones has virtue-theoretic conceptions of the value of knowledge in his sights, though his starting point is a much more general discussion of doxastic goods—'goods [that] correspond to the ways in which believing is a valuable state, or contributes to valuable states or things' (p. 139). These would include the feeling of well-being or happiness that a belief can give us and the explanatory

insight afforded by a belief. Jones observes that some doxastic goods can motivate belief only surreptitiously. It might bring relief to me to believe that my sick mother is recovering, but, if I recognize that the sole motive for believing is that it brings relief, my belief will tend to be weakened. Jones explores why it matters for issues concerning the values of exploration that some doxastic goods are surreptitious. He focuses on considerations about the relation between reasons and motivation that are more usually discussed in the literature on practical reasoning. Sympathetic to the idea that necessarily reasons are potential motivators, Jones notes that this idea can be interpreted strictly, in keeping with the view of Bernard Williams (1980) about reasons for action or in more liberal ways as in the writings of John McDowell (1995). He favours the stricter view and on that account resists the idea that surreptitious doxastic goods can provide reason to believe. Such goods are ‘at best, beneficial side-effects of believing’ (p. 149). Finally, he teases out the ramifications of his view for virtue-theoretic accounts of the value of knowledge focusing our attention on whether gaining credit for a true belief virtuously formed is a surreptitious good.

The starting point of Matt Weiner’s essay (Ch. 7) is the idea discussed by John Hawthorne (2004) that knowledge is tied to practical reasoning. As Weiner reads it the idea is that ‘it is unacceptable to use p as a premise in your practical reasoning if and only if you do not know that p ’ (p. 164). Offered a lottery ticket with a \$5000 prize at the price of one penny it would be crazy to argue that since if you buy the ticket it will lose, and you will therefore be out by a penny, you should not buy the ticket. This would be so even if the premises were true. Weiner is prompted to consider what it is for a premise in practical reasoning to be acceptable. He argues that there is no single way of understanding it so that all and only known premises are acceptable. Practical reasoning can be acceptable, from a certain point of view, even when it has a false premise, and thus proceeds from something that is not known. For instance, an inference with the premise that one will not have enough money to go on an African safari next year can be acceptable even though the premise is not known. It is not feckless to rely on the assumption that you will not be able to afford the safari. If that does not lead to the best outcome—you don’t buy the guidebook to Africa and do buy the local guide—then that is simply epistemic bad luck. But there can be different interests attached to a stretch of practical reasoning. In some contexts it matters whether the premises are true. In others what matters is that belief in the premises should be well enough justified.

In the light of these and related considerations, Weiner invites us to think of knowledge as being like a Swiss Army knife. Each of the tools in the knife are valuable in some practical contexts and not in others. There is nothing especially valuable about being a Swiss Army knife beyond its making each of the tools conveniently available. Weiner suggests that knowledge is in some respects like this. Knowing that p involves having a belief that p that satisfies various conditions. Each condition may be one that it is valuable to satisfy in some

contexts of practical reasoning but not in others. Knowledge is worth having because one does not know in advance which of its elements is going to be valuable in this or that circumstance.

The idea that epistemic evaluation is in some way dependent upon practical concerns is an important theme of recent epistemology. Pascal Engel (in Ch. 8) is concerned with this phenomenon of pragmatic encroachment. Engel considers, and finds wanting, a number of different arguments for a number of different forms of pragmatic encroachment. According to Paul Horwich, the value of true beliefs is to be explained in terms of their role in guiding people to satisfy their desires. Engel questions whether this account can properly accommodate the normative character of truth. The fact that it is pragmatically good to have a true belief, Engel contends, would not explain why it is desirable to believe something because it is correct and correct because it is true.

Engel then addresses Jeremy Fantl and Matthew McGrath's (2002) argument for a form of pragmatic encroachment upon justification, according to which it is possible that one person is justified in believing an empirical proposition, and another person is not, even though both people are in possession of the same evidence for the proposition. According to Fantl and McGrath, the difference in justification is explained by the fact that the practical costs of believing the proposition are higher for the second person than they are for the first. But Engel thinks that all this argument shows is that it is more important for the second person to be justified in believing the proposition than it is for the first.

Finally, Engel discusses Jason Stanley's (2005) argument for a form of pragmatic encroachment upon knowledge, according to which its being rational to act as if a certain proposition is true is a necessary condition for knowing the proposition in question. Engel thinks that Stanley's argument only gives us reason to endorse the weaker claim that knowing a proposition is a necessary condition for its being rational to act as if the proposition is true. He does not deny that practical factors play a role in epistemic enquiry. But he insists that this role is limited to what he calls 'pragmatic relevance': knowledge does indeed matter for practical reasoning and action but that is a far weaker claim than the claim that knowledge depends on pragmatic considerations.

Wayne Riggs commends a value-driven approach to epistemology on the grounds that becoming clear about the value of knowledge and why in practice we care about knowledge 'helps us understand how our cognitive lives are intertwined with our moral lives, our prudential lives, etc.' (p. 204). Against this background he suggests that if we can work out what is so bad about luckily true belief we can come to an understanding of why we value knowledge. His contribution takes a step in this direction by exploring in some detail how luck should be conceived if knowledge is taken to exclude luck. Pritchard's (2005) account of luck makes safety central. More specifically, on his account if an event is lucky then it does not occur in a wide class of the nearest possible worlds where the relevant initial conditions are the same, and, in addition, it is

an event that is significant to the agent, or would be if the agent were apprised of the relevant facts. Drawing on a discussion of Pritchard by Jennifer Lackey (2006), Riggs argues that neither of the specified conditions on luck are necessary: there are counterexamples in which there is a lucky combination of events or circumstances yet the salient features of the example obtain in many nearby possible worlds. Riggs also argues that Pritchard's conditions are not jointly sufficient, while recognizing that Pritchard does not intend his conditions to be taken to be jointly sufficient. Building on these considerations he finds fault with safety-based accounts of knowledge. He then defends an account of luck in terms of lack of agent control. He suggests that '[o]ne has control over some happening to the extent that the happening is properly considered something the agent has *done*' (p. 214) and thus both something that the agent meant to do and the product of the agent's power, abilities, and skills. He then argues that this conception handles the examples he directed against Pritchard's account. Finally, he addresses an objection to the lack-of-control account to the effect that there are far too many things that are not in one's control but which are also not lucky. This leads him to suggest that for an event to be lucky for a person the person must not have successfully exploited the event (prior to its occurrence) for some purpose.

6. THE CONTRIBUTIONS IN PART II

Michael Lynch (Ch. 10) presents two ways of understanding the claim that truth is a value. The first has it that truth is a proper end of enquiry in the sense that '[i]t is *prima facie* good that, relative to the propositions one might consider, one believe all and only those that are true' (p. 226). The second takes truth to be a norm of belief in the sense that '[i]t is correct to believe $\langle p \rangle$ if and only if $\langle p \rangle$ is true' (p. 228). Lynch asks what sort of meta-normative stance should, or indeed, can we take to the value of truth conceived in either of these ways. His particular critical target is expressivism. As applied to *alethic correctness* (the value implicated by the truth-norm), expressivism states that '[i]t is a non-factual matter whether beliefs are correct if and only if true, since to describe a belief as correct is not to state a fact about it but to express a sentiment or attitude' (p. 231). After arguing that a simple version of expressivism is self-undermining, Lynch considers a more sophisticated version that distinguishes between engaged and disengaged stances. Sophisticated expressivists can acknowledge that from an engaged standpoint we can use evaluative language to affirm all that the realist about values can affirm. To this end they can adopt a deflationary view of truth and from an engaged standpoint acknowledge that the truth-norm is true. At the same time they can acknowledge that from a disengaged stance we can tell an expressivist story about the function of evaluative thought according to which the truth-norm is neither true nor false but an expression of our desires and sentiments.

Lynch argues that the disengaged stance is not an option when it comes to alethic correctness. The reason is that the expressivist must invoke the natural fact that a given proposition is true to explain why it is correct to believe it but, being a deflationist, lacks a theory of truth that provides for truth to have an explanatory role. After considering possible responses on the part of the expressivist, Lynch argues that the expressivist about the value or good of having true beliefs fares no better. He concludes that a sceptical attitude towards the value of truth is not an option on either of the understandings of this value that have been considered

How can we best make sense of epistemic appraisal in the broad sense? Stephen Grimm (Ch. 11) engages with a teleological approach to this question on which truth is regarded as being of intrinsic epistemic value and the positive epistemic status that some beliefs have is explained in terms of the value of truth. A problem for this view is that it is far from obvious that intrinsic value attaches to true beliefs irrespective of their content. Suppose that to deal with this it is suggested that only some true beliefs are intrinsically valuable. Grimm points out that this raises the question, 'How should we make sense of our epistemic appraisals with respect to those beliefs (or, better, those topics) that apparently *lack* intrinsic epistemic value . . . ?' (p. 249). He considers Ernest Sosa's (2007) proposal that we should think of values in relation to domains of human experience. Performances within a domain are evaluated in terms of the values that are fundamental to that domain. Appraising performances within a domain involves no commitment to the domain-transcendent value of its fundamental values. Epistemic appraisal, on this approach to the matter, is made within a domain for which the fundamental value is the truth. But this involves no commitment to the view that the truth has an intrinsic value that is absolute (domain-transcendent). Grimm argues that this view fails to accommodate the fact that to judge a belief to be unjustified or irrational is to judge that the subject's attitude towards the content of the belief *should* be reconsidered, in some binding sense, which implies that the subject is culpable in some way. The sense of 'should' that is internal to a domain, he thinks, is too weak to capture this binding sense and thus is unable to capture the idea that the truth-perspective is non-optional. Grimm suggests that to make sense of epistemic appraisal we need to adopt a communal or social view of truth that can accommodate a version of the thesis that every truth has a special value. The key idea is that every topic might come to have value given the varied practical concerns people throughout a community might have. We must duly respect the truth because of the central role that truths play in the lives of others if not ourselves.

In the course of his discussion Grimm notes that some theorists think that we value certain truths for their own sake, irrespective of what practical ends they may serve, simply because of natural curiosity about certain matters. In his essay (Ch. 12) Michael Brady poses the question, 'How can we move from the claim that we are naturally curious to discover the answers to particular questions, to the claim that answers to those questions are valuable in themselves?' (p. 271).

There are reasons to think that this move is problematic. For instance, if our natural curiosity is compulsive it might generate enquiries leading to answers on which we place no value. From this and other considerations Brady concludes that 'it is implausible to suppose that we could have a general interest in the truth which is generated by our natural curiosity' (p. 274).

Curiosity, Brady suggests, is an emotion that by its nature involves selective attention. In particular, it alerts us to matters of intrinsic interest or fascination. Curiosity can be aroused by what is of little or no interest in this sense, but the standard of correctness or appropriateness for curiosity about some matter is that it should be of intrinsic interest or fascination. In the light of this one might think that true beliefs are valuable for their own sake because such beliefs constitute answers to intrinsically interesting or fascinating questions. But Brady argues that it is neither true that true beliefs are valuable only if they answer interesting or fascinating questions, nor that if true beliefs answer interesting or fascinating questions they are valuable. He concludes that the value of true beliefs is only sometimes contingently related to our natural curiosity. Nonetheless he thinks that 'if we are to equate the aim of inquiry with something that is genuinely valuable, then it should be identified with the goal of attaining interesting or fascinating or surprising true beliefs' (p. 281) and natural curiosity has a role in attaining this goal.

Epistemic monism, on Berit Brogaard's understanding (Ch. 13), is the view that truth is the highest epistemic goal. Brogaard argues that a problem arises if this view is understood to be tantamount to the thesis (EG) that for any proposition p , the epistemic goal is for you to believe that p if and only if it is simply true that p . A proposition p is simply true if and only if it is true at the actual world. Brogaard argues that the problem arises because on plausible assumptions (EG) entails a further thesis (FT) that for any proposition p , the truth-value of p is relative only to the world of evaluation. (FT) is linked to the idea that the objects of belief have their truth-values eternally and thus that the only factors that determine whether or not a proposition is true are factors pertaining to the world with respect to which the proposition is being evaluated. She explores examples suggesting that (FT) is false. Suppose, for instance, that when I say 'John is a firefighter' at t_1 , what I am saying is that at t_1 John is a firefighter, and when you say 'John is not a firefighter' at t_2 what you are saying is that at t_2 John is not a firefighter. It follows that if t_1 and t_2 are different times we are not contradicting each other. But, argues Brogaard, it is unnatural to suppose that if my utterance was at t_1 and yours was at t_2 ($t_1 \neq t_2$) we would not be contradicting each other. So, on the face of it, the object of the beliefs we would express by our utterances is not adequately specified by, respectively, the propositions that at t_1 John is a firefighter and that at t_2 John is not a firefighter.

Relativists respond to these and related considerations by arguing that there are relative expressions whose extensions vary with parameters determined not only by the context of use but also by a context of assessment, where '[a] context of

assessment is a context in which a sentence is considered or evaluated for truth' (p. 291). If this is right 'there are some propositions which have no truth-value relative only to a world and a time' (p. 292) and thus some that are not simply true. If this is right, argues Brogaard, then epistemic value monism, understood as above, is false. After considering various responses to this line of thought Brogaard suggests an alternative view, perspectivalism, which, she argues, is both compatible with a version of epistemic monism and can explain the linguistic data. On this alternative there is no need to distinguish between the context of use and the context of assessment; the data are accommodated by recognizing a judge parameter, the default value of which is the speaker.

7. THE APPENDIX

The symposium on Jonathan Kvanvig's *The Value of Knowledge and the Pursuit of Understanding* opens with a brief precis by the author. Two key themes of the book are that 'our ordinary assumption that knowledge is more valuable than its subparts is simply mistaken' (p. 311) and 'that understanding has a special and unique value that exceeds the value of its subparts' (p. 311).

In his critical essay John Greco argues that Kvanvig's presentation of the value problem shifts between different questions. He challenges Kvanvig's assumption that there is widespread pre-theoretical conviction that knowledge is more valuable than any of its proper subparts. If there is no such conviction then, he suggests, Kvanvig's demand that a solution to the value problem should show that knowledge is more valuable than any of its proper subparts is inappropriately demanding. Nonetheless, argues Greco, all of the questions Kvanvig poses can be answered in terms of a virtue-theoretic approach to value problems according to which knowledge is true belief achieved through the exercise of intellectual abilities that constitute intellectual virtue. Indeed, such an approach, he claims, can even explain why knowledge is more valuable than the sum of values of its parts. This is because success *through* virtue is more valuable than success *plus* virtue.

Catherine Elgin explores the idea, prominent in Kvanvig's book, that propositional understanding is factive. She argues that if epistemology is to accommodate science we need to consider different types of understanding. In particular, she argues that 'some bodies of information, even though they are not true, nonetheless display a measure of understanding' (p. 325). To make sense of the idea that progressive science involves advances in understanding, she suggests, we need to recognize that not all understanding is factive. She supports her case with reference to the role of idealization in science. Such idealizations, she contends, have a central and ineliminable role in scientific understanding. While they are fictions—felicitous falsehoods—they 'afford epistemic access to matters of fact that are otherwise difficult or impossible to discern' (p. 327).

Wayne Riggs considers that an adequate response to the problem posed by Plato's *Meno* does not require us to show that knowledge is more valuable than any combination of its proper subparts. Indeed, in setting up the problem he sees no reason to deny that some true belief could be as valuable or more valuable than any instance of knowledge. He targets Kvanvig's view that understanding is factive and that we can have understanding of X in the absence of knowledge of X. Further, he argues that '[a]s long as there is a possible source of epistemic value for a belief that is independent of whether that belief constitutes knowledge [e.g. its being a component of a subject's understanding of some subject matter] then there will be the possibility that the belief could derive enough value from that alternative source that it is equally or more valuable than it would be if it were known' (p. 337). Indeed, being an element in one's understanding of some subject matter is just such an alternative source. Riggs concludes that a better way to formulate the *Meno* requirement is that 'an item of knowledge must have a kind of epistemic value, or perhaps a source of epistemic value, that no belief that fails to count as knowledge has' (p. 338).

In his response to his critics Kvanvig begins by defending the view that all truths, irrespective of their content, are defeasibly valuable. His main focus, however, is on the nature and value of understanding and on the character of the *Meno* problem. As in his contribution to this collection in Part I (Ch. 4), Kvanvig gives special attention to objectual understanding, 'the kind of understanding in which the content of the attitude is an object of some sort (person, theory, part of reality, etc.)' (p. 341). He suggests that such understanding is quasi-factive, being dependent *in some measure* on true belief. He acknowledges that there is vagueness in the extent to which understanding can incorporate false belief and he exploits this vagueness to accommodate examples that critics have taken to show that understanding is not necessarily factive. Understanding X, he thinks, may incorporate some false beliefs about X.

In response to Elgin's discussion of idealization, he distinguishes between understanding a model or theory and understanding the phenomenon that the theory or model is about. Idealizations are indeed strictly false but they may themselves be part of the object understood, as when the object is *how a theory or model stands in relation to reality*. In response to concerns about his formulations of the *Meno* problem, Kvanvig emphasizes that the key issue for him always concerned the value of knowledge over any of its subparts. Responding to Greco's suggestion that there is no pre-theoretical conviction that knowledge is more valuable than any of its subparts, he suggests, among other things, that this is belied by our taking knowledge to close enquiry in a way that justified true belief does not. He agrees with Riggs's view that his way of formulating the *Meno* problem does not sit well with his view about the value of understanding, but contends that the formulation should be viewed as a working hypothesis about the nature of the problem. Finally, he takes issue with Greco's attempt to solve the *Meno* problem by arguing that Greco's account of knowledge is not adequate.

REFERENCES

- Craig, E. (1990). *Knowledge and the State of Nature*. Oxford: Clarendon Press.
- DePaul, M., and Grimm, S. (2007). 'Critical Study of *The Value of Knowledge and the Pursuit of Understanding* by Jonathan Kvanvig', *Philosophy and Phenomenological Research*, 74, 498–514.
- Gettier, E. (1963). 'Is Justified True Belief Knowledge?', *Analysis*, 23, 121–3.
- Greco, J. (2003). 'Knowledge as Credit for True Belief', in M. DePaul and L. Zagzebski (eds.), *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Oxford University Press.
- Hawthorne, J. (2004). *Knowledge and Lotteries*. Oxford: Clarendon Press.
- Kvanvig, J. (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge: Cambridge University Press.
- Lackey, J. (2006). 'Pritchard's *Epistemic Luck*', *Philosophical Quarterly*, 56, 284–9.
- McDowell, J. (1995). 'Might There Be External Reasons?', in J. E. J. Altham and R. Harrison (eds.), *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams*. Cambridge: Cambridge University Press. Reprinted in McDowell's *Meaning, Knowledge, and Reality*. Cambridge, MA: Harvard University Press.
- Riggs, W. (2002). 'Beyond Truth and Falschhood: The Real Value of Knowing that P', *Philosophical Studies*, 107, 87–108.
- Pritchard, D. (2005). *Epistemic Luck*. Oxford: Clarendon Press.
- (2007). 'Recent Work on Epistemic Value', *American Philosophical Quarterly*, 44, 85–110.
- Sosa, E. (1991). *Knowledge in Perspective*. Oxford: Blackwell.
- (2007). *A Virtue Epistemology: Apt Belief and Reflective Knowledge*, i. Oxford: Clarendon Press.
- Williams, B. (1980). 'Internal and External Reasons', in R. Harrison (ed.), *Rational Action*. Cambridge: Cambridge University Press. Reprinted in Williams's *Moral Luck*. Cambridge: Cambridge University Press, 1981.
- (2002) *Truth and Truthfulness: An Essay in Genealogy*. Princeton: Princeton University Press.
- Williamson, T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.
- Zagzebski, L. (2003) 'The Search for the Source of Epistemic Good', *Metaphilosophy*, 34, 12–28. Reprinted in M. Brady and D. Pritchard (eds.), *Moral and Epistemic Virtues*. Oxford: Blackwell, 2003.

This page intentionally left blank

PART I

THE VALUE OF KNOWLEDGE

This page intentionally left blank

Reliabilism and the Value of Knowledge

Alvin I. Goldman and Erik J. Olsson

1. WEAK KNOWLEDGE IS MERE TRUE BELIEF

It is a widely accepted doctrine in epistemology that knowledge has greater value than mere true belief. But although epistemologists regularly pay homage to this doctrine, evidence for it is shaky. Is it based on evidence that ordinary people on the street make evaluative comparisons of knowledge and true belief, and consistently rate the former ahead of the latter? Do they reveal such a preference by some sort of persistent choice behavior? Neither of these scenarios is observed. Rather, epistemologists come to this conclusion because they have some sort of conception or theory of what knowledge is, and they find reasons why people *should* rate knowledge, so understood, ahead of mere true belief. But what if these epistemological theories are wrong? Then the assumption that knowledge is more valuable than true belief might be in trouble. We don't wish to take a firm position against the thesis that knowledge is more valuable than true belief. But we begin this paper by arguing that there is *one* sense of 'know' under which the thesis cannot be right. In particular, there seems to be a sense of 'know' in which it means, simply, 'believe truly.' If this is correct, then knowledge—in this weak sense of the term—cannot be more valuable than true belief. What evidence is there for a weak sense of 'knowledge' in which it is equivalent to 'true belief'?

Knowledge seems to contrast with ignorance. Not only do knowledge and ignorance contrast with one another but they seem to exhaust the alternatives, at least for a specified person and fact. Given a true proposition *p*, Diane either knows *p* or is ignorant of it. The same point can be expressed using rough synonyms of 'know.' Diane is either *aware of* (the fact that) *p* or is ignorant of it. She is either *cognizant of* *p* or ignorant of it. She either *possesses the information* that *p* or she is uninformed (ignorant) of it.

To illustrate these suggestions, consider a case discussed by John Hawthorne (2002). If we ask you how many people in the room know that Vienna is the capital of Austria, you will tally up the number of people in the room who possess the information that Vienna is the capital of Austria. Everyone in the room who

possesses the information counts as knowing the fact; everybody else in the room is ignorant of it. It doesn't really matter, in this context, where someone apprised of the information got it. Even if they received the information from somebody they knew wasn't trustworthy, they would still be counted as cognizant of the fact, that is, as knowing it rather than as being unaware of it.

The point can be expressed by the following principle:

$$(\text{COMPL}) \quad \sim (K_{Sp}) = \text{IGN}_{Sp}$$

(COMPL) applies only where p is true, or factive. Given the truth of p , it says that ignorance and knowledge are complements of one another, that is, S is ignorant of p if and only if S doesn't know that p . How could this principle hold, however, if knowledge consisted in something more than true belief? Suppose, for example, that knowledge is justified true belief plus an anti-Gettier condition X . Then, assuming the truth of p , S 's failure to know p wouldn't imply his being ignorant of p . Instead of being ignorant of p , he might believe p unjustifiedly, or might believe it justifiedly but without fulfilling condition X . So, when p is true, failure to know p in a strong sense of knowing (e.g. JTB + X) would not imply ignorance. The correctness of (COMPL) implies that, at least in one sense, knowing is nothing more than having true belief.

We can illustrate the foregoing argument diagrammatically (Fig. 1.1). If knowledge is something like JTB + X , then the terrain is exhaustively captured by the set of possibilities displayed in the diagram. The complement of knowing is not knowing, but not knowing p (where p is true) can occur in any of three different ways: (1) by being ignorant of p (not believing it), (2) by believing p unjustifiedly, or (3) by believing p justifiedly but violating condition X . Under this concept of knowledge, no inference is licensed from not knowing p to being ignorant of p . We contend, however, that there is a sense of 'knowing' in which this inference is licensed. People commonly make this inference. The

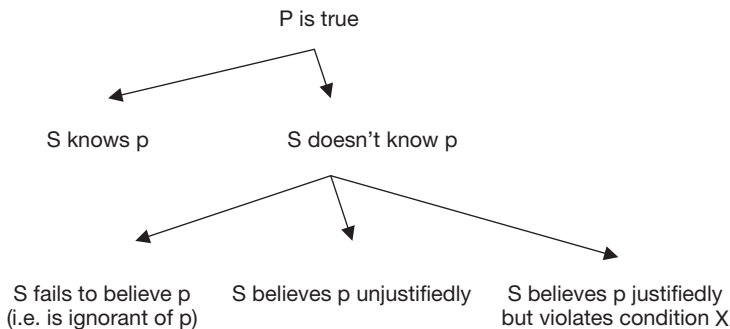


Figure 1.1.

only concept of knowledge compatible with this inference is the one in which knowledge = true belief.

Someone might challenge this conclusion by challenging the claim that ignorance of *p* (where *p* is true) is failure to believe *p*. The challenger might claim that there are three ways to be ignorant of *p*: by failing to believe it, by believing it unjustifiedly, or by violating condition X. If this were right, (COMPL) wouldn't imply that knowledge is mere true belief. But this claim about the meaning of 'ignorance' is plainly wrong. It is highly inaccurate, inappropriate, and/or misleading to characterize somebody who unjustifiedly believes (the fact that) *p* as being *ignorant* of *p*. Similarly, it is highly inaccurate, inappropriate, and/or misleading to characterize somebody who justifiedly believes *p* but fails to satisfy condition X as being ignorant of *p*. Thus, the exhaustiveness of the dichotomy between knowledge and ignorance is best explained by the thesis that knowing *p* (in one sense of 'knowing') is simply believing *p* where *p* is true. It does not consist in anything beyond true belief.¹

If there is a weak sense of 'knowledge' in which it is equivalent to true belief, then the unqualified thesis that knowledge is more valuable than true belief goes by the board. If a state of knowing, in this sense of 'know,' is nothing more than a state of true belief, then neither knowing nor truly believing can be more valuable than the 'other.' However, we do not maintain that weak knowledge is the *only* kind of knowledge, or the only sense of 'know.' In this respect we depart from the radical positions of Isaac Levi (1980) and Crispin Sartwell (1991, 1992), who both hold that 'know' *uniquely* means 'believe truly.' We cheerfully grant that there is a stronger sense of 'know,' which epistemologists have long pursued and which involves more than the two conditions of belief and truth. For this stronger sense of 'know,' the thesis that knowledge is more valuable than true belief is not so easily disputed. In the rest of the paper we shall be concerned with knowledge in its strong sense.

2. THE VALUE OF RELIABILIST KNOWLEDGE

Why is knowledge, in the strong sense, more valuable than mere true belief? The question was first raised in Plato's dialogue *Meno* where it was pointed out that a mere true belief seems instrumentally just as valuable as knowledge. What

¹ That there are both weak and strong senses of 'know' is advocated in Goldman (1999). According to Stephen Hetherington (2001), knowledge is a concept that admits of degrees and mere true belief is a minimal kind of knowledge. Similarly, Keith DeRose (2002) says that he is 'tempted' by the contextualist thesis that, in very low-standard contexts, nothing more than true belief is required for knowledge. The identification of knowledge with true belief has been defended in the German philosophical literature by Wolfgang Lenzen (1980), Franz von Kutschera (1982), and, more recently, Ansgar Beckermann (2001). For a critical discussion of Goldman's defense of the weak sense of knowledge, see Le Morvan (2005).

matters for someone who wants to get to Larissa is to have a *true belief* about its location. Satisfying the stronger requirement of *knowing* where Larissa is does not seem to make you any more likely to get there. Still we do believe that knowledge is somehow better.

The extra-value-of-knowledge (EVOK) problem can be used to test the adequacy of accounts of knowledge. Suppose an analysis of knowledge is incompatible with knowledge having an added value. That would be a pretty strong argument against the adequacy of that analysis. Recently a number of authors have argued that process reliabilism does not pass the value test.² According to process reliabilism, a subject S knows that p if and only if (1) p is true, (2) S believes p to be true, (3) S's belief that p was produced through a reliable process, and (4) a suitable anti-Gettier clause is satisfied. Ward Jones (1997) raises the value objection as follows:

In short, given the reliabilist's framework, there is no reason why we should care what the method was which brought about a true belief, as long as it is true. We value the better method, because we value truth, but that does not tell us why we value the true beliefs brought about by that method over true beliefs brought about by other less reliable ones. (1997: 426)

Richard Swinburne (1999) makes essentially the same point:

Now clearly it is a good thing that our beliefs satisfy the reliabilist requirement, for the fact that they do means that . . . they will probably be true. But, if a given belief of mine is true, I cannot see that it is any more worth having for satisfying the reliabilist requirement. So long as the belief is true, the fact that the process which produced it usually produces true belief does not seem to make that belief any more worth having. (1999: 58)

Similar arguments are presented by Linda Zagzebski (1996; 2000; 2003), Wayne Riggs (2002*a*; 2002*b*), Jonathan Kvanvig (2003), and Ernest Sosa (2003). As Kvanvig points out, the common element of these criticisms of reliabilism is the identification of the 'swamping effect' that the value of truth seems to have on the value of reliably acquired belief. Once truth is in place, its value appears to swamp the value of reliability, thus making the combination of truth and reliability no more valuable than truth itself. Accordingly, the argument is often referred to as the 'swamping argument' against reliabilism.

In response to the swamping argument one could point out that few reliabilists have claimed that knowledge amounts to nothing but true belief reliably produced. As we have already noted, reliabilists about knowledge usually insist on an anti-Gettier clause. Adding such a clause opens up the possibility that satisfaction of that clause is what gives reliabilist knowledge its additional value over mere true belief. Even if the value of true belief reliably formed does not exceed the value of true belief *simpliciter*, the value of a true belief reliably formed *in a way*

² For statements of process reliabilism, see Goldman (1979; 1986).

that satisfies the anti-Gettier clause could conceivably exceed the value of a mere true belief. Nevertheless, the idea of knowledge depending on the existence of a reliable connection is the central one behind reliabilism, and it would be unfortunate for the theory if that very component failed to produce an added value. In the following we will, with one exception, be concerned with a simple reliabilist theory according to which knowledge requires the satisfaction of only (1)–(3).

The standard conclusion of the swamping argument is that reliabilism must be rejected. This raises the question of whether other accounts of knowledge can solve the value problem. Swinburne and Kvanvig argue that certain internalist theories fare better in this regard. Kvanvig also thinks—as do Sosa, Riggs, and Zagzebski—that virtue epistemology holds special promise when it comes to accounting for the added value of knowledge. The basic idea here is that S knows that p only if S acquires her belief in p by exercising some epistemic virtue and, furthermore, that a person who knows can therefore be credited for his or her true belief in a way in which a person who has a mere true belief cannot.³ But is it really true that process reliabilism is incompatible with knowledge having an added value? We will, in the course of the paper, explore how process reliabilism can be defended against the challenge posed by the swamping argument.

3. THE SWAMPING ARGUMENT

The standard swamping argument, as endorsed by Jones, Swinburne, and others, runs simply as follows:

- (S1) Knowledge equals reliably produced true belief (simple reliabilism).
- (S2) If a given belief is true, its value will not be raised by the fact that it was reliably produced.
- (S3) Hence: knowledge is no more valuable than unreliably produced true belief.

Since (S3) is a highly counterintuitive conclusion and the argument appears valid, one of the premises must be false. The most common reaction is to reject (S1), that knowledge equals reliably acquired true belief.

Let us take a closer look at the swamping argument. While some theorists, for example, Swinburne, seem to think that this short argument is good as it is, others have tried to present some form of argument for why (S2), the characteristic swamping premise, should be considered true. In Linda Zagzebski's view,

[T]he reliability of the source of a belief cannot explain the difference in value between knowledge and true belief. One reason it cannot do so is that reliability per se has no value

³ Ward Jones (1997) takes the less common position of arguing that the value problem, while problematic, does not show reliabilism to be false.

or disvalue. . . . The good of the product makes the reliability of the source that produced it good, but the reliability of the source does not then give the product an additional boost of value. . . . If the espresso tastes good, it makes no difference if it comes from an unreliable machine. . . . If the belief is true, it makes no difference if it comes from an unreliable belief-producing source. (2003: 13)

What Zagzebski is saying is that the value of a good espresso is not raised by the fact that it was produced by a reliable espresso machine *if taste is all that matters*; and, likewise, the value of a true belief is not raised by the fact that it was produced through a reliable process *if truth is all that matters*. On this view, (S2) depends for its justification on the following additional premise:

(Veritism) All that matters in inquiry is the acquisition of true belief.

Hence, the swamping problem can be seen as arising from combining reliabilism with veritism.⁴ Once a true belief is in place, it does not matter whether it was reliably produced, provided attaining true belief is all we strive for in inquiry. Veritism has been advocated within a reliabilist framework by one of the authors of this paper whose theory is also one of the prime targets of swamping theorists.⁵

The standard swamping argument should be distinguished from the swamping argument offered by Kvanvig (2003). Kvanvig too argues that 'ordinary reliabilist theories of knowledge cannot explain the value of knowledge over true belief' (p. 44). His argument runs essentially as follows:

- (K1) Knowledge equals true belief produced by a process that normally produces true belief (simple reliabilism).
- (K2) Being produced by a process that normally produces true belief just means being likely to be true.
- (K3) The value of having a true belief that is likely to be true is no greater than the value of having a true belief *simpliciter*.
- (K4) Hence: the value of knowledge, reliabilistically construed, is no greater than the value of true belief *simpliciter*.

It is noteworthy that (S2) plays no role in Kvanvig's argument. It is not assumed that if a given belief is true, its value is not enhanced by the fact that it was reliably produced. The crucial premise in Kvanvig's reasoning is rather (K2), which says that being reliably produced just means being likely to be true.⁶ The fact that Kvanvig's version of the argument is essentially different from the standard version seems to have gone unnoticed in the literature.

⁴ For a clear statement of this point, see Percival (2003: 33). As for the role of veritism, see also Jones (1997: 424) and, following him, Riggs (2002b: 82).

⁵ See e.g. Goldman (2002: 53). Goldman's theory is the explicit target of Jones's (1997: 438) and Riggs's (2002b: 80) swamping reasoning. Zagzebski (2003: 14) mentions Alston, Plantinga, Sosa, and Goldman as advocating epistemological theories that are vulnerable to swamping problems.

⁶ Kvanvig commits himself to (K2) in the course of his chocolate analogy at 2003: 478.

Unfortunately for Kvanvig, however, premise (K2) is false. While it is plausible that being produced by a process that normally produces true beliefs implies being likely to be true, the implication does not go in the other direction. Being likely to be true does not imply the existence of a reliable process that produced the belief in question. John may have acquired his belief that he will contract lung cancer from reading tea leaves, an unreliable process, and yet if John is a heavy smoker, his belief may well be likely to be true.

Later in his book, Kvanvig seems to distance himself from (K2). There he equates reliability not simply with (objective) likelihood of truth but with 'objective likelihood derived from the process or methods employed' (p. 49), the suggestion being that reliability is but a 'special kind of objective likelihood' (*ibid.*). This, he goes on to say, does not save reliabilism from the swamping problem because 'once it is assumed that truth is present, this special kind of objective likelihood has no power to increase the value of the composite beyond that involved in true belief itself' (*ibid.*). However, Kvanvig's new proposal is not easy to make sense of. While reliability is a feature of a process—roughly speaking, the feature of leading to beliefs that are mostly true—objective likelihood is rather a property of a belief or proposition. Hence, in saying that reliability is but a special kind of objective likelihood Kvanvig seems to commit a category mistake.⁷

Kvanvig's formulation of the swamping problem is also afflicted by another difficulty. Actually, this problem may be shared by other writers' formulations as well, but it is particularly clear in Kvanvig's case. His formulation focuses on the error of allowing a property of an item whose value is *parasitic* on the value of another property of the item to add value to that item. Here is how the argument goes:

If we have a piece of art that is beautiful, its aesthetic value is not enhanced by having as well the property of being likely to be beautiful. For being likely to be beautiful is a valuable property because of its relationship to being beautiful itself. Once beauty is assumed to be present, the property of being likely to be beautiful ceases to contribute any more value to the item in question. Likelihood of beauty has a value parasitic on beauty itself and hence has a value that is swamped by the presence of the latter. (2003: 45)

Similarly,

... [W]hen the value of one property is parasitic on the value of another property in the way that the likelihood of X is parasitic on X itself, the value of the first is swamped by the presence of the second. So even if likelihood of truth is a valuable property for a belief to have, adding that property to a belief already assumed to be true adds no value to the resulting composite that is not already present in true belief itself. (2003: 45)

⁷ Consider the following claim: (K2*) Being produced by a process that normally produces true belief entails being likely to be true (unless we have explicit reasons for thinking that the belief is false). This claim is true. Moreover, substituting it for (K2) in Kvanvig's argument makes that argument difficult to separate from the standard swamping argument.

As stated, this argument doesn't work. Here is an example that shows why. Suppose you are offered a choice between options (A) and (B).

- (A) Having one thousand dollars.
- (B) Having one thousand dollars plus having a lottery ticket with a 10 percent chance of winning another thousand dollars.

A swamper who appeals to the principles Kvanvig lays down in the previously quoted passages could argue as follows. Option (B) offers outright possession of a thousand dollars plus a certain probability of acquiring a thousand dollars. But the property of having the chance of acquiring a thousand dollars is parasitic on the property of having a thousand dollars. So the value of this property cannot add to the value of the first property. Thus, option (B) is no more valuable than option (A).

This argument is absurd, of course. Where does it go wrong? It follows Kvanvig's formulation in posing the issue in terms of *properties* and their values, where the value of one property is parasitic on the value of another. The idea can be formulated as follows:

- (Property Parasitism) If the value of property P* is parasitic on the value of property P, then the value of P and P* together does not exceed the value of P.

As the money example demonstrates, however, this cannot be a correct formulation. The point of the principle is to avoid the mistake of *double counting*. If the value of one item is wholly *derived* from the value of a second, we don't want to count the derived value in addition to the original, or *fundamental*, value. Here is another example. Suppose you own a lump of gold, which you keep in a safe-deposit box at a bank. You receive a certificate from the bank that specifies that the contents of the box belong to you, and insures those contents. This certificate, in a sense, has value. But its value is wholly derived from the value of the lump of gold. The certificate doesn't add anything to the value of the box's contents. It would be a double-counting error to suppose that having the certificate doubles the value of what you own. The lottery ticket example, however, is entirely different. The probability of getting the second thousand dollars isn't derivative from possession of the first thousand dollars, because the two quantities of money are distinct and independent.

It's a delicate problem to identify an adequate anti-double-counting principle to replace Property Parasitism. We are inclined to think that it must involve some notion of property *instances* (or property exemplifications, or states of affairs) rather than properties per se. But we shall not try to formulate a more satisfactory principle at this juncture. We simply note that Kvanvig's formulation of the swamping argument appeals to Property Parasitism. He writes:

If we want to answer Plato's question about what makes knowledge more valuable than true belief, it is insufficient to cite a further property of knowledge beyond true belief *even*

if that property is itself valuable. The parts may each have value, but when put together, the whole still may have no more value than if one of the parts were missing altogether. (2003: 48)

When formulated in terms of properties and property composites, the swamping objection hinges on the acceptability of the Property Parasitism principle. But that principle, we have seen, is inadequate. So it isn't clear that a compelling challenge has been mounted against the process reliabilist account of knowledge. Nonetheless, we shall attempt to reply to this challenge. Clearly, there is at least a *potential* problem of double counting of value, even if this problem has yet to be formulated satisfactorily.

In the next two sections, we propose two distinct solutions to the EVOK problem from the vantage point of reliabilism. The solutions are independent, but they are also compatible with one another and perhaps complementary. We regard each of the solutions as well motivated, but we recognize that readers may favor one over the other. The authors themselves have different favorites (EJO favors the first solution and AIG favors the second).

Before proceeding, let us back up a bit. The central problem on the table is whether reliabilism can account for the extra value of knowledge as compared with true belief. This problem is common to *all* proffered theories of knowledge. Each must try to explain this extra value. Now, in the case of reliabilism it is generally presupposed that the only way reliabilism can solve the EVOK problem is to say that a reliable process itself has value, of one kind or another. This value can be added to that of the resulting true belief to yield a composite state of affairs (a knowledge state) with more value than the true belief alone. This presupposition about reliabilism's best prospect for solving the EVOK problem is shared by most critics of reliabilism, and perhaps by its proponents as well. If this presupposition is granted, the swamping or double-counting objection immediately comes into play, and supporters of reliabilism are obliged to answer this objection. It is possible, however, for reliabilists to *reject* this presupposition and try to solve the EVOK problem by circumventing the double-counting objection. This approach is adopted in our first solution, which doesn't try to show that the extra value of knowledge is derived from the (token) reliable process that produces the target belief. This solution doesn't solve the swamping or double-counting problem; it just sidesteps it.

4. THE CONDITIONAL PROBABILITY SOLUTION

According to our first solution, if a true belief is produced by a reliable process, the composite state of affairs has a certain property that would be missing if the same true belief weren't so produced. Moreover, this property is a valuable one to have—indeed, an epistemically valuable one. Therefore, *ceteris paribus*,

knowing that *p* is more valuable than truly believing that *p*. What is this extra valuable property that distinguishes knowledge from true belief? It is the property of *making it likely* that one's future beliefs of a similar kind will also be true. More precisely, under reliabilism, the probability of having more true belief (of a similar kind) in the future is greater conditional on *S*'s *knowing* that *p* than conditional on *S*'s *merely truly believing* that *p*. Let's call this proposed solution to the EVOK problem *the conditional probability* solution. Probability should here be interpreted objectively.

The solution can be illustrated in connection with the espresso example. If a good cup of espresso is produced by a reliable espresso machine, and this machine remains at one's disposal, then the probability that one's next cup of espresso will be good is greater than the probability that the next cup of espresso will be good given that the first good cup was just luckily produced by an unreliable machine. If a reliable coffee machine produces good espresso for you today and remains at your disposal, it can normally produce a good espresso for you tomorrow. The reliable production of one good cup of espresso may or may not stand in the singular-causation relation to any subsequent good cup of espresso. But the reliable production of a good cup of espresso does raise or enhance the probability of a subsequent good cup of espresso. This probability enhancement is a valuable property to have.

The following example shows that reliable production of true belief is no different from reliable production of good espresso when it comes to probability enhancement. Suppose you are driving to Larissa but are at loss as to which turns to take at various crossroads. On the way to Larissa there are two forks. If you choose correctly on both occasions, you will get to Larissa on time. If not, you will be late at best. Your only assistance in forming beliefs about the right ways to turn is the on-board computerized navigation system. We consider two situations differing only in that the navigation system is reliable in Situation 1 and unreliable in Situation 2. We assume that in both cases the navigation system tells you correctly how to turn at the first crossroads. In the first scenario this is to be expected, because the system is reliable. In the second it happens by chance. Suppose the correct information at the first crossroads is 'The best route to Larissa is to the right.' Hence in both situations you believe truly that the road to Larissa is to the right (*p*) after receiving the information. On the simple reliabilist account of knowledge, you have knowledge that *p* in Situation 1 but not in Situation 2. This difference also makes Situation 1 a more valuable situation (state of affairs) than Situation 2. The reason is that the conditional probability of getting the correct information at the second crossroads is greater conditional on the navigation system being reliable than conditional on the navigation system being unreliable.

We said that the conditional probability approach bypasses the swamping, or double-counting, problem. How does this transpire? As presented, the conditional probability approach is silent about the value that attaches to the reliable process

per se (as opposed to the value that attaches to the state of affairs of knowing). It is equally silent on the legitimacy of *adding* any value that attaches to that process to the value of the true belief in order to obtain a new value that exceeds that of the true belief. It simply doesn't address these issues. Instead, it looks directly at the composite state consisting of knowing (by means of a reliable process or method) and compares its value to the composite state consisting of truly believing (without arriving at that belief by means of a reliable process). The solution contends that, other things being equal, the former composite state has a valuable property that the latter composite state lacks. Moreover, we might remark, the value of this property is not already contained in the value of the true belief that helps constitute the knowledge state. Thus, there is no way for a critic of reliabilism to re-introduce the swamping problem for the conditional probability approach.

Obviously, the extent to which a knowledge state enhances the conditional probability of future true beliefs depends on a number of empirical regularities. One is that people seldom face unique problems. Once you encounter a problem of a certain type, you are likely to encounter a problem of the same type at some later point. Problems that arise just once in a lifetime are relatively rare. In our navigation example, the question of what is the best turn for driving to Larissa occurs more than once. Another observation is that if a particular method successfully solves a problem once, this method is usually available to you the next time around. In our example, you use the navigation system to solve the problem of what road to take at the first crossroads. This method is also available to you when the same question is raised at the second crossroads. A further empirical fact is that, if you have used a given method before and the result has been unobjectionable, you are likely to use it again on a similar occasion, if it is available. Having invoked the navigation system once without any apparent problems, you have reason to believe that it should work again. Hence, you decide to rely on it also at the second crossroads. Finally, if a given method is reliable in one situation, it is likely to be reliable in other similar situations as well. Let us refer to these four empirical regularities as *non-uniqueness*, *cross-temporal access*, *learning*, and *generality*, respectively.

To see what roles these regularities play, suppose S knows that p. By the reliabilist definition of knowledge, there is a reliable method M that was invoked by S so as to produce S's belief that p. By non-uniqueness, it is likely that the same type of problem will arise again for S in the future. By cross-temporal access, the method M is likely to be available to S when this happens. By the learning assumption, S is likely to make use of M again on that occasion. By generality, M is likely to be reliable for solving that similar future problem as well. Since M is reliable, this new application of M is likely to result in a new true belief. Thus the fact that S has knowledge on a given occasion makes it to some extent likely that S will acquire further true beliefs in the future. The degree to which S's knowledge has this value depends on how likely it is that this will happen. This,

in turn, depends on the degree to which the assumptions of non-uniqueness, cross-temporal access, learning, and generality are satisfied in a given case.

Clearly, no corresponding conclusion is forthcoming for unreliably produced true belief. While non-uniqueness and cross-temporal access are usually satisfied quite independently of whether or not the method used is reliable, there is no reason to believe that an unreliable method that yields a correct belief on its first occasion of use will also yield a correct belief on the second occasion. This blocks the step from the availability of the method on the second occasion to the likely production of true belief on that occasion.⁸

On the conditional probability solution, knowledge has its extra value provided that a number of empirical conditions are satisfied. They are the conditions of non-uniqueness, cross-temporal access, learning, and generality. These conditions are normally satisfied, but we would be hard-pressed to claim that those conditions always hold. When they fail to hold, knowledge fails to have an extra value in the present sense. Is this a flaw in the account? Should it follow from an account of the extra value of knowledge that knowledge *always* has this extra value? This is a matter of controversy. Several authors express satisfaction with weaker conclusions. Swinburne (1999: 64) concludes that knowledge has an added value by arguing that it has this value 'almost always.' Williamson (2000: 79) maintains that knowledge is more valuable provided that the cognitive faculties of the knower are in good order, a condition that may occasionally fail to hold. Finally, Percival (2003: 38) thinks that what needs to be shown is that knowledge has added value 'by and large.' Clearly, these authors think that the value problem can be solved without there being a need to show the greater value of knowledge in every single case. For these authors, the conditional probability solution should be deemed satisfactory in this respect. However, other authors disagree. Riggs conceives of the 'value principle' as '[k]nowledge is *always* more valuable than (mere) true belief' (2002*b*: 79, emphasis added). Kvanvig similarly insists that what is to be certified is the '*unqualified* value of knowledge over true belief' (2003: 57, emphasis added). These authors would consider the conditional probability solution incapable of accounting for the whole of the value of knowledge.

⁸ Bits and pieces of what we have called the conditional probability solution can be found in different places in the literature. It is noted in passing by Armstrong (1973) in his reply to an objection raised by Deutscher. There Armstrong also acknowledges the importance of generality. Williamson (2000: 100–2) presents an account similar to our conditional probability account. However, he focuses on the special case of beliefs with temporally related contents. The approach has rarely been invoked in connection with the swamping problem for reliabilism. For instance, there is no discussion in Kvanvig (2003), which is otherwise well informed by the existing literature. Jones (1997) does attend to the present strategy but fails to see its full import. In Jones's view, the proposal, which he misrepresents as a purely social account of the value of knowledge, can 'only explain the value we place on knowledge in other people' (1997: 430) and is 'of no help in explaining why we value our own knowledge' (ibid.). Our Larissa example shows Jones to be in error. The fact that you have reliabilist knowledge of which way to turn at the first crossroads is better *for you*, the agent, for it makes it more likely that you will acquire further true beliefs in the future.

Still, it is far from clear that our pre-systematic thinking on the matter demands that knowledge always be more valuable than mere true belief. Most generalizations we subscribe to are arguably of a 'typical' rather than an 'absolute' kind. Money is a valuable thing to have, yet rich people are sometimes killed because of their wealth; so for them money was actually something bad. Birds fly, yet birds with feathers covered in oil don't; and so on. The generalizations we make in our daily lives are not universal generalizations in the sense of predicate logic but elastic generic claims that can survive a limited number of counter-instances. If so, why should our claim that knowledge is more valuable than mere true belief be any different?⁹

5. VALUE AUTONOMIZATION

The conditional probability solution explains why reliabilist knowledge is *normally* but not *always* more valuable than mere true belief. We have also given some significant reasons for thinking this is all that needs to be shown. Still, some philosophers think that we are *always* prepared to attribute greater value to knowledge than to mere true belief. If this is a general attributional tendency of ours, does reliabilism have the resources to explain why this is so? In this section we try to offer an explanation of this sort. Notice that we have also slightly shifted the question to the question of value *attribution*. A rationale for this shift will be presented later in this section.

The crucial feature of our proposed explanation is an elucidation of the psychological mechanisms whereby reliable belief-forming processes come to be accorded 'autonomous' value. Although the ascribed value of reliable processes is initially derivative from the ascribed value of the true beliefs they cause, reliable processes ultimately acquire autonomous value—value that isn't dependent, on a case-by-case basis, on the value of resultant true beliefs. A subsidiary part of the solution is hinted at by a phrase in the preceding sentence ('on a case-by-case basis'), which indicates that the value of a token reliable process isn't a function solely of that token's own consequences. Rather, the value imputed to a token process is inherited from the value imputed to its type, a possibility that seems to be ignored by the swamping argument. This approach might be labeled 'type-instrumentalism' as opposed to 'token-instrumentalism.'

The swamping argument assumes that, according to reliabilism, the value of a token reliable process derives from the value of the token belief it produces.

⁹ Cf. Jones (1997: 434): 'I value going to fairs because I have fun when I go to them, even though I can distinctly remember occasions when I got sick on the rides and did not have any fun at all. The fact of my having fun at fairs is responsible for the value I place on fairs, but my having fun is only a contingent property of my attending them. Knowledge is like fairs. We value them both even though we do not always get what we want from them.'

The argument doesn't highlight the type-token contrast, but it hinges on the token interpretation. The argument contends that the value accruing to a token reliable process cannot be *extra* value, because it is wholly traceable to the token belief that it causes. So the value of the process token is analogous to the value of the certificate in our safe-deposit box example, which is wholly traceable to the lump of gold. As in the certificate-gold example, it would be double-counting to add the value of the process token to the value of the true belief token.

Our present solution to the swamping problem, by contrast, involves the claim that instrumental value isn't imputed exclusively because of a singular causal relation between a token instrumental event and a token result. There is a second kind of instrumentalism-based value inheritance. When tokens of type T_1 regularly cause tokens of type T_2 , which has independent value, then type T_1 tends to inherit (ascribed) value from type T_2 . Furthermore, the inherited value accruing to type T_1 is also assigned or imputed to each token of T_1 , whether or not such a token causes a token of T_2 .

An example of value inheritance that fits the type-instrumentalist pattern is money. Money (especially paper money) is not the sort of thing that has fundamental value. But since possessing money (in non-trivial amounts) frequently produces events or states of affairs that do have fundamental value (pleasure, satisfaction, etc.), possessing money comes to be viewed as an instrumentally valuable type of state. Furthermore, each *token* of this type inherits instrumental value from the type, even if some tokens don't actually cause events or states of affairs with fundamental value. (Either the money isn't spent at all, or it's spent on things that bring no pleasure, happiness, satisfaction, etc.)

How does type-based value inheritance allow us to solve the swamping problem? Does it allow us to say that whenever a reliable process causes a true belief, the compound state of affairs consisting of the process token and the true belief has more value than the state of affairs consisting of the true belief alone? Consider an analogy. Taking aspirin regularly causes headache relief. Since headache relief is valuable, taking aspirin is instrumentally valuable. Now consider a particular token of headache relief produced by taking aspirin. Does the value associated with taking an aspirin *add anything* to the value of the headache relief? Does the compound state consisting of aspirin taking plus headache relief have more value than headache relief alone? It seems not. Contemplate a possible world in which you undergo headache relief at a certain time through no cause at all, and then contemplate another possible world exactly like the first except that the headache relief is caused by aspirin taking. Is the second possible world more valuable than the first? Given a choice, would you prefer to live your life in the second world rather than the first? Presumably not. Spontaneous headache relief is just as good as headache relief caused by aspirin taking. On analogy with this case, why should the use of a reliable process *add* value to that of a true belief? (Thanks to Dennis Whitcomb for this challenge.)

What these cases suggest is that instrumental value, even of a *type*-derived sort, does not *generally* combine with fundamental value so as to increase overall value. However, to say that it doesn't *generally* increase overall value does not entail that it *never* does so. Under what conditions might it do so?

The main possibility we suggest is that a certain type of state that *initially* has merely (type-) instrumental value eventually acquires *independent*, or *autonomous*, value status. We call such a process *value autonomization*. Consider the relationship between (morally) good actions and good motives. The primary locus of moral value, quite plausibly, is actions, for example, acts of duty fulfillment or acts of altruism. Actions, however, are not the only things regarded as morally good or valuable. We also value good motives, for example, a desire to help others in need or a desire to do one's duty. Why do we value such motives? A straightforward explanation is that such motives regularly bring about corresponding actions, actions which themselves are valuable. It is therefore plausible that there is a pattern of inheritance by which value passes from certain types of actions to corresponding motive types, which regularly produce those actions. Notice that a token motive of an appropriate type is regarded as good or valuable even if it fails to produce a good action. For example, a token motive may not generate any action, because there are conflicting considerations that yield indecision. Or it might produce an action that doesn't really assist the intended party, or isn't really what duty requires. Despite failing to produce good consequences of a standard sort, the token motive is still good or valuable, presumably because such value is inherited from the type of which it is a token. And this value it retains autonomously, even without triggering an independently good action.

Good motives or good intentions are naturally thought of as good *in themselves*. Famously, Kant held that *only* a good will, or a good intention, has fundamental moral value. We don't have to join Kant in holding that *only* a good will has such a status. But it is very plausible that good motives or intentions are *among* the things rated as *independently* good. This is confirmed by intuitive judgments to the effect that a compound state consisting of a good motive and a good action is (morally) *better* than a compound state consisting of the same good action done from a bad (or non-good) motive. Apparently, a good motive's value can be added to the value of a good action, suggesting that a good motive doesn't have merely (type-) instrumental value, in contrast with the aspirin case.

Conjoining the elements of type-instrumentalism and value autonomization, we have the main ingredients for an explanation of the greater (assigned) value of knowledge as compared with true belief. Some wrinkles remain to be ironed out, however. According to process reliabilism, it doesn't suffice to turn true belief into knowledge that a reliable process *precede* the true belief; it is critical that the true belief be *caused* (or causally sustained) by such a process. What we have said thus far, however, does not fully accommodate this. A compound state of affairs consisting in a reliable process followed by a true belief will be more valuable

than the same true belief not preceded by a reliable process, and this is so even if there is no causal relation between the two, and hence no knowledge.¹⁰

This is not a severe problem. What we need to ensure is that true belief *caused* by a reliable process is more valuable than true belief merely *preceded* by a reliable process. This can be handled in a fairly ad hoc way, though; it isn't essential to a solution to the swamping problem. We can simply note that our valuations are sensitive to causal linkages between suitable pairs of states. For example, good actions that are *caused* by good motives get higher moral marks than good actions that are merely *preceded* by good motives. The valuation of knowledge comports with this pattern.

Let us return now to the central explanatory construct, value autonomization. Value autonomization is a psychological hypothesis, which concerns our practices of *ascribing* or *attributing* value to various states of affairs. Readers may object that the issue posed by the swamping problem is not a problem of value *attribution*. It concerns objective value, not how people *assign* valuational status. The real question is whether a state of affairs with the status of being merely instrumentally good can objectively change its status to being non-instrumentally, or fundamentally, good. That's not something on which human psychology can shed any light.

Here we beg to differ. In claiming that this or that state of affairs has fundamental or extrinsic epistemic value, what evidence do epistemologists have to rely on? Presumably, their own intuitions and those of others. But are these intuitions (intuitional states) wholly a function of the *objective* value status of various states of affairs? Are we methodologically entitled to help ourselves to that assumption about intuitions? We think not. Although we don't wish to deny categorically that intuitions track objective value facts, the epistemological enterprise needn't collapse if that assumption is mistaken. In explaining epistemic intuitions and how they relate to one another, it makes sense to utilize human psychological propensities to form linked patterns of evaluation. Just as understanding the sources of moral intuitions may profit from an understanding of human psychology, so understanding the sources of epistemic intuitions may profit from an understanding of human psychology.¹¹ The value autonomization hypothesis allows that some states of affairs that at one time are assigned merely instrumental value are 'promoted' to the status of independent, or fundamental, value. If this is right, it is compatible with the hypothesis that such promotion occurs in the case of reliable belief-forming processes. While

¹⁰ How could a belief-forming process fail to cause a belief? Simply through some sort of malfunction.

¹¹ For example, recent studies in psychology and cognitive neuroscience by Joshua Greene and by Jonathan Haidt suggest that moral judgments (or intuitions) are often a product of emotion and affect, even when the respondent considers subtly different hypothetical scenarios. Other studies suggest that moral judgments are influenced by framing effects, a well-documented phenomenon in cognitive psychology. See Greene and Haidt (2002) and Doris and Stich (2006).

many such processes are originally regarded as merely instrumentally valuable to true-belief attainment, they are later upgraded to the status of independent value, thereby accommodating the legitimacy of *adding* their value to that of true-belief outcomes.¹²

Let us now step back from our two solutions and reflect on their general characteristics. Reliabilism is generally considered a form of naturalistic epistemology. What is meant by 'naturalism,' of course, varies widely from writer to writer and from topic to topic. In the context of epistemology at least two kinds of naturalism may be distinguished: *metaphysical* and *psychological* naturalism. Metaphysical naturalism holds that epistemic evaluative facts supervene on natural facts. On this understanding, our conditional probability solution to the EVOK problem is very congenial to metaphysical naturalism. This is because it seeks to explain the value of knowledge by reference to the objective (conditionality) probability of obtaining further true beliefs, given the satisfaction of the four conditions we specified in Section 4.

What does psychological naturalism consist of in epistemology? It's an approach that tries to explain our commonsense epistemic valuations in a scientific fashion, especially by appeal to psychology. The autonomization solution is a proffered explanation of our evaluative practices vis-à-vis knowledge. Although the idea of autonomization hasn't received rigorous empirical support, it's put forward here in the spirit of an empirical treatment of human valuational activities. As such it is obviously congenial to psychological naturalism.

6. OTHER APPROACHES TO THE VALUE PROBLEM

In this final section, we examine selected solutions to the value problem that other writers have offered and explain why ours are preferable. Given the size of the literature, it's impossible to discuss all the work that merits discussion. We highlight those approaches that offer either a sharp contrast with ours or some interesting overlap.

Proponents of virtue epistemology have been in the forefront of emphasizing the challenge to reliabilism posed by the value problem. However, virtue epistemology has many varieties. Keeping things simple, it is instructive to distinguish radical and moderate forms of virtue epistemology (VE). Radical VE tries to distance itself sharply from 'naturalistic' approaches to epistemology such as reliabilism, and models the study of epistemology quite closely on the study of ethics. Moderate VE, by contrast, has closer affiliations with reliabilism; it tends

¹² We concede that we don't have a fully detailed story about when value autonomization occurs, that is, under what conditions a state of affairs initially viewed as instrumentally valuable is subsequently upgraded to independently valuable. Notice that such a change may not occur within an individual's ontology, but may be more of a historical-cultural transition.

to explain epistemic virtues, to a substantial degree, in terms of truth-getting skills. Moderate VE—sometimes called ‘virtue reliabilism’—does not pursue so tight a liaison between ethics and epistemology.

At the most radical end of the spectrum is Zagzebski’s approach to VE, which models VE quite closely on virtue ethics. In our view, this makes for an awkward fit, especially where Zagzebski gives excessive emphasis to motivation and love in the theory of knowledge and exaggerates (by our lights) the role of the voluntary in the epistemic domain. Taking aim at the ‘machine-product’ concept of knowledge that she associates with reliabilism (but never explains very carefully), Zagzebski embraces a much ‘loftier’ agent-act conception of knowledge and epistemic value that requires credit-worthiness and even admirability. She imposes ‘a motivational requirement for getting credit for the truth that involves love of truth’ (2003: 19). This strikes us as an unduly heavy burden on many cases of knowledge, especially unreflective knowledge by animals and young children. Much knowledge with which we credit people and animals is of a fairly rudimentary sort, acquired by unmotivated perception or spontaneous recall that operates independently of any ‘love of truth’. If a dog remembers where it buried a bone, we readily grant that the dog ‘knows’ where its bone is, and this piece of knowledge is valuable. But does this bit of knowledge require some sort of canine ‘agency’? Must the dog perform epistemic ‘acts’ out of a love of truth?¹³ Although our example concerns a non-human animal, we don’t think there’s a big difference between humans and dogs when it comes to (this kind of) memory knowledge; nor do we think that our ordinary conception of such knowledge distinguishes the two.

Zagzebski introduces a motivational theme in her treatment of knowledge, but it’s not clear that she gets any mileage out of it for solving the EVOK problem. She complains that reliabilism’s (alleged) commitment to the ‘machine-product’ model of knowledge precludes a proper account of the relation between knowledge and true belief because the value of a mere cause cannot be transferred to its effect. A reliable process, she says, must be ‘external’ to the true belief it causes, whereas a motive can be ‘internal’ to the agent on which it confers value. Thus, only in the case of a motive can value be transferred to its act. As Philip Percival notes, however, these remarks of Zagzebski’s are ‘little more than gestures. She gives no guidance as to *how* an “internal” connection between motive and act, or a “part–whole” relationship between act and agent, can result in the value of a motive being transferred to its effect’ (2003: 34). Moreover, it is becoming well recognized that an external, or ‘extrinsic,’ event *can* transfer value to another event or object. A widely cited illustration is Princess Diana’s dress, which has more value than an exact duplicate simply because it once belonged to

¹³ Hilary Kornblith’s (2002) treatment of knowledge as a natural phenomenon provides a good antidote to the excesses of Zagzebski’s overly intellectualist picture (although Kornblith does not specifically critique VE).

Diana. Having once belonged to Diana is an extrinsic (or external) rather than intrinsic (or internal) property of the dress.¹⁴

Let us turn now to moderate VE. Some themes in moderate VE are fairly congenial to at least the first approach advanced here. For example, John Greco's (1999; 2000) agent reliabilism emphasizes the stability, or 'non-fleetingness,' of a cognitive skill as essential to knowing. He associates such stability with epistemic virtues and agency, claiming that this requirement goes beyond 'generic' reliabilism. The stability requirement is friendly to our conditional probability solution to the value problem. The greater the stability of a reliable source, the greater the probability that it will be used again in the future in similar cognitive tasks. So this element of moderate VE is congenial to part of our approach.

Notice, however, that stability does not cleanly separate virtue reliabilism from generic reliabilism. Although early forms of process reliabilism (e.g. Goldman 1979; 1986) placed no emphasis on stability, nothing in the spirit of generic reliabilism prevents incorporation of stability into its framework. If the VE idea hadn't made an appearance in the 1990s, nobody would have been surprised if stability had instead surfaced within the ambit of generic reliabilism.¹⁵ On the other hand, Greco hasn't persuaded us that stability, or non-fleetingness, of a cognitive source is a strictly necessary condition for knowing. We can easily generate cases in which knowledge occurs through a fleetingly possessed method or skill, such as cases in which a cognitive skill or method is newly acquired and successfully applied to produce a true belief but then promptly lost through death, stroke, onset of Alzheimer's disease, etc.

Another facet (or family of facets) of virtue reliabilism is its emphasis on agency, attributability, and credit-worthiness. According to Riggs (2002*b*), the value assigned to an epistemic state is a function of the credit deserved by the agent. Such credit is deserved only when the state is arrived at in a sufficiently non-accidental way, a way that constitutes an 'achievement.' This theme has pervaded Ernest Sosa's writing from the early 1990s (Sosa 1991) to his most recent writings. Riggs uses the analogy of two holders of Olympic gold medals: Maude possesses one because she won it at the Olympics, whereas Martin possesses one

¹⁴ The example is due to Rabinowicz and Roennow-Rasmussen (1999).

¹⁵ In discussing Greco's requirement of a stable disposition, Berit Brogaard (2006) discusses Greco's (1999) example of a character Rene, who is reliable only through the mediation of an epistemic guardian angel. Since Rene lacks a *stable* disposition or faculty for getting the truth, says Greco, virtue reliabilism denies him knowledge, whereas generic reliabilism would have to concede him knowledge. Brogaard plausibly argues that such examples do not favor virtue reliabilism over generic reliabilism. She introduces David Lewis's (1980) example of prosthetic vision, and compares it to Rene with his guardian angel. Since virtues can be acquired, according to Greco, and needn't be under our control, virtue epistemology should credit a possessor of a prosthetic eye with knowledge derived by the use of such prosthetic vision. But if that is correct, Brogaard implies, why not say the same for Rene? In short, virtue reliabilism doesn't draw a principled distinction between sources of belief grounded in virtuous abilities and those that are not.

because he found it while taking a stroll through the woods. Clearly, Martin doesn't deserve the same degree of credit (if any) as Maude does for having a gold medal. The situation is parallel for two cases of true belief, one acquired by luck versus another acquired by cognitive skill. Only the former deserves epistemic credit. Sosa's (2003) analogy is that of an archer who hits his target by skill, that is, his virtue, versus hitting it via a lucky gust of wind that carries the arrow off its initial path into the target. In the former case, success is attributable to the archer *qua* agent; in the latter, it isn't so attributable. Analogously, says Sosa, what is (most) valuable in the epistemic sphere is attaining truth by one's own performance rather than by luck or accident.

However, these points seem pretty congenial to generic reliabilism; they are not the special preserve of virtue reliabilism. Even if generic reliabilism doesn't use the same *language* as VE—the language of agency, credit, attributability, etc.—it certainly seeks to exclude luck or accidentality by some permutation of the reliability theme. In particular, reliability theories propose either sensitivity, safety, or the absence of relevant alternatives as forms of non-accidentality required for knowledge. Consider Gettier's (1963) disjunction case, for example. Smith makes an entirely justified sequence of inferences from the evidence that Jones owns a Ford to a belief that Jones actually owns one, and thence to a belief that either Jones owns a Ford or Brown is in Barcelona. The inferences used are highly reliable types of inference. Still, it's only accidental that his final, disjunctive belief turns out to be true. It isn't true because Jones owns a Ford, but because Brown, by sheer coincidence, happens to be in Barcelona. Modal reliabilists try to capture what goes wrong here by invoking either a sensitivity condition, a safety condition, or a relevant alternatives condition. These sorts of conditions are standard tools in reliabilism's toolkit. Although they typically figure as a fourth condition for knowledge rather than a third condition, they are squarely within the spirit of generic reliabilism, and not the special preserve of virtue reliabilism. Virtue reliabilism introduces special *language* to describe these cases, but the nuts and bolts of explaining the conditions usually proceeds in a generic reliabilist fashion, using concepts like belief, truth, and various possible-world permutations.

Moreover, insofar as there is a difference between virtue reliabilism and generic reliabilism, the difference resides in the former trying to specify non-accidentality in terms of the agent's credit-worthy performance (or the like). This added element, however, doesn't really work as an account of non-accidentality that captures knowledge. Consider two cases: a Gettier case like the one just discussed and an exactly similar case in which everything goes smoothly—that is, all of the agent's beliefs are true. The latter case is an instance of knowledge whereas the former isn't. Can this be explained in terms of a difference in credit-worthiness of the agent's performance? No. The agent performs just as well in both cases. He infers the same conclusion from the same justified premises using the same inferential procedures. So, contrary to the claim of virtue reliabilism, we cannot

explain all differences between knowledge and merely (justified) true belief in terms of differences in credit-worthiness of performance.¹⁶

Finally, some virtue reliabilist theories of value seem to embrace ingredients quite similar to those advanced in the present paper. In particular, Sosa (2003) comes quite close to making such commitments. The details of his theory are not entirely clear to us, so we remain unsure just how close are its commitments to ours.¹⁷ At a minimum our proposals might be important supplements to the ingredients he uses explicitly, so we commend them to Sosa as amplifications of his brand of virtue reliabilism (amplifications that don't really depart, as we see it, from generic reliabilism).

Sosa says that our epistemic evaluations express a preference for attaining truth (true belief) by our own virtuous performance rather than by mere happenstance. Thus, there is such a thing as 'performance value.' What is good about performance value, he says, is to be understood in a truth-connected way: '[W]hat is good about this performance value cannot be understood independently of the fundamental value of true believing, and especially of true believing that hits the mark of truth attributably to the agent' (2003: 177). This suggests that performance value is a kind of derivative value. Since Sosa also says that a performance can have performance value even when it doesn't attain its characteristic end (e.g. when it doesn't occur in a proper environment), he implies that performance value attaches in the first instance to a *type*. Here he is in the neighborhood of our type-instrumentalist account. However, Sosa also proceeds to regard performance value as a kind of *intrinsic value*. He writes, 'we seem plausibly committed to the *intrinsic* value of such intellectual deeds. . . . We want . . . to attain truth by our own performance, which seems a reflectively defensible desire for a good preferable not just extrinsically but intrinsically' (2003: 175, italics in the original). Later he speaks of a '*eudaimonist, intrinsic* value of true believing where the agent hits the mark of truth as his own attributable deed' (2003: 177, emphasis in the original). Somehow—Sosa doesn't tell us how—the performance value of a cognitive performance seems to rise from the level of extrinsic value to the level of intrinsic value. We are not sure if Sosa means this, but we feel that the account would be clearer and more persuasive if it were supplemented by the autonomization story we presented above. Autonomization would explain *why* performance value gets to be an intrinsic type of value, although it isn't initially characterized this way. Whether or not Sosa would endorse the autonomization story, we feel that his virtue account would be more

¹⁶ Dennis Whitcomb (2006) makes a similar criticism of virtue reliabilism.

¹⁷ The article in question (Sosa 2003) runs through a number of possible positions on the value problem and distinguishes several kinds of epistemic value. A novel type of value, 'praxical' value, is initially introduced as important to a solution to the value problem, but it seems to fade in importance toward the end of the paper. We are not sure we fully capture Sosa's position vintage 2003 here, but it's the one that seems most promising. For a new formulation (on which we won't try to comment), see (Sosa 2007: ch. 4).

compelling if this story were added. Thus, to the extent that his account is on the right track, it does not differ sharply from the one offered here.

REFERENCES

- Armstrong, D. M. (1973). *Belief, Truth and Knowledge*. Cambridge, MA: Cambridge University Press.
- Beckermann, A. (2001). 'Zur inkohärenz und Irrelevanz des Wissensbegriffs—Plädoyer für eine neue Agenda in der Erkenntnistheorie,' *Zeitschrift für philosophische Forschung*, 55: 571–93.
- Brogaard, B. (2006). 'Can Virtue Reliabilism Explain the Value of Knowledge?', *Canadian Journal of Philosophy*, 36 (3): 335–54.
- DeRose, K. (2002). 'Assertion, Knowledge, and Context,' *Philosophical Review*, 111: 167–203.
- Doris, J., and Stich, S. P. (2006). 'Moral Psychology: Empirical Approaches,' in Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (spring 2008 edition), URL = <<http://plato.stanford.edu/entries/moral-psych-emp/>>.
- Gettier, E. (1963). 'Is Justified True Belief Knowledge?', *Analysis*, 23: 121–3.
- Goldman, A. I. (1979). 'What is Justified Belief?', in G. Pappas (ed.), *Justification and Knowledge*; Dordrecht: Reidel. Reprinted in A. Goldman (1992), *Liaisons: Philosophy Meets the Cognitive and Social Sciences*. Cambridge, MA: MIT Press.
- (1986). *Epistemology and Cognition*. Cambridge, MA: Harvard University Press.
- (1999). *Knowledge in a Social World*. Oxford: Oxford University Press.
- (2002). 'The Unity of the Epistemic Virtues,' *Pathways to Knowledge*. New York: Oxford University Press.
- Greco, J. (1999). 'Agent Reliabilism,' in J. Tomberlin (ed.), *Philosophical Perspectives*, 13, Atascadero, CA: Ridgeview Press.
- (2000). *Putting Skeptics in Their Place*. Cambridge, MA: Cambridge University Press.
- Greene, J., and Haidt, J. (2002). 'How (and Where) Does Moral Judgment Work?', *Trends in Cognitive Sciences*, 6 (6): 517–23.
- Hawthorne, J. (2002). 'Deeply Contingent A Priori Knowledge,' *Philosophy and Phenomenological Research*, 65 (2): 247–69.
- Hetherington, S. (2001). *Good Knowledge, Bad Knowledge*. Oxford: Oxford University Press.
- Jones, W. E. (1997). 'Why Do We Value Knowledge?', *American Philosophical Quarterly*, 34 (4): 423–39.
- Kornblith, H. (2002). *Knowledge and its Place in Nature*. Oxford: Oxford University Press.
- Korsgaard, C. M. (1983). 'Two Distinctions in Goodness,' *Philosophical Review*, 92: 169–96.
- Kvanvig, J. L. (1998). 'Why Should Inquiring Minds Want to Know? *Meno* Problems and Epistemological Axiology,' *Monist*, 81: 426–51.
- (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge, MA: Cambridge University Press.

- Le Morvan, P. (2005). 'Goldman on Knowledge as True Belief,' *Erkenntnis*, 62 (2): 145–55.
- Lenzen, W. (1980). *Glauben, Wissen und Wahrscheinlichkeit*. Vienna: Springer-Verlag.
- Levi, I. (1980). *The Enterprise of Knowledge*. Cambridge, MA: MIT Press.
- Lewis, D. (1980) 'Veridical Hallucination and Prosthetic Vision,' *Australasian Journal of Philosophy*, 58, 239–49.
- Percival, P. (2003). 'The Pursuit of Epistemic Good,' *Metaphilosophy*, 34 (1/2): 29–47.
- Pritchard, D. (2007). 'Recent Work on Epistemic Value,' *American Philosophical Quarterly*, 44: 85–110.
- Rabinowicz, W., and Roennow-Rasmussen, T. (1999). 'A Distinction in Value: Intrinsic and for its Own Sake,' *Proceedings of the Aristotelian Society*, 100 (1): 33–49.
- Riggs, W. D. (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.
- Sartwell, C. (1991). 'Knowledge is Merely True Belief,' *American Philosophical Quarterly*, 28: 157–65.
- (1992). 'Why Knowledge is Merely True Belief,' *Journal of Philosophy*, 89: 167–80.
- Sosa, E. (1991). *Knowledge in Perspective: Selected Essays in Epistemology*. Cambridge, MA: Cambridge University Press.
- (2003). 'The Place of Truth in Epistemology,' in M. DePaul and L. Zagzebski, (eds.), *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Oxford University Press, 155–79.
- (2007). *A Virtue Epistemology*. Oxford: Oxford University Press.
- Stich, S. (1990). *The Fragmentation of Reason*. Cambridge, MA: MIT Press.
- Swinburne, R. (1999). *Providence and the Problem of Evil*. Oxford: Oxford University Press.
- von Kutschera, F. (1982), *Grundfragen der Erkenntnistheorie*. Berlin: De Gruyter.
- Whitcomb, D. (2006). 'Knowledge, Virtue, and Truth' (manuscript in preparation, Department of Philosophy, Rutgers University).
- Williamson, T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.
- Zagzebski, L. (1996). *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge, MA: Cambridge University Press.
- (2000). 'From Reliabilism to Virtue Epistemology,' in G. Axtell (ed.), *Knowledge, Belief, and Character*. Lanham, MD: Rowman and Littlefield.
- (2003). 'The Search for the Source of Epistemic Good,' *Metaphilosophy*, 34: 12–28.

2

Is There a Value Problem?

Jason Baehr

According to the standard conception of the ‘value problem’ in epistemology, the problem originates with a compelling pretheoretical intuition to the effect that knowledge is more valuable than true belief.¹ Call this the ‘guiding intuition.’ The guiding intuition is then thought to motivate a constraint on an analysis of knowledge such that any plausible analysis must entail that knowledge is more valuable than true belief. The ‘problem’ emerges in light of two additional considerations or commitments. The first is that knowledge is roughly justified or warranted true belief.² The second is that, on certain popular accounts of knowledge, the value of justification or warrant appears to be *derivative* from and thus *reducible* to the value of true belief.³ Therefore, if knowledge is justified true belief, these accounts apparently fail to entail that knowledge has value over and above the value of true belief, and so fail to satisfy the relevant constraint. Defenders of the accounts in question have tended to respond by arguing that the value of justification as they conceive of it is *not* entirely derivative from the value of true belief, and hence that their theories do in fact satisfy the relevant constraint.⁴

¹ This is implied by most discussions of the value problem, which treat the claim that knowledge is more valuable than true belief as a kind of truism. See e.g. Linda Zagzebski’s various treatments of the value problem (1996; 2000; 2003). John Greco (this volume) makes this point explicit. Two of the earliest contemporary treatments of the value problem are Zagzebski 1996 and Jones 1997. But traces of the problem go as far back as Plato’s *Meno*.

² Gettier concerns may be set aside in the present context. For it is implausible to suppose that the added value of knowledge relative to true belief consists in the kind of non-accidentality that anti-Gettier conditions are aimed at capturing. For a discussion of this point, see Zagzebski (2000: 117).

³ Reliabilism is the easiest target here. As I explain in more detail below, reliabilists define justification in terms of reliability or truth-conduciveness, the value of which is apparently reducible to the value of true belief. A similar point can be made in connection with several other accounts of justification, including *evidentialism*—at least insofar as the value of having good evidence for one’s beliefs is understood (as it often is—see e.g. BonJour (1985: 7–8), in terms of the resulting likelihood that the beliefs in question will be true. For if justification amounts to the possession of good evidence and having good evidence is valuable because it increases the probability that one’s beliefs will be true, then the value of justification is apparently reducible to the value of true belief.

⁴ See e.g. Riggs (2002), Greco (2003), and Sosa (2003).

I argue here that the value problem conceived in the foregoing way is unmotivated and thus that a good deal of the literature on this problem is fundamentally off-track. Specifically, I argue that when the content of the guiding intuition is clarified it becomes evident that this intuition is incapable of motivating a constraint on an analysis of knowledge. If this is correct, then there is no problem (or at least no problem posed by the guiding intuition) with an analysis that fails to entail that knowledge is more valuable than true belief. I go on to argue, however, that the guiding intuition does motivate a 'value problem' of sorts. This is the problem or project of identifying the complete *range* of epistemic values, that is, the complete range of ways (beyond being true) in which a belief or other epistemic state might be valuable. But this account of the problem differs dramatically from the standard one.

1. THE VALUE PROBLEM AND THE ANALYSIS OF KNOWLEDGE

I begin by further clarifying the nature of the constraint allegedly motivated by the guiding intuition. We noted that according to this constraint any plausible analysis must entail that knowledge is more valuable than true belief. But what does 'entail' amount to here? In short, an analysis of knowledge 'entails' that knowledge is more valuable than true belief if and only if one or more of the conditions specified by this analysis has what might be called 'truth-independent' value, or value that is not reducible to the value of true belief.⁵ Take reliabilism as an example. According to this view, knowledge is essentially justified true belief, and a belief is justified just in case it is produced by a reliable process or faculty.⁶ The problem is that the value of justification conceived in this way is apparently derivative from and thus reducible to the value of true belief. Forming beliefs in a reliable manner is valuable, it seems, simply because doing so makes it likely that these beliefs will be true. Thus reliability is aimed at and apparently gets its value from the value of true belief. It appears, then, that none of the conditions specified by a reliabilist account of knowledge has any truth-independent value and that reliabilism therefore fails to entail that knowledge is more valuable than true belief.⁷

⁵ This may in fact be too narrow a characterization, since the relevant added value of an item of knowledge could, say, supervene on certain of some subset of its parts, the value of which, taken individually, is not independent of the value of true belief. But the present formulation makes the relevant point nicely; and nothing in my argument hangs on the sort of possibility just identified.

⁶ See Goldman (1986) for a classic statement of this view.

⁷ It is not obvious that reliabilism is incapable of overcoming this problem. For instance, Greco (2003) argues that additional value supervenes on the event of reaching the truth *through* or *as a result of* one's cognitive faculties. Knowledge as he describes it is a sort of *achievement* that has value over and above the value of true belief.

It is also critical to note that the constraint in question is purely *formal*, in the sense that it does not require anything of an analysis of knowledge beyond that one or more of the conditions specified by this analysis have truth-independent value. It does not require that the condition or conditions in question have any additional characteristics, for example, that they be 'internal' versus 'external,' a function of a belief-state rather than of how the belief was formed, etc. Again, all that matters is that the analysis identify knowledge with one or more features that *in fact* have truth-independent value.

This aspect of the constraint is evident in the standard methodology for responding to the value problem. Replies to this problem typically amount to little more than an attempt to show that knowledge as the author in question conceives of it has truth-independent value.⁸ The underlying assumption is that, once this has been accomplished, the value problem has been 'solved' or overcome. Again, it does not matter which quality an author identifies as 'the added value of knowledge,' as long as this quality does in fact have truth-independent value. But this is the right way of approaching the value problem only if the constraint in question is formal in the sense just noted. This is an important point, and one that we shall have occasion to return to below.

Now let us examine more closely the *motivation* for this constraint. We have seen that the constraint is thought to be motivated by the guiding intuition. The idea, again, is that from a pretheoretical standpoint, knowledge seems clearly to be more valuable than true belief; and this, it seems, is something that an adequate analysis of knowledge ought to 'accommodate' or 'explain' in the sense just described.

Two important observations regarding this claim are in order. The first is that if the value problem is generated by a constraint on an analysis of knowledge, and this constraint is motivated by the guiding intuition, then clearly there had better *be* a guiding intuition. That is, there had better be a reasonably compelling and widespread intuition to the effect that knowledge has value over and above the value of true belief. Otherwise, there will be no value problem (at least as it is standardly conceived).⁹ Moreover, it is important that the intuition in question be a product of *pretheoretical* or *commonsense* thinking about knowledge: that is, that it not be a mere theoretical *result*. For if we found the claim that knowledge is more valuable than true belief plausible only *after* accepting a particular theoretical account of knowledge, reliabilists and others whose theories

⁸ See e.g. Zagzebski (2000), Riggs (2002), and Greco (2003).

⁹ As the parenthetical remark suggests, my target here is strictly the value problem understood as a problem motivated by the guiding intuition. This is certainly the standard conception. If, however, a different and more principled motivation were available, and if this motivation were to satisfy (or circumvent) the conditions discussed below, then a version of the problem with implications for an analysis of knowledge might exist. I will not pursue this possibility here. I will note in passing, however, that for reasons partially indicated below in connection with the possible generality of the guiding intuition, I am less than optimistic about such an alternative motivation.

of knowledge are threatened by the value problem could simply reject this claim as a product of a mistaken theory. If the guiding intuition were not theoretically neutral, these authors would be under no obligation to try to accommodate it.¹⁰ Therefore, if the value problem (as ordinarily conceived) is genuine, there must be a *widespread, reasonably compelling, pretheoretical intuition to the effect that knowledge is more valuable than true belief*. While I do not wish to deny the existence of such an intuition, it is important to be clear about the critical role it occupies within the standard conception of the value problem.

A second observation is that, while the existence of such an intuition is necessary for motivating the relevant constraint, it is not sufficient. For the specific *content* of the guiding intuition must satisfy certain standards as well. First, this content must be entirely *general*, in the sense that it must (at least implicitly) cover or be applicable to *all* instances of knowledge.¹¹ It must (at least implicitly) be an intuition to the effect that knowledge is *always* or *categorically* more valuable than true belief. This is because an analysis of knowledge is aimed at specifying only the *necessary* or *defining* features of knowledge. It is not aimed at specifying any of its merely contingent or accidental features. Therefore, if the guiding intuition motivates a constraint on an analysis of knowledge, it must be built into the content of this intuition that knowledge is categorically or necessarily more valuable than true belief. This point bears emphasis. Again, if the guiding intuition were just an intuition to the effect that knowledge is *sometimes* more valuable than true belief, or that certain *kinds* or *instances* of knowledge are more valuable than true belief, or that knowledge is more valuable than true belief on account of certain features that some kinds or items of knowledge might *lack*, then it would fail to indicate that knowledge is *always* or *categorically* more valuable than true belief, and thus would fail to motivate a constraint on an *analysis* of knowledge, which again is an attempt to specify the *necessary* (and sufficient) features of knowledge.¹²

¹⁰ Interestingly, this is not how reliabilists or other epistemologists have tended to respond to the value problem. As indicated earlier, many prominent reliabilists have regarded the value problem as a serious challenge and have taken significant pains to show that their preferred versions of reliabilism can overcome it. This suggests that they at least *think* that there is a compelling commonsense intuition to the effect that knowledge is more valuable than true belief.

¹¹ The qualifier 'implicitly' is meant to allow for the possibility that the guiding intuition might motivate the relevant constraint even if it does not *obviously* or *explicitly* have the features necessary for doing so.

¹² It might be thought that the content of the guiding intuition is like that of the judgment that 'dogs have four legs' or some similar generalization. This is not merely a judgment about certain 'kinds or instances' of dogs; nor is it, in any obvious way, a judgment about a certain *subset* of dogs. And yet neither is it (even implicitly) a judgment about *every* dog. This, then, might be thought to suggest a way in which the guiding intuition could motivate a constraint on an analysis of knowledge *without* being general in my sense. But in fact, if the guiding intuition were relevantly similar, it would *not* generate a constraint on an analysis of knowledge. For again, an analysis of knowledge is aimed at specifying the *necessary* conditions for knowledge. And while we accept that 'dogs have four legs,' we would not think to include 'four-leggedness' in an account of the *necessary* conditions for being a dog. Thanks to an anonymous referee for raising this issue.

Second, the content of the guiding intuition must be entirely *formal*, in the sense that it must not provide any indication of *why* or *that in virtue of which* knowledge is more valuable than true belief. This requirement is a result of the fact that the constraint that the guiding intuition is thought to motivate is itself entirely formal. As noted above, the value problem is thought to arise from a purely formal constraint on analysis of knowledge: one that makes no demands on an analysis of knowledge beyond that one or more components of this analysis have truth-independent value. But if the guiding intuition were *substantive*, then presumably the constraint motivated by this intuition would be substantive as well. Suppose, for instance, that the guiding intuition were an intuition (at least implicitly) to the effect that knowledge is more valuable than true belief on account of some reasonably specific feature *F* (e.g. on account of being supported by good evidence). If this were the case, then it would be incumbent upon an analysis of knowledge, not merely to incorporate *a* component that has truth-independent value, but rather to incorporate *F in particular* (e.g. to make the possession of good evidence a necessary condition for knowledge). It would make little sense to hold that knowledge intuitively is more valuable than true belief on account of *F*, that an analysis of knowledge must therefore entail that knowledge is more valuable than true belief, but that in doing so it need not make any reference to *F*. This would be to ignore the very basis of the intuition in question. Therefore, if the constraint at the heart of the value problem is entirely formal, the content of the intuition that motivates this constraint must be entirely formal as well.¹³

We have seen, then, that the value problem (at least as it is ordinarily conceived) is a genuine problem only if (a) there is in fact a widespread commonsense intuition to the effect that knowledge is more valuable than true belief and (b) the content of this intuition is general and formal in the senses just noted.¹⁴

¹³ A value problem of sorts might arise if the guiding intuition had a certain *kind* of substantive content. For instance, if from an intuitive standpoint knowledge appeared more valuable than true belief on *account* of its being the product of a reliable process, this would present a certain puzzle or problem, since again the value of reliability seems not to exceed that of true belief. But the problem here is very different from the value problem as it is standardly conceived. First, for reasons already noted, the relevant intuition would fail to have any implications for an *analysis* of knowledge unless it were also *categorical* in nature. Second, even if the intuition were categorical in nature, the resulting constraint obviously would not be *formal* in nature. Thanks to Adrian Haddock for prompting this point.

¹⁴ Some recent work by Jonathan Kvanvig (2003) might be thought to suggest an alternative motivation for the value problem understood as involving a constraint on an analysis of knowledge, a motivation that requires neither (a) nor (b). At times, Kvanvig seems to endorse the idea that the relevant constraint is motivated directly by what we are calling the guiding intuition (e.g. pp. x, 4). Elsewhere, however, his discussion suggests something like the following principle: an analysis of knowledge (vs. some other concept) is worthwhile only if the value of knowledge exceeds the value of any subset of its parts (pp. x, xiv, 11, 109, 185). If correct, this principle would require that a plausible or worthwhile analysis of knowledge entail that knowledge is more valuable than true belief. But why endorse the principle to begin with? It is by no means obviously true. Kvanvig

2. THE CONTENT OF THE GUIDING INTUITION

I have already indicated that I shall not call into question whether there is a guiding intuition, or an intuition of the sort specified by (a) above.¹⁵ Rather my concern in the present section lies with (b). I argue that the guiding intuition does not have the qualities noted in (b), that is, that its content is neither relevantly general nor relevantly formal, and thus that the value problem as it is ordinarily understood is unmotivated.

Before proceeding, an important methodological concern must be addressed. It might reasonably be wondered whether one can do much in the way of *arguing* for a certain account of the specific content of the guiding intuition (or *any* intuition for that matter). It might be thought that one can do little more than *describe* this content as it appears to one, with the hope that the resulting description will resonate with one's interlocutors. While I agree that the resources for defending a certain view of the precise content or character of the guiding intuition are limited, they still amount to something. Indeed, in what follows, I employ two additional strategies in defense of my account of the content of the guiding intuition. First, I consider several *implications* of conceiving of this content in one way rather than another: implications that are plausible or implausible in their own right. Second, in Section 3, I sketch an alternative account of the intuitive origin of the value problem that coheres well with and helps to explain several of the points pertaining to the content of the guiding intuition arrived at in the present section. When combined with something like the descriptive strategy noted above, these considerations comprise a cogent case for thinking that the content of the guiding intuition is neither general nor formal.

2.1. The Generality of the Guiding Intuition

We may begin by considering how plausible it is to think of the content of the guiding intuition as *general* in the relevant sense. That is, how plausible is it to suppose that when we think about knowledge as more valuable than true belief

himself says very little in support of this principle beyond an appeal to something like the guiding intuition (4). Moreover, even if there *were* an intuition that satisfied the condition noted in (b), Kvanvig's principle would not be well supported. For the guiding intuition thus conceived would entail merely that knowledge has value over and above the value of *true belief*, not over and above the value of 'any subset of its parts.' (See Greco, this volume, for a similar point.) Finally, if the guiding intuition is ultimately supposed to underwrite Kvanvig's principle, then in fact this principle does not yield an *alternative* motivation for the constraint in question at all.

¹⁵ I do have doubts about how univocal the intuition in question is, which could present a major obstacle to thinking of this intuition as motivating a constraint on an analysis of knowledge. But I will set this worry aside in the present context.

in the relevant intuitive or commonsense way, we think of it as *necessarily* or *categorically* more valuable? There may be some initial support for thinking of the guiding intuition in this way, for we do not tend to qualify our acceptance of the claim that knowledge is more valuable than true belief by *saying* or *explicitly thinking* that this claim holds only sometimes or only in relation to certain kinds or instances of knowledge. This, however, is not sufficient for showing that the content of the guiding intuition is general in the relevant sense. At a minimum, it leaves open the possibility that this content has certain *implicit* limits. Specifically, it leaves open the possibility that we regard knowledge (at least implicitly) as more valuable than true belief on the basis of certain features that we do not (intuitively or otherwise) think of as *necessary* features of knowledge. I might, for example, regard knowledge as more valuable than true belief at the relevant intuitive level on account of its being accompanied by good evidence or a kind of 'reflective epistemic perspective,' while denying that such a perspective, strictly speaking, is *required* for knowledge. In this case, my intuition to the effect that 'knowledge is more valuable than true belief' would at bottom be an intuition to the effect that a certain *kind* of knowledge (e.g. 'reflective knowledge') or knowledge that involves certain *properties* (properties which I do not regard as essential to knowledge) is more valuable than true belief. Again, nothing about the apparent character of the guiding intuition seems to rule out this possibility.

In fact, I take it that there is something *prima facie implausible* about the suggestion that when we think of knowledge as more valuable than true belief in the relevant intuitive way, we think of it as necessarily or categorically more valuable. We do not, in other words, tend to treat the relevant claim as an *exceptionless* or *necessary* truth: as applying to any *possible* instance of knowledge. Instead, the judgment in question seems to be something like a broad *generalization*, the content of which may very well be *indeterminate* with respect to whether *in every case* knowledge is more valuable than true belief. This suggests that the alternative, *exceptionless* characterization is too strong, that it attributes more than is warranted to the content of the guiding intuition. Thus, while the guiding intuition may be somewhat general, it does not appear to be completely or relevantly general.

This conclusion is reinforced by the idea that there is nothing independently or inherently counterintuitive in the suggestion that there might exist, say, at least *one* item of knowledge the value of which fails to exceed that of the corresponding item of true belief. I take it, that is, that nothing in our commonsense or pretheoretical way of thinking about knowledge rules out such a possibility. But again, if the guiding intuition is a product of common sense, and if its content is completely general, this suggestion should generate a clash of intuitions; it should strike us as questionable or implausible that *any* item of

knowledge might fail to be more valuable than the corresponding item of true belief. Such a suggestion should be on a par with the suggestion that there might exist a married bachelor or that two plus three might fail to equal five. But clearly it is not.¹⁶

It is, then, at least initially implausible to think that the content of the guiding intuition is relevantly *general*. Thus this intuition apparently fails to yield a reason for thinking that knowledge is *always* more valuable than true belief, which in turn implies that it cannot motivate a constraint on an *analysis* of knowledge (that is an attempt to specify the *necessary* features of knowledge).

Moreover, even if the content of the guiding intuition *were* relevantly general, it would not automatically or unproblematically generate a constraint on an analysis of knowledge. This is true in part because its content must also be *formal*, a point which we shall return to shortly. It is also true, however, because *as a matter of fact*, knowledge is *not* always more valuable than true belief. This is evident, first, in cases in which a demand for knowledge rather than mere true belief would require the forfeiture of certain other important goods. For instance, if trying to flee a certain foreign location in the face of some impending catastrophe, I might do better simply to trust my hunch (which happens to be accurate) about the appropriate way out than to stick around and do what it takes to acquire knowledge about this matter (e.g. find a map, get directions from a reliable passerby, etc.). This suggests that knowledge is sometimes *less* valuable than mere true belief. In response, it might be said that while, from a practical or an all-things-considered perspective, I am better off with mere true belief in this case, *other things being equal*, I am better off with knowledge. But this is far from obvious. Why think I would be better off *knowing*, say, that to escape the imminent catastrophe I need to turn right at this corner and left at the next than I would simply having a true belief about this matter?¹⁷ The subject matter is not of any epistemic interest to me: I care about it only because I care about preserving my life. In light of this, and of the fact that true belief serves

¹⁶ Of course the guiding intuition need not be on a par with these other claims in *all respects*. The idea is simply that we regard the latter claims as *exceptionless*, such that the suggestion that there might be an unmarried bachelor or that two plus three might equal something other than five is *clearly*, indeed *obviously*, problematic. Thus, if we (even implicitly) regard the content of the guiding intuition as exceptionless, the suggestion that there might be a single item of knowledge the value of which fails to exceed that of the corresponding true belief should strike us as similarly problematic.

¹⁷ This is, of course, very similar to the question raised by Socrates in the *Meno*, 97a–98a. Socrates' eventual response to the question (concerning the *logos* or 'account' which he says is a feature of knowledge but not of mere true belief), while perhaps showing why in *general* knowledge is more valuable than true belief, does not show that it is more valuable in the present case. It matters not, for instance, whether my belief about how to escape the present predicament is 'tied down' in the Socratic sense.

my practical interests just as well as knowledge, it is unclear why in this case knowledge should be preferable to mere true belief.

Consider as well cases of so-called 'trivial knowledge,' for instance, knowledge about the number of blades of grass on one's front lawn or the number of names listed under 'C' in the local phonebook. Here the subject matter is likely to be of no interest to me at all (it lacks even the practical payoff of the belief just considered).¹⁸ I am likely to regard knowledge of this subject matter as epistemically neutral, and perhaps even as a waste of cognitive resources.¹⁹ But if the subject matter in question is not an epistemically worthy one, why should knowledge of it be superior or preferable to mere true belief? It might be replied that if I settle for mere true belief in such cases I shall forfeit an opportunity to achieve the good of knowledge. But this, of course, is an unacceptable response, since one of the questions at issue is whether knowledge *is* always a good.

The present point can be made even clearer vis-à-vis other kinds or instances of knowledge. Consider, for instance, much of what is peddled on the evening news (at least here in Los Angeles): the latest scoop about which Hollywood 'celebs' are sleeping with (or divorcing) each other; how these people are dressing their children; or the unsavory details of the latest sex scandal in the local schools. Here it would be an understatement to say that the knowledge in question 'lacks value.' Such knowledge seems positively *disvaluable*: I want *not* to know about the things in question. And the reason is not strictly a moral one. Acquiring or possessing this knowledge need not, say, involve a violation of the subjects' privacy (some of the subjects might even welcome the attention). Even from a strictly 'epistemic standpoint,' or relative to the standard of an excellent intellectual life, such knowledge is problematic: it is apparently a waste or misappropriation of cognitive capacity and resources.²⁰ But if an item of knowledge lacks epistemic value, or indeed is epistemically *disvaluable*, then surely it is not *more* valuable than the corresponding item of mere true belief.

In response, the following question might be raised. Suppose I were forced to choose between mere true belief and knowledge about one of the states of affairs just mentioned (e.g. where Britney Spears was spotted shopping today). Would I not prefer to have knowledge? That is, if I did not have the luxury of

¹⁸ This subject matter *could* be of interest to me: e.g. if I were a telecommunications historian. However, since the claim at issue is whether knowledge is *always* more valuable than mere true belief, we need only identify one instance in which knowing something would fail to be more valuable than having a mere true belief about it. A similar point applies to the other kinds or instances of knowledge discussed below.

¹⁹ Indeed, this is, I take it, the gist of many recent discussions about the 'epistemic goal,' and specifically about whether true belief is *always* valuable. See DePaul and Zagzebski (2003). It is widely held that *not* all true belief is valuable; and instances of trivial knowledge are often cited in support of this claim.

²⁰ Moreover, the point is not merely that it would be a waste of resources *given* the other ways in which these resources might be spent; rather, the suggestion is that the subject matter itself fails to warrant or is unworthy of these resources.

simply ignoring and not forming any beliefs about the relevant subject matter, wouldn't it be better for me to have knowledge of this subject matter rather than mere true belief? It might be thought that knowledge would be preferable on the grounds that otherwise I would be settling for a second-rate cognitive situation with respect to the beliefs in question. But again, this is so only if knowledge is *always* more valuable than mere true belief, which is precisely the issue under discussion. Indeed, from an intuitive standpoint, I see no reason to think that I should prefer knowledge to mere true belief in this case. Since I want *not* to know about these things, I might even regard mere true belief as the lesser of two cognitive evils.

This leads to a further, more principled consideration. It is plausible to think that knowledge tends to make a greater draw on our cognitive capacity, resources, abilities, and the like, than mere true belief.²¹ It is also plausible to think that for subject matters that are epistemically unworthy or disvaluable, we would do well (*ceteris paribus*) to devote as little of our cognitive capacity, as few of our resources, etc., as possible to such subject matters. While these claims could be explored in more detail, I take it that they have considerable intuitive appeal. If they are correct, then relative to many of the subject matters in question, knowledge is *less* valuable than mere true belief.

The foregoing considerations provide good reasons for thinking that knowledge is not always or categorically more valuable than true belief. It follows that even if the content of the guiding intuition *were* general in the relevant sense (and again we have examined some good reasons for thinking it is not), this intuition could not be taken at face value; it could not immediately or unproblematically motivate a constraint on an analysis of knowledge.²²

2.2. The Formality of the Guiding Intuition

Let us now turn to consider whether the content of the guiding intuition is relevantly *formal*. How plausible is it to suppose that when we think of knowledge as more valuable than true belief at the relevant intuitive level, we are thinking of it *purely* in the abstract, without any (even implicit) reference to any of the features *in virtue of which* it is apparently more valuable? In fact this seems quite implausible. That is, it is implausible to think that the relevant intuitive

²¹ The plausibility of this claim may depend on how exactly one is thinking about knowledge. But it would seem to hold for at least many of the more popular accounts of knowledge. For instance, on a 'success through cognitive virtues or abilities' account of knowledge (à la Sosa), knowledge involves an investment or contribution by one's cognitive virtues or abilities that 'mere true belief' may very well lack. Similarly, on a grounds-based or 'evidentialist' account of knowledge, knowledge is likely to require greater effort and resources (those required for having good grounds) than mere true belief. Thus I take it that the claim is broadly plausible.

²² Whether a modified version of the constraint might still be motivated is, I take it, an open question. However, I will not pursue this question here. I am grateful to Rusty Jones for a helpful conversation about the issues dealt with in this section.

preference for knowledge over true belief has *nothing* whatsoever to do with any reasonably specific *features* that we take knowledge to have. This is not to deny that our grasp or awareness of these features may be somewhat implicit or fuzzy; nor that the content of the guiding intuition may to some extent be indeterminate on this point. But all of this is entirely consistent with the possibility that when we value knowledge at the relevant intuitive level we do so (at least implicitly) on the *basis* of its seeming to us to be a certain way or to have certain reasonably specific features. Thus, while the content of the guiding intuition may be somewhat formal, it is apparently not *entirely* or *relevantly* formal. If so, it cannot motivate a purely formal constraint on an analysis of knowledge.

This point can be drawn out in several ways. First, we can ask whether it might ever be reasonable to object to a proposed reply or solution to the value problem for any reason *other* than that the reply identifies the added value of knowledge with a feature that lacks truth-independent value. For instance, consider Linda Zagzebski's reply to the value problem.²³ She defines knowledge (roughly) as true belief that results from virtuous intellectual motives and actions; and she maintains that the motives in question have intrinsic value. Accordingly, she claims that knowledge is more valuable than true belief on account of these motives. Suppose we grant that the motives in question are intrinsically valuable. Are there no possible grounds for objecting to Zagzebski's proposed solution to the value problem? Or rather, couldn't someone object by claiming, say, that the relevant motives—valuable as they may be—cannot really explain or make sense of the intuitive added value of knowledge, in the sense that they cannot really be what we have in mind when we regard knowledge as more valuable than true belief in the relevant intuitive way? I see no principled problem with this sort of objection (indeed it strikes me as quite plausible). However, if the guiding intuition were entirely formal, this objection would make little sense. For it assumes that we regard knowledge as more valuable than true belief (at least implicitly) on account of certain of its specific features but not others.

A second and related way of drawing out the relevant point is to ask whether it makes sense to think that we might make progress in our understanding of *why* knowledge is more valuable than true belief by further reflecting on or scrutinizing the content of the guiding intuition. Here the question is whether the content of the guiding intuition *itself* might provide at least some indication of the relevant added value of knowledge. Again, I see no reason to rule out such a possibility; in fact, it seems to me an entirely plausible way to proceed. If we intuitively regard knowledge as more valuable than true belief, and want to understand why it is more valuable, it seems entirely sensible to focus on the intuitive basis of this judgment: for example, to ask what it is about knowledge as we conceive of it that apparently leads us to think of knowledge as more valuable

²³ Zagzebski (1996: 300–4; 2000; 2003).

than true belief. But again, if the content of the guiding intuition were purely formal, this would be a hopeless and misdirected endeavor, for *ex hypothesi* it would fail to provide us with any of the sought-after information.

A third point is related to the foregoing two. If the value problem is rooted in a purely formal intuition to the effect that knowledge is more valuable than true belief, it follows (implausibly) that the two possibilities just considered are even more misguided than has been suggested thus far. To see why, note that the possible reply to Zagzebski and the possible method of ascertaining the added value of knowledge just considered both presuppose that there is *an* added value of knowledge, that is, that there is some reasonably determinate property or fairly limited set of properties in virtue of which we intuitively regard knowledge as more valuable than true belief. The problem is that, if the value problem is rooted in a purely *formal* intuition to the effect that knowledge is more valuable than true belief, then there are in fact *any number* of possible 'solutions' to this problem: any number of 'right answers' to the question of what makes knowledge more valuable than true belief. Put another way, there is no very determinate or univocal 'added value of knowledge' at all. It follows that the questions of whether it makes sense to reply to Zagzebski in the manner noted above and whether we might hope to make some progress in our understanding of why knowledge is more valuable than true belief by reflecting further on the content of the guiding intuition are fundamentally confused. For again, both of these questions presuppose that there *is* a reasonably determinate and univocal 'added value of knowledge.' But this is a very implausible implication, for neither of the relevant questions or proposed ways of addressing the value problem *seems* confused or misguided at all. This, then, is a further indication that it would be a mistake to think of the content of the guiding intuition as entirely formal.

A fourth and final point in favor of this conclusion emerges in connection with cases like the following. Imagine a world in which, owing to the work of a mildly malevolent Cartesian demon, human beings are capable of nothing more than so-called 'animal knowledge,' or knowledge of the sort possessed by higher animals and small children; such knowledge represents the upper boundary of human cognitive achievement. Does the guiding intuition hold relative to worlds like this? If human cognition were limited in this way, would it still be clear from the relevant commonsense standpoint that knowledge is more valuable than true belief?²⁴ I take it that, at a minimum, the intuition would be considerably less firm than it is in relation to the actual world. That is, it would at least be questionable from a commonsense standpoint whether

²⁴ Perhaps it *would* be clear from certain theoretical standpoints: e.g. one's preferred philosophical analysis of knowledge might entail (as perhaps Greco's in 2003 does) that even animal knowledge has value over and above the value of true belief. But our concern at present is with the content of the guiding intuition and so is limited to the value of knowledge examined from a pretheoretical or commonsense standpoint.

knowledge is in fact more valuable than true belief. If so, this shows that our actual, rather firm intuitive preference for knowledge over true belief is (at least implicitly) a response to certain epistemic features absent from animal knowledge but instantiated by a higher-grade, 'human' or 'reflective' kind of knowledge. And again, if this is right, then the content of the guiding intuition cannot be entirely formal.

We have examined several good reasons for denying that the content of the guiding intuition is formal or general in the relevant senses. It appears, then, that this intuition cannot motivate a constraint on an analysis of knowledge. This in turn shows that the value problem as it is standardly conceived is unmotivated and thus that a good deal of the literature on the value problem is fundamentally off-track. Before moving on, it is worth noting that this conclusion follows even if the content of the guiding intuition were to lack just *one* of the two features discussed above. In other words, the viability of the value problem requires that the content of the guiding intuition be *both* relevantly general *and* relevantly formal. Therefore, even if one were convinced that, say, the content of the guiding intuition is in fact general in the required sense, one would not be entitled to conclude that this intuition motivates a constraint on an analysis of knowledge.

2.3. A Diagnosis

If the characterization of the guiding intuition that has emerged in this section is correct, what explains the nearly universal tendency to think of the value problem as involving a constraint on an analysis of knowledge? I think an answer is suggested by the fact that those who have addressed the problem appear not to have thought very carefully about the specific character and content of the guiding intuition; instead they have tended to treat it as a kind of necessary, self-evident truth.²⁵ This is suggested by the standard method of addressing the value problem. As alluded to earlier, the standard approach is to offer a quick and uncritical acknowledgment of the fact that 'knowledge is more valuable than true belief' and then to turn immediately to the project of showing that knowledge as one conceives of it has value over and above the value of true belief. If the guiding intuition were something like a necessary, self-evident epistemic principle, then this approach would make good sense. It would also warrant thinking of the value problem as involving a constraint on an analysis of knowledge, for if the guiding intuition had this status, then presumably it would be general and formal in the relevant senses. We have

²⁵ Strictly speaking, the claim is that the authors in question treat the *object* of the guiding intuition as if it were a necessary, self-evident truth. The intuition itself, as I am thinking of it here, is a psychological state that involves grasping or apprehending this proposition (it is not itself a proposition).

seen, however, that it is a mistake to think of the guiding intuition in such terms. So again, the problem is apparently that epistemologists have failed to pay sufficient attention to the specific character of the guiding intuition and the implications of this character for a proper conception of the value problem.

3. AN ALTERNATIVE VERSION OF THE VALUE PROBLEM

Before dismissing the value problem as nothing more than a pseudo-problem, it is worth considering what a more plausible conception of the guiding intuition might look like and whether this conception might motivate an alternative version of the value problem.

3.1. An Alternative Conception of the Guiding Intuition

In some recent work on the value problem, John Greco (following Socrates in the *Meno*) argues that the problem begins, *not* with the question of why knowledge is more valuable than true belief, but rather with a more basic question: namely, ‘Why is knowledge valuable? Or perhaps better, What is it that makes knowledge valuable?’²⁶ One obvious reply is that knowledge is valuable because it involves true belief (or, as Greco describes it, ‘true information’), which itself is valuable. But, as Greco (again following Socrates) points out, this cannot be a fully adequate answer, for ‘we think that knowledge is more valuable than *mere* true information, or true information that is not knowledge.’²⁷ On Greco’s view, then, the value problem begins with a question about the value of knowledge *per se* and then *shifts* to a question about the comparative value of knowledge and mere true belief.

What does Greco’s characterization imply about the *intuitive basis* of the value problem? It implies, first and most obviously, that the ‘guiding intuition,’ which again is the intuition that is supposed to get the value problem up and running, is not merely a judgment about the comparative value of knowledge and true belief, but rather an intuition rooted in a judgment about the value of knowledge *per se*. In particular, this intuition is rooted in something like the rather ordinary and familiar pretheoretical judgment or general conviction that ‘knowledge is a valuable thing,’ that it is an estimable and worthy human good. For it is presumably a judgment of this sort—and not, say, some abstract, technical, or philosophical judgment—that motivates Greco’s initial question concerning the value of knowledge.²⁸ This initial, commonsense judgment leads to a judgment

²⁶ Greco (this volume, 313). ²⁷ Ibid.

²⁸ Admittedly I am reading into Greco’s discussion somewhat, since he does not discuss the underlying intuition as such. However, it is entirely plausible to think that what he does say

about the comparative value of knowledge and mere true belief once we attempt to identify *why* or *that in virtue of which* knowledge is valuable. Again, this is because, while the fact that knowledge involves true belief may *partly* explain our high regard for it, this fact fails to provide a *complete* explanation. Put another way, it is because, at the relevant pretheoretical or commonsense level, we think of knowledge as valuable in ways that go *beyond* the value of true belief.

Notice how different this characterization of the guiding intuition is from the standard characterization. Here the guiding intuition clearly does not amount to an apprehension of a necessary, self-evident principle to the effect that 'knowledge is more valuable than true belief.' It is rather a more pedestrian or 'folksy,' but also more complex, psychological state. Indeed, it amounts to a kind of intuitive *inference*: an inference based upon a very ordinary pretheoretical judgment about the value of knowledge together with an additional judgment to the effect that this value is not exhausted by the value of true belief. The inference has the following general structure: (1) Knowledge is highly valuable; (2) this value apparently is not reducible to the value of true belief; (3) therefore, knowledge apparently has value over and above the value of true belief.

This way of thinking about the guiding intuition coheres well with certain of the conclusions reached earlier concerning its content. First, when understood in the present way, it is highly doubtful that the content of the guiding intuition is entirely *general*. For there is little reason to think that, the relevant, rather ordinary and folksy judgment about the value of knowledge (corresponding to (1) above) is itself entirely general. That is, there is little reason to think that, when we judge knowledge to be valuable in the relevant intuitive way, we take this judgment (even implicitly) to apply to any and every *possible* item of knowledge. Rather, this judgment is more plausibly understood as a kind of rough *generalization* or judgment to the effect that *in general* knowledge is highly valuable. If this is right, then the scope of the resulting *comparative* judgment (captured by (3)) is limited as well. Second, it is also highly doubtful that the guiding intuition, when understood in the suggested way, is entirely *formal*. For it is plausible to think that when we regard knowledge as valuable at the relevant intuitive or commonsense level (as per (1)), we do so on account of one or more reasonably specific *features* that we take knowledge to have. We do not value knowledge simply in the abstract, but rather, presumably, on account of what (in particular) it *is* or on account of how it is *related* to other things we value. But again, if this is right, then the resulting judgment to the effect that knowledge has value over and above the value of true belief is at least somewhat substantive. It is not purely or relevantly formal.

presupposes the present account of this intuition. For a similar account of the underlying intuition, see Kvanvig (2003: 4), and Zagzebski (2000: 122).

3.2. The 'Value Pluralism' Conception of the Value Problem

Suppose, then, that we think of the guiding intuition along the lines just sketched. Is there a 'value problem'? It depends on how broadly we are willing to use this term. Let us refer to the standard conception of the value problem (criticized above) as the 'formal constraint' conception, since its thrust is a formal constraint on an analysis of knowledge. Clearly the guiding intuition conceived in the present way does *not* motivate the formal constraint conception, for again, the content of this intuition lacks the required generality and formality. The intuition does, however, motivate an alternative problem. For it provides at least an initial indication that truth is not the sole epistemic value. Thus it gives rise to the following question: *What is the full range of epistemic values? That is, what are the various ways in which an item of knowledge might be valuable (beyond its being true)?* This is an interesting and substantive question, and one that seems clearly to indicate a 'value problem' of sorts. Let us then refer to this as the 'value pluralism' conception of the value problem.

This conception is considerably broader and more open than the formal constraint conception. First, there is no reason to think its focus must be limited to the *essential* or *defining* features of knowledge. It is not necessarily concerned only with ways in which an item of knowledge *must* be more valuable than the corresponding item of true belief, but also with ways in which it *might* be more valuable. In fact, the central question could easily be restated without any reference to knowledge at all. It might be put thus: What is the full range of ways (beyond being true) in which a *belief* or related epistemic state might be valuable? Alternatively: What is the full range of values that might accrue to such a state? Second, while on the value pluralism conception the value problem is concerned with the various ways in which a belief or related state might be valuable, there is no reason to think that the ways in question must themselves be 'epistemic.' They might, for example, be instances of pragmatic, moral, aesthetic, or other kinds of value. For again, the concern here is with the *full* range of ways (whether epistemic or not) in which beliefs or related states might be valuable.²⁹ Third, on the value pluralism conception, work on the value problem is less likely to be fraught with a certain kind of theory-laden disagreement. Suppose that I hold to a certain account of knowledge A, that according to A a belief counts as knowledge only if it has some feature F, and that F has truth-independent value. I will naturally draw the conclusion that knowledge is more valuable than true belief on account of its involving F. Suppose, however, that many of my fellow epistemologists reject A in favor of other accounts of knowledge, according

²⁹ As this suggests, when I speak of the full range of 'epistemic values,' I am referring to values that might be instantiated by *epistemic* states; the values *themselves* need not be of an epistemic variety. If this were somehow too permissive, the value pluralism conception could easily be restricted accordingly.

to which F is *not* an essential feature of knowledge. Within the framework of the formal constraint conception, my colleagues are likely to reject the thesis that 'knowledge is more valuable than true belief on account of its involving F' because they deny that F is an essential feature of knowledge in the first place. Under the value pluralism conception, however, this disagreement need not arise. For as long as they agree that F has truth-independent value, my colleagues can readily agree that instantiating F is one way (beyond being true) in which a belief can accrue value. Their agreement does not commit them to thinking of F as an essential feature of knowledge. This difference between the two conceptions is a function of the fact, noted above, that on the formal constraint conception, but not on the value pluralism conception, the value problem is concerned strictly with certain essential or defining features of knowledge.

Central to the value pluralism conception of the value problem, then, is the project of elucidating the full network of values that might be instantiated by a belief or related epistemic state, regardless of whether these values are instantiated by *every* instance of knowledge or are themselves 'epistemic' in nature. The aim is to identify all of the relevant values, to get a better sense of their character, and to grasp their interrelations. Understood in this way, the value problem is well motivated and stands to inspire some interesting and original work in epistemology.³⁰

REFERENCES

- BonJour, L. (1985). *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.
- DePaul, M., and Zagzebski, L. (eds.) (2003). *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Oxford University Press.
- Goldman, A. (1986). *Epistemology and Cognition*. Cambridge, MA: Harvard University Press.
- Greco, J. (2003). 'Knowledge as Credit for True Belief,' in DePaul and Zagzebski 2003. — (2009 this volume). 'The Value Problem.'
- Jones, W. (1997). 'Why Do We Value Knowledge?' *American Philosophical Quarterly*, 34: 423–39.
- Kvanvig, J. (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge: Cambridge University Press.

³⁰ This paper has been in the works for some time and has benefited from many conversations with colleagues, friends, and a few (patient) family members. Of special note are extended and especially helpful conversations with Michael Pace, Dan Speak, and Dan Yim. Versions of this paper were read at the 2004 Southern California Philosophy Conference and at a conference on epistemic value at the University of Stirling, Scotland, in August 2006. Feedback received on these occasions was also very helpful. Finally, work on an earlier version of this paper was supported by a 2004 Summer Research Grant from Loyola Marymount University.

- Riggs, W. (2002). 'Reliability and the Value of Knowledge,' *Philosophy and Phenomenological Research*, 64: 79–96.
- Sosa, E. (2003). 'The Place of Truth in Epistemology,' in DePaul and L. Zagzebski 2003.
- Zagzebski, L. (1996). *Virtues of the Mind*. Cambridge: Cambridge University Press.
- (2000). 'From Reliabilism to Virtue Epistemology,' in G. Axtell (ed.), *Knowledge, Belief, and Character*. Lanham, MD: Rowman and Littlefield.
- (2003). 'Intellectual Motivation and the Good of Truth,' in DePaul and Zagzebski 2003.

3

Testimony and the Value of Knowledge

Martin Kusch

Knowledge is in the end based on acknowledgement.

(Ludwig Wittgenstein)

At a certain point philosophy needs to make way to history, or, as I prefer to say, to involve itself in it.

(Bernard Williams)

1. INTRODUCTION

In this paper I seek to defend a *communitarian* form of *value-driven* epistemology. ‘Value-driven epistemology’ studies the epistemic value of various cognitive achievements, such as knowledge, understanding, or wisdom.¹ The favourite question of value-driven epistemologists is whether knowledge is epistemically more valuable than other cognitive states; for instance, whether knowledge is epistemically more valuable than true belief (*simpliciter*), reliably produced true belief, or justified true belief. A form of epistemology qualifies as ‘communitarian’ if its investigations are guided by the assumption that human cognizers are ‘highly gregarious and deeply interdependent’ creatures.² Putting the two elements together, a communitarian form of value-driven epistemology seeks to understand the values of various cognitive states in relation to the needs and actions of human beings in social interaction with one another.

In the present context I shall not try to offer anything like a ‘comprehensive’ communitarian value-driven epistemology. My main focus will be on the idea

¹ The expression ‘value-driven epistemology’ was coined by Riggs (2007). Pritchard (2007) and Riggs (2007) provide useful reviews of the work to date. Wisdom has received least attention. For a useful discussion, see Whitcomb (2006).

² I have defended a version of (and the term) ‘communitarian epistemology’ in Kusch (2002). I take the formula ‘highly gregarious and deeply interdependent’ from Barnes (2000: p. ix) who in turn borrows it in the form ‘highly gregarious and deeply interdependent primates’ from Gagnier and Dupré (1998).

that knowledge is valuable as a collective good. In giving content to this idea I shall draw on two traditions that to date have played little or no role in value-driven epistemology: the ‘genealogy’ of epistemological concepts and virtues (Craig 1990; Williams 2002), and the ‘sociology of scientific knowledge’ (especially Barnes 1995; 2000; and Shapin 1994). I shall use these traditions in order to construct an alternative to Jonathan Kvanvig’s highly—and deservedly—influential version of value-driven epistemology (Kvanvig 2003). In my view, Kvanvig’s work on the value of knowledge is too static, too abstract, not sufficiently historical, and not suitably attentive to the social nature of knowledge. Put differently, in his recent work on the value of knowledge, Kvanvig seems to have forgotten the most important insight of his earlier study of intellectual virtue (Kvanvig 1992). In 1992 Kvanvig wrote that ‘divorcing epistemological concerns from the realities of social interaction generates an epistemology built on answers to questions as relevant to the life of the mind as “desert island” cases are in ethics’ (1992: 178). And he insisted that epistemologists should focus on ‘bunches of people rather than isolated individuals, bodies of knowledge rather than individuated propositions, and experienced, processual chunks of time rather than abstracted, individual time-slices’ (1992: 186). I wholeheartedly applaud these views. And I am ready to defend them even against (the more recent) Kvanvig himself.

Lest some readers will be disappointed by the sketchy nature of what follows, let me conclude my introductory comments with a warning. This study is more about hammering out a rough position than about fine conceptual distinctions or detailed defences against all conceivable objections. As Kant once inimitably put it in defence of the intricacy and complexity of his epistemological writings: ‘Hammer and chisel are perfectly fine for working raw lumber, but for copperplate one must use an etching needle’ (Kant 2004: 9). As far as communitarian value-driven epistemology is concerned, we are still in the era of raw lumber; the time of copperplate and needlework is still to come. For the moment, at least in this area, we have to learn ‘how to philosophise with the hammer’ (Nietzsche 1990).

2. KVANVIG’S VALUE-DRIVEN EPISTEMOLOGY

Kvanvig’s recent book *The Value of Knowledge and the Pursuit of Understanding* (2003) has put the issue of epistemic value on a new and high level of content and clarity. Kvanvig’s starting point is the insistence that a satisfactory account of knowledge must explain both its *nature* and its *value* (2003: p. x). We have captured the *nature* of knowledge if, and only if, with respect to the *concept* of knowledge we have succeeded in constructing an explicit and correct *intension* for our intuitive *extension*.³ We have identified the *value* of knowledge if, and

³ This way of putting the task comes from Craig (1990: 1).

only if, the reconstructed intension can explain why we are prone to think that knowledge itself—or some of its components, or some concepts in the vicinity of knowledge—is epistemically valuable.

Kvanvig defends a form of *value-scepticism* about knowledge; he maintains that there is no *knowledge-specific* epistemic value. There is no distinctive epistemic value that all and only the elements of the extension of *knowledge* share. (From here on, I shall write *concepts* in italics, ‘words’ in inverted commas, and leave elements of the extension unmarked.) Kvanvig argues by elimination: he shows that no existing account of the intension of *knowledge* can explain why we find all and only elements of the intuitive extension of *knowledge* epistemically valuable in a distinctive and unique way. Here I shall briefly mention just three proposals and why they fail; my aim is to give (the novice to value-driven epistemology) a flavour, rather than a detailed summary, of Kvanvig’s reasoning.

(a) *Knowledge is true belief plus a property that renders the truth of the belief objectively likely*

This formula captures a number of positions, amongst them reliabilism. The reliabilist thinks that knowledge is more valuable than true belief since a true belief produced by a reliable truth-tracking method is epistemically better than a true belief produced in some other, arbitrary, way. After all, a reliable truth-tracking method makes it objectively highly likely that its products will be true. Kvanvig throws out this thought on the basis of the ‘swamping argument’. The epistemic value of a *true belief that is objectively likely to be true* can never be greater than the epistemic value of a *true belief simpliciter*. The epistemic value of a true belief ‘swamps’ the value of using a method that makes the belief likely to be true (Kvanvig 2003: 47–9). The case is analogous to the following. Assume that the taste of the cup of green tea I am currently drinking is perfect; in no possible world does it taste any better than it tastes here. As it happens, this cup was prepared by a total novice in tea-making (i.e. me) who achieved the perfect taste largely by accident. However, that this perfect taste was produced by accident does not diminish it. Perfect taste is perfect taste, whatever its causes. Even if, counterfactually, this perfect-tasting cup of tea had been brewed by a Japanese ‘Grand Tea Master’, the pleasure of the taste would not be greater. The fact that the Grand Tea Master is objectively more likely (than me) to bring about the perfect taste has no bearing on the taste of this very cup. The perfect taste of this cup swamps whatever value there might be in having one’s tea prepared by a Grand Tea Master. The swamping argument works not only against reliabilism but also against any other third condition that is meant to be *instrumental* in getting us to the truth.

(b) *Knowledge is subjectively justified true belief*

Kvanvig accepts that—at least under a certain reading of ‘subjective justification’—subjectively justified true beliefs are epistemically valuable. On the

‘certain reading’ in question, subjective justification is not merely of *instrumental* value for reaching the truth. To be instrumentally valuable is *just one way* of being *extrinsically* valuable; where *A* is extrinsically valuable with respect to *B* if, and only if, the value of *A* is explicable in terms of the value of *B*. In the present case, the extrinsic but non-instrumental value of subjective justification has to do with the fact that, as mental states, subjective justifications are *reflectively accessible*: we usually can tell by means of reflection alone whether we have a justification or not. In this respect justification differs from truth. Merely reflecting on a belief does not (usually) help one decide whether or not the belief is true (2003: 66). It is reflective accessibility that gives subjective justifications an extrinsic but non-instrumental value. The epistemic value of reflectively accessible subjective justifications must be explicated in terms of the value of truth—after all, justifications are supposed to help with reaching the truth. But subjective justifications bear a feature that truth lacks: reflective accessibility. And hence the epistemic value of subjective justification is not swamped by the epistemic value of truth. Unfortunately, all this does not mean that we have found a distinctive epistemic value for knowledge. The two concepts *subjectively justified true belief* and *knowledge* are not co-extensive. It seems that to get from subjectively justified true belief to knowledge we need to add a fourth, Gettier-case-blocking, component.

(c) Knowledge is subjectively justified true belief plus a fourth condition blocking Gettier cases

This suggestion picks up where the previous left off. Kvanvig now argues that no fourth condition is able to close the gap between knowledge and justified true belief. The key difficulty is that all existing proposals for a fourth condition fail the value-test: none of them excludes a form of accidentalness that we find ‘intuitively disvaluable’. The point can also be put in terms of a dilemma: trying to solve the problem of the nature of knowledge, we are led, as far as the fourth condition is concerned, to ever more ‘complex, ad hoc and gerrymandered proposals’. And yet, attempting to make progress with the value of knowledge, we are forced to look towards ‘simpler proposals in which the value of the condition is intuitively obvious’. Kvanvig’s conclusion: ‘The Gettier problem shows that no component-based account of the value of knowledge will be successful’ (2003: 117).

Having established his value-scepticism regarding knowledge, Kvanvig offers a more upbeat assessment concerning the distinctive epistemic value of ‘factive understanding’. We take understanding factively when we accept that to understand that *p* implies *p* (e.g. ‘John understands that he is still a child’ implies ‘It is a fact that John is still a child’). On Kvanvig’s analysis, factive understanding is similar to knowledge in one respect, but radically different from knowledge in another respect. Factive understanding is similar to knowledge if knowledge

is rendered according to internalist coherence theories of justification. In other words, understanding ‘requires the grasping of explanatory and other coherence-making relationships in a large and comprehensive body of information’ (2003: 192). The radical dissimilarity is that, unlike knowledge, understanding does not require a condition that excludes epistemic luck. Understanding is distinctly epistemically valuable because: first, true beliefs are valuable; second, grasped coherence relations are relations of subjective justification, and, as such, reflectively accessible; third, grasped coherence relations enable both the finding and organizing of new truths and the development of our thinking about a given subject matter; fourth, having an organized system of beliefs is pragmatically useful for action; and fifth, elements of thought that are organized in this way allow for the ‘intrinsically satisfying closure of the process of inquiry’ (2003: 202).

3. CRAIG’S GENEALOGY OF KNOWLEDGE AND EPISTEMIC VALUE

It is striking and surprising to note that among the proposals considered by Kvanvig only one is part of the recent ‘social turn’ in epistemology.⁴ The sole representative of this turn is a social-pragmatic version of the theory of epistemic virtue. Kvanvig (2003: 83) ascribes this view to an unpublished paper by a past self of John Greco. According to the young Greco, we value epistemically virtuous believers on the prudential grounds that they make for good informants. Kvanvig repudiates this approach on several grounds. His most important point is that true believers are better informants than intellectually virtuous believers. Given the brevity of Kvanvig’s discussion, it is impossible to make out what considerations led Greco to his conception, or whether Kvanvig’s criticism of Greco is fair. Fortunately, in two short footnotes Kvanvig (2003: 83) registers his impression that Greco’s stance is a ‘close cousin’ of the theory advanced in Edward Craig’s book *Knowledge and the State of Nature: An Essay in Conceptual Synthesis* (1990). I therefore turn to a brief reconstruction of Craig’s main line of thought before returning to Kvanvig’s critical assessment.⁵

Craig’s book introduces a new method for understanding key concepts in epistemology. I shall follow Bernard Williams (2002: 32)—about whom more below—and call this method ‘genealogy’. Craig himself speaks of ‘conceptual synthesis’ in the title of his book. Craig’s synthesis contrasts with traditional philosophical analysis. Take the concept *knowledge* (and its cognates). By an

⁴ e.g. Code (1995), Elgin (1996), Fricker and Hornsby (2000), Goldman (1999), Kusch (2002), Schmitt (1994).

⁵ Craig’s work is discussed and developed in Fricker (1998), Hanfling (2000), Jones (1997), Lane (1999), Neta (2006), Schmitt (1992), and Weinberg (2006). The position explored in Reynolds (2002) is similar to Craig’s. I have profited most from studying Fricker, Lane, and Neta.

'analysis' of *knowledge* Craig means the enterprise of trying to find components of the concept that are individually necessary and jointly sufficient. Work of conceptual analysis proceeds by trying to provide an explicit or fully articulated intension for an intuitive extension. It is not part of conceptual analysis to ask why delimiting *this* rather than some other extension is of value to the concept-users. Conceptual synthesis inverts the route taken by analysis: it seeks to illuminate *knowledge* by providing a hypothetical historical narrative of the process in which this concept was put together. This narrative is constrained by two ideas. The first idea is that of an epistemic 'state of nature', that is, of an imaginary, early, social community composed of language-using human beings who are co-operative though not kin, and whose conceptual and reflective powers are somewhat weaker than our own. The genealogical narrative must make intelligible why these creatures found it useful or valuable to introduce an ancestor of our concept of *knowledge*. This aspect of genealogy situates it in the proximity of value-driven epistemology. Craig himself writes that he is 'asking after the roots of the value of knowledge' (1990: 7). The second constraint is social-developmental: the genealogical narrative must explain—invoking social change as the central cause—why the ancestor concept was eventually replaced by our concept. For ease of reference, I shall refer to the ancestor concept and word (and their cognates) as *protoknowledge*, 'protoknowledge', *protoknower*, etc.

The first part of Craig's genealogical story—why is *protoknowledge* found useful in the state of nature—goes as follows. In the state of nature, individuals depend upon one another for information. Distinguish between the roles of 'inquirer' and 'informant'. The inquirer needs information that she is currently unable to directly obtain herself; the informant offers such information. Inquirers must be able to separate good from bad informants. And it is natural to assume that meeting this need will involve concepts. Assume that *protoknower* is the central conceptual tool for dealing with this problem. Which conceptual components should *protoknower* contain? What should we hypothesize our imaginary ancestors to want this concept for? Craig's answers are that our ancestors want this concept as a tag for good informants and that the concept *protoknower* (whether *p*) comprises the following elements: (i) being as likely to be right about *p* as the inquirer's current needs require; (ii) being honest; (iii) being able to make the inquirer believe that *p*;⁶ (iv) being accessible to the informant here and now; (v) being understandable to the inquirer; and (vi) being detectable as a good informant concerning *p* by the inquirer. To elaborate briefly only on the last point, the inquirer needs to find properties (= *X*), 'indicator-properties', that she is able to detect and that correlate closely and in a lawlike fashion with holding a true belief, or telling the truth, as to whether *p* (1990: 25, 135). 'Being at the top of a tree' might be such a property

⁶ Craig allows that the informant can make the inquirer believe that *p* without that the informant herself believes that *p* (Craig 1990: 14–15).

X for some inquirers in the state of nature when p is the proposition that a tiger is approaching the village. Usually more than one X will be involved. The properties that make Fred a medical protoknower are not one but many. Craig is adamant that (i) to (vi) are not necessary and sufficient conditions. While all of these elements are present in 'prototypical' situations, the concept has a use even when some of the elements are missing. Finally, *protoknowledge* differs from *knowledge* in that: (a) only the former is closely tied to testimony; (b) *protoknowledge* is not a fully public concept insofar as it is indexed to the capacities and needs of specific inquirers (1990: 90); (c) *protoknowledge* can be ascribed only to others but not to oneself; and (d) *protoknowledge* is not undermined by accident or luck: users of *protoknowledge* lack the intellectual sophistication to distinguish between accidental and non-accidental fulfilment of the conditions of *protoknowledge*.

This brings us to the second half of the genealogical just-so story: the hypothetical social-historical narrative that takes us from *protoknowledge* to *knowledge*. Craig speaks of this development as a process of 'objectivization' of *protoknowledge*. Key steps in this objectivization are the following. First, *protoknowledge* comes to be used in self-ascription. In response to the question 'who knows whether p ?' group members start to investigate their own indicator-properties. Second, inquirers begin to recommend informants to others. This can be done in a helpful manner only if the perspectival or indexical character of *protoknowledge* is weakened. The recommended informant must be good in the eyes of both the recommender and the recipient of the recommendation. Further movement in this dimension—recommending an informant to ever more inquirers—makes *protoknowledge* increasingly hard to get. The endpoint is the idea of 'someone who is a good informant as to whether p whatever the particular circumstances of the inquirer . . . That means someone with a very high degree of reliability, someone who is very likely to be right—for he must be acceptable even to a very demanding inquirer' (1990: 91). And a very demanding inquirer will not accept epistemic luck or accident. Third, inquirers begin to use 'being recommended' as an indicator property. This move dilutes the original detectability requirement. Inquirers begin calling someone a 'protoknower' even when none of the original 'natural' indicator-properties is in sight. Fourth, in the context of group action inquirers cease to care whether the needed information is accessible to them as individuals; they are satisfied if it is accessible to someone in the group. As a result they will speak of *protoknowledge* even outside the context of testimony. The process of objectivization ends up with our concept of *knowledge*: 'The concept of knowing . . . lies at the objectivized end of the process; we can explain why there is such an end, and why it should be found worth marking in language' (Craig 1990: 90–1).

To sum up, Craig's overall message is that our concept of knowledge has a history and that if we want to understand its current contours, we must make a detour through an imaginary genealogy. Under the constraints of some natural

assumptions about the kind of state of nature in which our ancestors must have found themselves, a concept of *protoknower* as a detectable good informant is likely to have arisen. *Protoknowledge* had a distinctive epistemic value: the value of having true beliefs coming from such a detectable good informant. Subsequently, the concept of *protoknowledge* underwent a conceptual development resulting in our concept of *knowledge*. Craig does not say whether our concept still delimits something of a distinctive value.

4. OBJECTIONS AND MODIFICATIONS

I have already mentioned that Kvanvig is critical of the young Greco's and Craig's 'pragmatic accounts' of the value of knowledge (2003: 83–6). Time has come to expand on his concerns. I shall distinguish between four 'Kvanvagian' objections. The first two come directly from Kvanvig's *The Value of Knowledge and the Pursuit of Understanding* (2003). The last two cannot be found explicitly in Kvanvig's texts, but seem to me to be very much in his spirit. Since the four objections provide the structure for the remainder of this paper, I shall be a little pedantic in setting them out.

4.1. The *Swamping Objection* (Kvanvig 2003: 86)

Craig's account of epistemic value succumbs to a version of the swamping argument against reliabilism. He seeks to identify a protoknowledge-specific epistemic value. In other words, Craig holds that protoknowledge is epistemically more valuable than mere true belief. Consider the following two situations:

Situation A Inquirer Fred has just formed the *true belief* that *p*. He has done so on the basis of testimony offered by John. Fred had previously *tagged* John as a detectably good informant on whether or not *p*.

Situation B Inquirer Fred has just formed the *true belief* that *p*. He has done so on the basis of testimony offered by Otto. Fred did *not* previously *tag* Otto as a detectably good informant on whether or not *p*.

On Craig's account, Fred's cognitive state concerning *p* in *Situation A* is of greater epistemic value than his cognitive state concerning *p* in *Situation B*. This is because in *A*, but not in *B*, Fred got the information that *p* from a protoknower. But what does 'information coming from a protoknower' mean? On Kvanvig's reading it means 'information coming from someone who is objectively likely to be right about *p*'. And now it should be clear how the swamping argument kicks in. That John is an approved source of information (for Fred) regarding *p* might make it objectively more likely that Fred's belief that

p is true. However, if Fred's belief is already true, then the objective likelihood of it being true adds nothing of epistemic value. It is difficult to see how Craig can avoid this difficulty. After all, Craig wants to tell us why protoknowledge is valuable. And in developing an answer, Craig explicitly adopts the strategy of explaining the value of protoknowledge in terms of a surplus of epistemic value, a surplus above and beyond the epistemic value of true belief (1990: 7). Craig's genealogical narrative then goes on to locate the extra value in those qualities of a given informant that make it objectively more likely that the given inquirer will get the accurate information she needs. The fact that Craig's analysis is in terms of protoknowledge rather than knowledge does not lessen or divert the difficulty.

Kvanvig's second objection is close kin to the first:

4.2. *The Conflation Objection* (Kvanvig 2003: 85)

Craig's account is based on a conflation of our needs in the 'context of discovery' with our needs in the 'context of the final product'. Distinguish between two contexts. In the 'context of discovery' we need to make out who are good informants on whether or not p . This is because we do not yet know ourselves whether or not p . And thus in the context of discovery we require the 'indicator-properties' discussed by Craig. In this context the indicator-properties are of enormous value. In the 'context of the final product' we already have a true belief concerning p . And since we do, we no longer have any need for indicator-properties. They are now of no value to us. Craig thinks that indicator-properties add value to a true belief. Ergo, he must be conflating the two contexts.

4.3. *The Social Objection*

Craig's account is insufficiently social in that: (a) it ignores the social consequences of (proto-)knowledge attributions; (b) it disregards the fact that the institution of testifying is a collective good; and (c) it neglects the reciprocal conceptual needs of the informant.

It is with tongue in cheek that I call the *Social Objection* 'inspired by Kvanvig's work'. And yet, I am inclined to think that the Kvanvig of 1992—the wonderful time-slice of a man that so vehemently insisted on a social understanding of knowledge—could well have put this objection. The idea behind it is that Craig's genealogy is oddly one-sided. It centres on the conceptual needs of the inquirer but has nothing to say about the motivations and conceptual desiderata of the informant, or about the social effects of classifying people as protoknowers. How can Craig be so sure that these aspects have not had a decisive influence on the content of the ancestors of our concept of knowledge? And why is it implausible to assume that these aspects have 'left a mark' on our concept?

4.4. The Relevance Objection

Craig does not offer a convincing argument for why we should favour his particular genealogy over others. It is not obvious that *knowledge* bears marks of the good informant. Indeed, it is quite unclear why history would solve any problems of value-driven epistemology.

Craig never considers alternative genealogies and in support of his own he merely suggests that our concept of knowledge 'still bear[s] certain marks of its origin' or that 'the concept of the informant' has left 'its mark, so to speak, on the concept of knowledge' (1990: 95, 97). Alas, Craig fails to identify these marks for us. However, I imagine the *Relevance Objection* to lament this point only in passing. Let us assume that Craig is right, and there really once was a community operating with the concept *protoknowledge*, a concept that delimited items united by a distinctive epistemic value. What does that tell us about the epistemic value of knowledge? The answer surely is 'little, or nothing'. In order to make progress regarding the epistemic value of our concept we have to do what Kvanvig does in *The Value of Knowledge and the Pursuit of Understanding*; we have to investigate our intuitions regarding our concept. Borrowing a line from the early (unhistorical) Wittgenstein, we might ask: What is history to do with us? Ours is the first and only concept of knowledge.⁷

I am not suggesting that these four objections are the only possible ones—I shall mention some further worries in my Conclusion. But one cannot do everything at once, and thus here I shall be satisfied if I manage to weaken the appeal of the listed four. To give a brief preview of what is to come, I shall try to answer these objections in good part (though not exclusively) by modifying Craig's genealogy. Williams, Barnes, and Shapin will be crucial resources in this project. The three key modifications can already be listed here.

Modification One (inspired by Williams) Giving good information to others is a social institution and a collective good. It can exist only if it is underpinned by a network of intrinsic values.

Modification Two (inspired by Barnes and Williams) Actions of honouring and shaming are central to maintaining the institution of testifying (qua collective good). Protoknowledge attributions are attributions of honour.

Modification Three (inspired by Barnes and Shapin) Attributions of protoknowledge are intertwined with attributions of freedom: (a) Attributions of honour are attributions of freedom from (certain forms of) interference. (b) Attributions of protoknowledge are attributions of discretion over a practice. (c) Informants have the conceptual need to mark out inquirers who can be (rationally) influenced by good information. The concept of *protofreedom* plays this role.

⁷ 'What has history to do with me? Mine is the first and only world!' (Wittgenstein 1961: 82).

Modification One will be important for answering the *Swamping* and the *Conflation Objections*. *Modifications Two* and *Three* will first help defuse the *Social Objection* and then be instrumental in weakening the force of the *Relevance Objection*.

5. EXCURSUS: IN PRAISE OF CRAIG

In what follows I shall increasingly deviate from Craig's original position. Nevertheless I think of this paper as *developing* rather than as *replacing* Craig's genealogy. *Knowledge and the State of Nature* has given us a new paradigm or exemplar for doing epistemology and the fact that the book has picked up few citations over the past fifteen years should not blind us as to its value and originality. One of the most valuable aspects of genealogy is its systematic use of the idea that the evolution of concepts and the development of social relations are inseparable. Every step in the evolution from *protoknowledge* to *knowledge* is explicated in terms of changed needs of the group or changed forms of interaction. For instance, Craig explains that the detectability requirement becomes diluted because people begin to recommend informants to others or because they start to engage in certain new forms of group action. The idea that conceptual and social relationships are inseparable, or 'internally related', is of course not new—its most forceful articulation can be found in Peter Winch's Wittgenstein-inspired classic *The Idea of a Social Science and its Relation to Philosophy* (1958).⁸ But Craig's study is the only investigation in epistemology I know of that puts this idea to good systematic use.

There are other respects too in which it is natural to think of Craig's genealogy as influenced by Wittgenstein. (I take such influence to constitute a merit.) Some interpreters of Wittgenstein have suggested that his approach to meaning and intentional content involves three key ideas (Kripke 1982; Kusch 2006). The first is to focus on utterances in which meaning and content are attributed to others. The second recommendation counsels against trying to capture the significance of meaning-or-content attributions in terms of necessary and sufficient truth conditions. Truth-conditional analysis is to be replaced with a study of rough and ready assertability—or appropriateness—conditions.

⁸ [Consider] 'the . . . first introduction of the concept of a germ into the language of medicine. This was . . . a . . . radically new departure, involving not merely a new factual discovery within an existing way of looking at things, but a completely new way of looking at the whole problem of the causation of diseases, the adoption of new diagnostic techniques, the asking of new kinds of question about illnesses, and so on. In short it involved the adoption of new ways of doing things by people involved, in one way or another, in medical practice. An account of the way in which social relations in the medical profession had been influenced by this new concept would include an account of what that concept was. Conversely, the concept itself is unintelligible apart from its relation to medical practice. . . . a new way of talking sufficiently important to rank as a new idea implies a new set of social relationships' (Winch 1958: 121–3).

Finally, the third piece of advice urges us to ask: 'What is the role, and the utility, in our lives of our practice of asserting (or denying) the form of words [e.g. of meaning attributions] under these conditions?' (Kripke 1982: 73). Or elsewhere: 'granted that our language game permits a certain "move" (assertion) under certain specifiable conditions, what is the role in our lives of such permission? Such a role must exist if this aspect of the language game is not to be idle' (Kripke 1982: 75).

There is of course much in Craig that resembles this Wittgensteinian methodology. Craig rejects an analysis of *knowledge* in terms of necessary and sufficient conditions; he focuses primarily on the attribution of knowledge to others; he does not attempt to identify truth conditions for knowledge attributions; and he emphasizes the importance of identifying, or hypothesizing, the role, or the utility, of the concept of knowledge for us—or of *protoknowledge* for our ancestors.

Naturally, there is also an interesting parallel between Craig's genealogy and Wittgenstein's occasional use of simple or simplified language games. Recall for instance that the *Philosophical Investigations* (1953) begin with a series of language games involving two builders. Later elements of the series are more complex than earlier ones both as far as linguistic resources and as far as forms of interaction are concerned. And yet Wittgenstein calls each stage of the series a 'language', or a 'language game', adding the qualification 'primitive' for early stages (1953: § 7). As these passages suggest, Wittgenstein is entirely comfortable with imaginary scenarios, even when these involve more primitive and invented scenarios. Wittgenstein speaks of these imaginary scenarios as functioning as '*objects of comparison* which are meant to throw light on the facts of our language by way not only of similarities, but also of dissimilarities' (1953: § 130). And he insists that 'we are not engaged in natural science, and not even in natural history—since we can also surely provide fictitious natural history of our purpose' (1953: II. xii).

Finally, Craig applies Wittgenstein's idea of family resemblance concepts to *knowledge*.⁹ As already mentioned, Craig rejects the notion that the meaning of knowledge can be captured in terms of necessary and sufficient conditions. To this we can now add that Craig does not seek to assimilate all uses of 'knowledge' to his 'good informant' account. In discussing expressions like 'knowing how',

⁹ I am referring of course to the idea of family resemblance concepts, famously introduced by Wittgenstein in §§ 66–7 of the *Philosophical Investigations* (1953):

66. Consider for example the proceedings that we call 'games'. I mean board-games, card-games, ball-games, Olympic games, and so on. What is common to them all? . . . if you look at them you will not see something that is common to all, but similarities, relationships, and a whole series of them at that. . . .

67. I can think of no better expression to characterize these similarities than 'family resemblances'; for the various resemblances between members of a family: build, features, colour of eyes, gait, temperament, etc. etc. overlap and criss-cross in the same way.

'knowing people', or 'knowing places' he shows himself open to the possibility of an non-reducible variety of uses (1990: 140–61). Perhaps one should even insist that the strength of a genealogical analysis of a given term can be measured by its ability to make sense of a possibly irreducible variety of uses.¹⁰ If the same term can have different functions, then this variety calls for an explanation. And often the best explanation will be historical—that is, genealogical. There will be reasons why it was natural for language users to have the same term cover the different functions. At the same time, it cannot be expected that the elements of the extension(s) of a family resemblance concept like knowledge all come with the same epistemic value.

Especially as far as the last-mentioned issue is concerned, I find Craig's methodology superior to the standard approach by epistemologists, an approach that is well represented by Kvanvig (2003). Kvanvig is eager to distinguish between those uses of 'knowledge' that are 'central to epistemological inquiry, related as it is to theoretical concerns' and those uses that are not. The borderline he has in mind coincides with the divide between factive and non-factive uses of 'knowledge'. As Kvanvig has it, utterances of the form 'it used to be known that the earth is flat' are instances of 'misspeaking' typical of 'undergraduates' at the beginning of their philosophical education (2003: pp. xi, 190); expressions of the form 'I just *knew* I was going to fail' (when the speaker in fact succeeded) conflate knowledge and psychological certainty; and locutions involving the 'current state of scientific knowledge', or 'talk about the present state of knowledge (even though we know that some of what falls under that rubric is false)' are merely 'honorific' (Kvanvig 2003: 201; Kvanvig 2009: 342). Kvanvig goes so far as to say that 'if I were to offer a theory of knowledge, I would not expect it to answer to locutions involving the current state of scientific knowledge'. Kvanvig takes such honorific uses to belong not to the semantics but to the 'pragmatic dimension of epistemic terminology' (2009: 342).

I shall return to the issue of 'honorific uses' below. Here I am more interested in Kvanvig's excluding of 'misspeaking', that is, his excluding of undergraduates' talk of past and rejected knowledge. This case is interesting for us since Kvanvig is wrong in confining such uses of the term 'knowledge' for past institutionalized

¹⁰ What then can a genealogical analysis of knowledge tell us about the non-factive uses? Consider only the case of non-factive use found among social scientists. Returning to Craig's state of nature, and sticking to the demand that conceptual evolution must be in step with developing social relations, it seems natural to suggest that the non-factive use is the result of a further form of objectivization. Craig's objectivization of protoknowledge takes us from a concept that is relative to a single individual to a concept that is public; the standards for the good informant have been raised so that by using 'protoknowledge' one can recommend an informant to anyone in one's own community. The obvious next step is to consider a meeting of different communities of different degree of development and with diverging taxonomies and informational needs. Reflective observers of one tribe, say tribe A, might well notice that the other tribe, say B, has its own division of cognitive labour, its own testimonial system, and thus its own label for approved sources of information. These reflective members of A might then come to use their own label 'protoknower' for someone who—in tribe B, by members of B—is regarded as a detectable good informant.

or past taken-for-granted beliefs to ignorant undergraduates. Historians and sociologists of knowledge employ the term very much in the same way—and have done so at least since the beginning of the twentieth century (e.g. Durkheim 1983; Bloor 1991; Shapin 1994). To accuse generations of social scientists of misspeaking on the grounds that their use is not that favoured by mainstream epistemology strikes me as a misplaced attempt at linguistic legislating. Equally importantly, to dismiss these uses of ‘knowledge’ as simple mistakes is to block one’s view to what sociologists might be able to contribute to one’s own enterprise.

Lest this disagreement seems merely like a clash of intuitions, note that Kvanvig’s two criteria for excluding non-factive uses of ‘knowledge’ are not compelling. According to the first criterion, non-factive uses are ‘[not] central to epistemological inquiry, related as it is to theoretical concerns’. This is unconvincing. Clearly the non-factive use of ‘knowledge’ favoured by social scientists is very much ‘related to theoretical concerns’. Social scientists aim to develop theories that explain why certain types of belief—or indeed, certain types of *knowledge*—are found credible in certain types of communities. So Kvanvig had better have a different reading of his criterion. The only other reading I can think of is to interpret ‘theoretical concerns’ as ‘concerns with the truth of the beliefs’. Alas, this reading of ‘theoretical concerns’ renders the whole argument circular; it boils down to saying that non-factive uses should be ignored because they are not factive.

Kvanvig’s second attempt at linguistic purification distinguishes between the semantics and pragmatics of ‘knowledge’. Factive uses belong in the former, non-factive uses in the latter. This dividing line is not helpful either. Debating it properly would take us back to the debate whether G. E. Moore’s ‘I know I have a hand’ is a meaningful statement. Moore (1993) and later Paul Grice (1989), John Searle (1969), and Keith DeRose (1997) famously argue that the statement has meaning and thus is true or false; this is so, they maintain, since ‘know’ (and the other words occurring in the utterance) have a context-independent determinate meaning. Norman Malcolm (1949; 1977), Wittgenstein (1969) and, more recently, Charles Travis (1991), have supported the opposite view; to wit, that there is no sensible distinction between the semantics and pragmatics of knowledge attributions, and that hence, without a specified context, Moore’s utterance is meaningless. If the latter view is right, then Kvanvig’s second go at conceptual cleansing fails as well.

6. MODIFICATION ONE: WILLIAMS ON TESTIMONY, FREE-RIDING, AND INTRINSIC VALUES

Williams’s book *Truth and Truthfulness: An Essay in Genealogy* (2002) adjusts Craig’s genealogical method in one crucial respect.¹¹ According to Williams,

¹¹ It is astonishing to note that epistemologists have paid so little attention to this book. The exception that proves the rule is Elgin (2005). I have greatly profited from her discussion.

Craig's 'imaginary genealogy' cannot be the whole story; it needs to be complemented by 'real genealogy', that is, by an engagement with historical and cultural contingent realities. Williams insists that in moving from imaginary to real genealogy we do not leave philosophy behind: 'philosophy cannot be too pure if it really wants to do what it sets out to do' (2002: 39). At the same time, Williams defends the role of *imaginary* genealogy. Imaginary genealogy is crucial for making sense of the possibility of the development of 'cultural phenomena' in those cases where the historical record contains no helpful data—we cannot identify a past or present community that did not 'always already' have these cultural phenomena (2002: 35).

Although Williams does not put it in these very terms, it is inviting to read his book as offering an imaginary and real genealogy of the cultural phenomenon of testimony. Thus Williams shares with Craig both the subject matter and (much of) the method. But whereas Craig focuses on *concepts*, Williams is more interested in the details of the *values*, *motivations*, and *virtues* that underpin the phenomenon. According to Williams, the central virtues of testifiers are 'accuracy' and 'sincerity'. Accuracy refers to an individual's disposition to seek out the truth and to report it in a way that is sensitive to the recipient's circumstances, needs, and abilities. Unlike Craig, Williams does not take it for granted that individuals in the state of nature are co-operative and eager to offer information to others. Williams recognizes that the *institution of testimony*—as I shall subsequently call it—is a collective good. Individuals who are rational in a purely self-interested way will try to 'free-ride': they will seek to obtain accurate and sincere testimony from others without offering anything in return. After all, collecting useful information usually involves costly 'investigative investments' (2002: 88).

How is the problem of collective action solved? The core of Williams's solution to the Free-Rider problem concerning the institution of testimony is the suggestion that accuracy and sincerity (and with them the institution itself) must come to be regarded by community members as *shared intrinsic*—rather than as merely *instrumental*—*values* (2002: 90). For community members to have trust in each others' reports, they must be convinced that accuracy and sincerity are non-negotiable. And this implies that these values are not—except under extreme circumstances—weighed against, and possibly outweighed by, other interests and values. As Williams sees it, to say that accuracy and sincerity cannot be weighed against other interests and values is to say that accuracy and sincerity cannot be rendered instrumentally. The uses of instruments are routinely assessed as to their costs and benefits, and arguments in favour of their use can always be outweighed by other considerations. Williams also emphasizes that adherence to accuracy and sincerity must be 'stable under reflection' (2002: 91). It is not enough that everyone merely 'pretends' to have adopted accuracy and sincerity as intrinsic values. Pretence too can be outweighed. And finally intrinsic values must make sense to the agent 'from the inside, so to speak: it must be possible

for them to relate trustworthiness to other things that they value, and to their ethical emotions' (2002: 91–2).

The expression 'relate trustworthiness to other things that they value' points us to the fact that Williams's solution to the Free Rider problem involves an interpretation of intrinsic value that differs fundamentally from the well-known conception proposed in G. E. Moore's *Principia Ethica* (1993). Moore famously believed that we identify intrinsic value by 'the method of absolute isolation': 'it is necessary to consider what things are such that, if they existed *by themselves*, in absolute isolation, we should yet judge their existence to be good' (1993: 236–7). Against Moore, Williams holds that values come in socially shared webs and systems, that intrinsic values occupy a central position in such webs or systems, and that values make sense only in their essential relations to one another. Given the significance of this viewpoint for much of what follows, it is worth quoting the central passage at some length:

What we want . . . is some insight into these values, some account of their relations to other things that we know that we need and value, but an insight which does not reduce them to the merely instrumental. What we want . . . is explanation without reduction (2002: 90)

it is in fact a sufficient condition for something (for instance, trustworthiness) to have an intrinsic value that, first, it is necessary (or nearly necessary) for basic human purposes and needs that human beings should treat it as an intrinsic good; and, second, they can coherently treat it as an intrinsic good. This means that it is stable under reflection . . . What is essential for this to be so is that the agent has some materials in terms of which he can understand this value in relation to other values that he holds, and this implies, in turn, that the intrinsic good (in this case trustworthiness), or rather the agent's relation to it, has an inner structure in terms of which it can be related to other goods. (2002: 92)

Williams's view of intrinsic values relates directly to his distinction between imaginary and real genealogies. While no community can exist without values like accuracy and sincerity, different communities embed such values in different 'wider range[s] of values'. And what these wider ranges of values 'varies from time to time and culture to culture' (2002: 93). The bulk of Williams's study is an investigation into the links between trustworthiness and other values in the Western tradition. I shall take up one such link below: the link between trustworthiness and freedom.

7. MODIFICATION ONE AS A REPLY TO THE SWAMPING OBJECTION

Modification One adds to Craig's genealogy the ideas that the institution of testifying is a collective good and that this institution and the virtues underpinning

it are of intrinsic value to group members. I shall now show that *Modification One* enables us to answer Kvanvig's *Swamping* and *Conflation Objections*. As Kvanvig would say, Craig's work on protoknowledge seems to come down to the following claim: protoknowledge is epistemically more valuable than mere true belief since protoknowledge is true belief produced by an informant who is objectively likely to produce true beliefs in her audiences or recipients. Alas, whatever the epistemic value of this objective likelihood, it is swamped by the epistemic value of the true belief. The reason why the objective likelihood is swamped in this way is that it is of instrumental value only. The detectably good informant is an instrument—is of instrumental value—for achieving true beliefs. But once the goal has been reached, the instrument's value is cancelled out, and only the value of the goal remains.

Kvanvig considers two ways of avoiding the swamping problem in the context of justification. One way is to give justification an intrinsic value, the other way is to make justification extrinsically but non-instrumentally valuable. Kvanvig dismisses the first idea on the grounds that we just cannot make sense of justification as valuable in isolation, that is, in isolation from having true beliefs. And, as previously summarized, Kvanvig accepts the notion that justification is extrinsically but non-instrumentally valuable.

Williams's account of collective action on behalf of the institution of testimony in general, and his new understanding of intrinsic value in particular, decisively changes the balance of arguments between Craig and Kvanvig. First, if Williams is right to insist that the institution of testimony is a collective good *supported by intrinsic values*, then the swamping argument begins to dry out. Conceived along the lines suggested by Williams's analysis, the institution of testimony is not of merely instrumental value. Its very existence is of central *intrinsic* value to us. To see talking to others—and that includes informing them and being informed by them—as of merely instrumental value is to lose hold of much of our value system. It is to begin a process of radically redefining who we are. Put differently, *there is* a value gap between true belief and protoknowledge: protoknowledge is true belief produced (or made likely) by the institution of testimony, an institution that is of epistemic and non-epistemic value to us both as a tool for reaching true beliefs and as something that makes us—as knowers, as believers, as humans—who we are. Second, if Williams is right to maintain that intrinsic values cannot be identified via Moore's method of absolute isolation, if Williams is correct to hold that intrinsic values are intelligible and reflectively stable only in relation to one another, then even Kvanvig's argument against justification as an intrinsic value no longer appears compelling. Justification can be an intrinsic value despite the fact that its value is related to the value of true belief.

Of course all of this depends on whether, concerning intrinsic value, Williams is right, and Kvanvig and Moore are wrong. To properly adjudicate this question would demand a separate lengthy investigation. Here I can offer no more than

crass appeals to the reader's intuition and my own introspection. Consider the list of intrinsic values once compiled by William Frankena:

life, consciousness, and activity; health and strength; pleasures and satisfactions of all or certain kinds; happiness, beatitude, contentment, etc.; truth; knowledge and true opinions of various kinds, understanding, wisdom; beauty, harmony, proportion in objects contemplated; aesthetic experience; morally good dispositions or virtues; mutual affection, love, friendship, cooperation; just distribution of goods and evils; harmony and proportion in one's own life; power and experiences of achievement; self-expression; freedom; peace, security; adventure and novelty; and good reputation, honor, esteem, etc. (Frankena 1973: 87–8; quoted from Zimmerman 2004)

Ask yourself the question of how you make sense of the valuable character of these items. Is it by means of Moore's 'method of absolute isolation' or is it by means of Williams's network test? If you are at all like me, then the answer will be 'the latter'. It is not just that I spontaneously use Williams's method; I am not even sure how to make sense of the method of absolute isolation regarding intrinsic values like love, freedom, peace, or, indeed, knowledge. The moment I begin to reflect on what makes any of these valuable, I find myself pulled and pushed towards further items on Frankena's list.

8. MODIFICATION TWO: BARNES AND WILLIAMS ON KNOWLEDGE AND HONOURING

Williams teaches that the collective good of the testimonial institution can exist only if the virtues of accuracy and sincerity (amongst others) become intrinsic values. But how can communities make sure that their members look upon these virtues as intrinsic values? Williams has an answer to this problem as well. He tells us that 'people may be discouraged or encouraged, sanctioned, shamed, or rewarded with respect to this' (2002: 44), that a crucial motivation is the 'fear . . . of disgrace in one's own eyes, and in the eyes of people whom one respects and who one hopes will respect oneself' (2002: 116), that 'the motivations of honour and shame play an important part' (2002: 120), and that the structure 'of mutual respect and the capacity for shame in the face of oneself and others, is a traditional, indeed archaic, ethical resource, but it is still very necessary' (2002: 121). In a nutshell, Williams's idea is that a system of sanctioning—of honouring and dishonouring—encourages community members to constitute and maintain the institution of testimony and thus the virtues of accuracy and sincerity.

Here one might be tempted to complain that Williams in fact explains one collective good in terms of another. Is not sanctioning on behalf of the institution of testimony itself a collective good? And does not its existence pose exactly the same sorts of problems as were posed by the existence of

the institution of testimony? Why don't individuals seek to free-ride on the institution of sanctioning? It is surprising to note that Williams does not see this problem.

Fortunately, though, there is a response to this worry. This response was developed by the sociologist Barry Barnes in a different context. Barnes's proposal is best appreciated against the background of other proposed solutions to Free Rider problems. The first standard type of solution is the *individualism* of Rational Choice Theory. According to this conception, the tendency of self-interested rational individuals to free-ride can be held in check by offering additional and selective benefits. (Self-interested rational car-owners would buy catalytic converters if the government gave them special tax-credits.) The second type of solution is *functionalism* or *normative determinism*: individuals co-operate because they have internalized the same norms, and thus have developed the same dispositions to act. There is widespread agreement that neither of these solutions works. Individualism is refuted by the empirical observation that we often engage in collective action even when no selective benefits are in sight. Voting in general elections is a case in point. Functionalism suffers from the inability to explain the psychological processes underlying internalization, and from failing to explain how individuals can meaningfully apply internalized norms under various circumstances (Barnes 1995: chs. 1–2).

The problems besetting individualism and functionalism have increased the attractiveness of a third solution to social order problems: 'interactionism'. Barnes defends a version of this view. The key move of this approach is to deny that interacting with others—talking and sanctioning, for example—are costly 'instrumentalities'. Basic features of social interaction are, as it were, 'free' resources; that is, resources to which normally no costs can meaningfully be attached. Talking to others about things we (should) value, reminding them of the importance of our collectively beneficial institutions, and sanctioning them in cases of non-compliance are things we do without much calculation. It is in our nature to engage in these activities. This highly plausible element of interactionism can help to deflect the cart-before-the-horse objection to Williams's proposal that a system of sanctioning helps to establish the testimonial system. If the sanctioning is a free resource then no circularity is involved.

It is worth adding that Barnes, like Williams, gives a central role to honour, shame, and related emotions. Sanctioning in support of the collective good works through the 'deference-emotion system' (Barnes 1995: 72–3, 82–3; Scheff 1988). The precondition of this system is our emotional need to continuously monitor how others treat us and think of us. We respond to our assessments of this treatment by changing our position on an internal scale that ranges from pride to shame. When we believe that others treat us with deference, when we believe that others honour us, we feel pride (and related feelings) and move ourselves up on the pride–shame scale. When we suspect that such deference and honouring are missing, we tend to feel bad about ourselves and slide downwards

towards the shame end of the scale. This emotional dependence on others is exploited by the deference-emotion system. The granting or withholding of deference constitutes a subtle system of social sanctions, a system that we barely notice. And the operation of the deference-emotion system, is inseparable from our ongoing conversation about the collective good. Working with and through the deference-emotion system, this conversation continuously re-establishes the importance of the collective good in everyone's mind.

So far in this section I have sought to explain—following Barnes—that honouring on behalf of the collective good of the testimonial system is a free resource. With this proposal in place, I can now introduce a slightly more original move. This is the idea that attributions of (proto-)knowledge (and their cognates) play a key role in the collective action that constitutes the institution of testimony. They do so by *honouring* the informants. In keeping with the genealogical method, let us first see how the link between protoknowledge attributions and honouring works under the simplified conditions of the state of nature. To publicly apply the concept *protoknower* to someone is not only to classify them as a reliable source of information, it is also to honour them, or to encourage others to do likewise. To classify someone as a protoknower is to praise them for their contribution to the institution of testimony, and thereby for their contribution to the well-being of the community itself. After all, the community cannot survive without the institution of testimony. By means of attributions of protoknowledge members of the community honour good informants for contributing their part to the existing and flourishing of the community. *Mutatis mutandis*, withholding or denying protoknowledge is a way of censoring and dishonouring. It is to mark someone as not willing, not worthy, or not able to participate in the constitution of the collective good, and thus as not fit to be a (working) part of the group. In that sense, to deny someone *protoknowledge* is to expel them from the group.

My suggestion connects Craig's focus on conceptual needs of the inquirer with Williams's emphasis on the motivational problems of the informants. By using the concept of protoknowledge to both classify and honour informants, the inquirer manages to serve two key goals at one and the same time: the goal of tagging good informants for future reference (to herself and others), and the goal of motivating community members to make, or keep making, investigative investments.

9. MODIFICATION TWO AS A PARTIAL REPLY TO THE SOCIAL AND RELEVANCE OBJECTIONS

It should be obvious that *Modification Two* addresses the *Social Objection* at least in part. We are no longer just focusing on the conceptual needs of the inquirer but also on the motivational needs—the need for honour—of the informant. At

the same time, it is clear that we have only *begun* to make progress in correcting for the one-sidedness of Craig's state-of-nature scenario. We shall make further headway once we turn to *Modification Three*.

The *Relevance Objection* expresses doubts about the use of state-of-nature stories for understanding our own concepts—in particular it is sceptical of the possibility that narratives about *protoknowledge* help us to illuminate features of *knowledge*. Proponents of this objection demand to see the 'marks' of *protoknowledge* in our concept.

My main concern in this section is to show that such marks are visible as far as the honouring aspect of protoknowledge attributions is concerned. But before I turn to this issue, it seems worth pointing out that Craig could reasonably invoke linguistic evidence for the claim that *knowledge* has (some of its central) roots in the idea of testifying. There is for instance the observation that 'to know' (and its cognates) takes interrogative constructions as complements. The same is not true for 'to believe' (and its cognates). Compare, for example:

- (1) I know who she is. She knows where to go. He knows when to come.
- (*2) I believe who she is. She believes where to go. He believes when to come.

The sentences in (1) are all well formed, those in (*2) are not. This observation has prompted one communitarian epistemologist to claim that 'we use the notion of knowledge in describing our aims when we address questions to other people' (Welbourne 1993: 9; cf. Vendler 1979; Kusch 2002). Or consider the frequent use of the metaphor of testimonial transmission (or its failures) in processes of learning and knowing:

- (3) The tyre tracks *tell* us that the assassin arrived by car after it rained.
- (4) The appearance of a distant ship's mast before its hull *shows* that the Earth's surface is not flat but curved.
- (5) I couldn't *believe* my eyes.
- (6) I refused to believe the *testimony of my senses*.
- (7) Two cemeteries *bear witness* to the congregation's history.
- (8) Unless my ears *deceive* me, my daughter is crying downstairs.
- (9) Jones is very well *informed* about this insect: he has studied it for years.

(3) is one example of the many cases where non-human things are treated as testifiers. It usually is evidence that tells us things. In such cases 'show' is a near synonym (cf. 4). While less obviously a testimony word than 'tell', 'show' still carries the implication of someone showing you something. But used in its common metaphor it is again evidence showing you something. Interestingly enough, the classical Greek word for evidence (*martyrion*) derives directly from the Greek word for witness (*martyrs*) (Lloyd 1979: 252–3). (9) is

interesting given that the term 'informant' plays such key role in Craig's account. We say that someone is well informed even when there has been *no person* doing the informing. Note that similar comments could also be made about many other terms relating to inquiry or cognition. For instance, 'discovery' first meant the act of revealing something one already knows to someone else. Only later did it come to refer to one's own encounter with something heretofore unknown (Bennett 2002). And 'intelligence' means both a property of the capacity to understand and reason, and the conveyance of information. (Hence the joke that 'military intelligence' is an oxymoron (Schaffer 1994).) Needless to say, all this is not conclusive—but for all that, there is some impressive linguistic evidence for the central position of testimony in our thinking about knowledge.

Returning to my main concern of using *Modification Two* as a partial reply to the *Relevance Objection*, I cannot here present linguistic evidence for the honouring function of knowledge attributions. To bring out this function demands a detailed study of the context of the utterance, and there is no room for this here. But fortunately there is other evidence, too; there is the evidence of the intuitions and hypotheses of epistemologists and sociologists of knowledge. In using such data, I am once more following and developing Craig's lead. He writes that 'we can include amongst our *explananda* . . . the various analyses of the concept of knowledge that philosophers have given' (1990: 6–7).

First, in one of the perhaps most fruitful uses of Craig's and Williams's work to date, Miranda Fricker (1998) has argued in defence of the category of 'epistemic injustice'. It is epistemically unjust when someone with a true belief on whether or not *p* is not allowed to testify concerning *p*. Examples of such epistemic injustice are racist societies in which, say, blacks or Jews are barred from appearing as witnesses in court, or sexist societies in which even when, say, the experiences of menstruation or childbirth are in question, the opinion of the male doctor counts for more than the reports of women. Fricker is entirely right to speak of epistemic injustice in such cases. But I would like to add a footnote. The footnote is that to deny members of ethnic minorities and women the status of knowers involves more than a purely epistemic injustice—it is more than a misclassification. To deny these groups the status of knowers is to *dishonour* them and to exclude them from full membership in the community. That denials of knowledge have this function is evidence for the claim that (many) knowledge attributions carry an element of honouring.

Second, various virtue epistemologists have emphasized either or both of the following ideas: that to attribute knowledge to someone is to declare them deserving of 'credit' (Sosa 1991; Greco 2004; 2005; Riggs 2002), and that any 'separation of knowledge from moral concerns' is unhelpful and misleading (Zagzebski 1996: 336). I do not need to say much about the former point: it is obvious that to give someone 'credit' for the way in which they have arrived at

their true belief is to honour them in some way.¹² Concerning the latter point Linda Zagzebski summarizes her argument in the following passage:

Virtue is a heavily social concept . . . Knowledge is the result of acts of intellectual virtue by both the agent and others in her epistemic community upon whom she relies in forming a subset of her beliefs. This aspect of knowledge makes it something for which we have social responsibility. And, in fact, it is one of the main reasons knowledge is something for which we are *morally* responsible. Others in our society have the moral right to expect us not to harm them, and passing on to them something that is not knowledge is one way of harming them. (1996: 319)

Zagzebski links epistemology and ethics via the idea that not to exercise one's intellectual virtues is often to expose others to harm. This point reinforces the link between knowledge attributions and honouring. We honour good informants not just because of the abstract consideration that they help maintain the collective good of the testimonial system. The collective good of the testimonial system is somewhat distal from our ordinary everyday point of view. More proximate is the insight that the good informant (often) helps us to avoid harm or enables us to fulfil our desires. Thus the honour we bestow is not exclusively epistemic—it is both epistemic and moral at the same time.

Third, recall that Kvanvig speaks of so-called 'honorific' uses of 'knowledge': uses such as 'the current state of scientific knowledge'. Kvanvig thinks that sometimes we wish to pay a special compliment to certain systems of beliefs that, although they contain falsehoods, have been assembled through praiseworthy epistemic efforts. When used honorifically, 'knowledge' refers to something which in fact is not knowledge, but which resembles knowledge in respects that we find worth honouring. Here I shall not raise the question whether Kvanvig is right to insist that 'the current state of scientific knowledge' refers to something which is not really knowledge. At this point I only wish to stress that Kvanvig's proposal is explained by my suggestion according to which knowledge attributions often contain an element of honouring.

Fourth, and finally, the mentioned links between knowledge, morality, credit, and honour are a frequent topic in the sociology of knowledge. For instance, in lectures criticizing the individualism of William James's pragmatism, no one less than Émile Durkheim once wrote of the 'moral obligation' to acquire knowledge: 'Truth cannot be separated from a certain moral character. In every age, men have felt that they *were obliged* to seek truth' (1983: 73). 'We feel obliged to adhere to truth. We see our certainty as something that is not personal to us, and that it is to be shared by all men' (1983: 101). David Bloor makes a similar point when he suggests that scientific knowledge can be analysed in terms of the Durkheimian category of 'the sacred' (1991: 51–3).

¹² Of course there are exceptions. No one would think that I deserve credit for coming to know what the weather is like, if all I need to do, in order to find out, is to look out of the window in front of me (cf. Lackey 2007). Remember that we are not searching for necessary and sufficient conditions.

10. MODIFICATION THREE: BARNES AND SHAPIN ON KNOWLEDGE AND FREEDOM

Modification Three is influenced by Barnes's social theory of freedom and Steven Shapin's social-historical work on testimony and gentlemen in seventeenth-century English natural philosophy (Barnes 2000; Shapin 1994). *Modification Three* is concerned with the ways in which attributions of freedom are intertwined with attributions of knowledge. To repeat, the three crucial claims are: (a) attributions of honour are attributions of freedom from (certain forms of) interference; (b) attributions of protoknowledge are attributions of discretion over a practice; (c) informants have the conceptual need to mark out inquirers who can be (rationally) influenced by good information. The concept of *protofreedom* plays this role. I shall take these three claims in turn.

Re (a) To honour someone is to give them a status that increases their chances of not being interfered with by other members of the community. The higher the honorific status of a group member, the smaller the range of others who can tell her what to do. This should not be a controversial claim. After all, dictionary entries on 'honour' link it to respect, esteem, reverence, deference, approbation, and high rank. All these concepts concern a status that protects one—at least to a degree—from being controlled and directed by others. I have previously suggested that attributions of protoknowledge do honour to the testifier. It therefore follows that attributions of protoknowledge confer and sustain a (culturally and contextually varying) degree of freedom. One important form of non-interference that comes with the status of (proto-)knower is protection from the demand to investigate things further. If we attribute to Fred the knowledge that *p*, then we accept that there is no further room to negotiate whether or not *p* (Lynch 1993: 76). Or, as Kvanvig puts it: 'One of the platitudes about the functional role of knowledge ascriptions is that it is a legitimator of inquiry closure. Nothing similar can be said about belief, true belief, or justified or virtuous true belief' (Kvanvig 2009: 344–5). (In passing I add two further links between knowledge, honour and freedom. First, *honour* is due primarily, or perhaps even exclusively, for actions that people *freely* chose to do; honour is due only for actions of which the actors *knew* what they were doing. Second, no group member can occupy a status of honour in a group unless all, or at least most, group members know of it. Indeed, it is the knowledge of the status that constitutes it in the first place. To occupy a social status is to be known to occupy it.)

Re (b) Assume in our state-of-nature community Nick is the best informant about where to find the tastiest wild mushrooms. It surely would then be a good

idea to give Nick discretion over our practices of mushroom hunting. This will maximize our chances of getting the delicacies we want, and it will show Nick additional respect. I suggest generalizing the point by saying that attributions of protoknowledge come with the acknowledgement that the testifier has at least a *prima facie* case for directing our practices in the relevant area. Put differently, when Fred attributes protoknowledge to Nick he thereby restricts his own discretion in favour of Nick's.

Re (c) Claims (a) and (b) both deal with the social consequences of protoknowledge attributions. Suggestion (c) tries to remove an *asymmetry* in Craig's treatment of the inquirer–informant interaction. Craig's theoretical focus is exclusively on the conceptual needs of inquirers, that is, on their needs to tag good informants. Alas, Craig has nothing to say on the corresponding conceptual needs of informants, that is, on their needs to tag good inquirers (i.e. recipients). However, in the absence of any argument to the contrary, it seems likely that both sets of concepts, or both conceptual needs, have left their marks on our understanding of key epistemic terms and their interrelations.

What distinction regarding good and bad inquirers should be important to informants? This is where my train of thought makes contact with Barnes's social theory of free will and determinism (2000). In a thought experiment that is genealogical in all but name, Barnes asks us to consider what kinds of action and decision classifications must be important to highly gregarious and socially interdependent creatures like us. A little reflection shows that such creatures must find it important to distinguish between two kinds of decisions by others: decisions by others they can influence by means of verbal interventions; and decisions by others that they cannot so influence. I can bring about your decision to fetch me a glass of water by asking for it or by explaining how important the water is to me. 'Fetching a glass of water' is a type of action that typically can be brought about by symbolic intervention. However, if you are pathologically obsessive about washing your hands, then typically I cannot stop you from doing so by asking you to stop. Your decision to wash your hands thus falls on the other side of the divide. Barnes suggests that we see our distinctions between voluntary and involuntary actions, and between free and coerced decisions, in light of the distinction between decisions that can be influenced by communication, and decisions that cannot be so influenced. What—at least in the simplified situation of state-of-nature scenario—unifies our criteria for attributing free decisions and voluntary actions is the idea of being susceptible to verbal intervention. A careful look at the decisions that we end up classifying as either free or coerced suggests that susceptibility to change on the basis of symbolic intervention is central. Put differently, it seems that for us a decision is free if it 'could have been otherwise if symbolic intervention had occurred' (Barnes 2000: 73).

Barnes is not primarily concerned with the conceptual needs of informants in tagging good and bad recipients or audiences. But the point carries over. If

protoknowledge is the concept for picking out good informants, *protofreedom* is the concept for tagging good recipients. The protofree recipient of testimony is open to the information conveyed to her, and willing and able to modify her behaviour in predictable ways in its light. Take Nick, our (state-of-nature) informant on the location of tasty mushrooms. Assume that Nick can give his information either to Fred or to Otto. Fred is notorious for his inability to act in predictable (i.e. rational) ways. Tell Fred that a mushroom is safe to eat, and he will throw it away. Let Fred know that a mushroom is poisonous, and he will run over to your children to let them bite off a piece. Compare Fred with Otto. Otto is thankfully predictable and makes precisely the use of information that Nick would expect. According to the proposal made here, Nick will find it important to tag Otto and Fred in different ways: Otto as protofree, Fred as proto-unfree. I write 'protofree' rather than 'free' since it would clearly be contentious to claim that *protofreedom* is identical with our *freedom*. Barnes only makes the weaker claim that our *freedom* (with all its surrounding conflicting intuitions) has developed out of *protofreedom*. This genealogy lies beyond the scope of this paper.

11. MODIFICATION THREE AS A FURTHER REPLY TO THE SOCIAL AND RELEVANCE OBJECTIONS

It hardly merits mentioning that *Modification Three* immunizes genealogy against the *Social Objection*. After all, *Modification Three* develops genealogy in exactly those respects in which the *Social Objection* detects a deficit. Hence I shall here concentrate on the contribution of *Modification Three* to the question of relevance.

I want to address this question by trying my hand at linking together imaginary and real genealogy. In doing so I seek to improve on Williams's vague hints concerning a link between knowledge and freedom. Picking up on themes in Primo Levi's reflections on natural science, Williams insists that to lie to another person is to violate their freedom; that to follow the 'dictates' of nature is not to be un-free; and that the 'freedom to believe the truth must be shared' (2002: 146). Williams does not tell us whether these links are the product of real or imaginary genealogy.

My central source of inspiration here is Shapin's study *A Social History of Truth: Civility and Science in Seventeenth-Century England* (1994).¹³ Shapin studies the role of testimony and gentility in the shaping of the investigative culture of English natural philosophy during the lifetime of Robert Boyle. In

¹³ Shapin (1994) further develops a theme that was already important in Shapin and Schaffer (1985).

order to highlight the general relevance of Shapin's study for our concerns here, I need hardly do more than quote parts of one of its key early paragraphs:

Knowledge is a collective good. In securing our knowledge we rely upon others, and we cannot dispense with that reliance. That means that the relations in which we have and hold our knowledge have a moral character . . . [T]he fabric of our social relations is made of knowledge—not just knowledge of other people, but also knowledge of what the world is like—and, similarly . . . our knowledge of what the world is like draws on knowledge about other people—what they are like as sources of testimony, whether and in what circumstances they may be trusted. (1994: pp. xxv–xxvi)

In the present context I want to highlight a single central theme of Shapin's book: the notion that in order to be a good and reliable testifier—and thus in order to be a fully trustworthy producer of knowledge—one must be free. Natural philosophers of seventeenth-century England accepted—notwithstanding the Royal Society's motto *Nullius in verba* (on no man's word)—that testimony was needed to make natural knowledge, and that finding good testifiers was a difficult practical problem. Seventeenth-century literature on the topic suggested any number of maxims, such as '(i) assent to testimony which is plausible; (ii) assent to testimony which is multiple' or '(iii) assent to testimony which is consistent' (1994: 212). Alas, no sooner was a maxim proposed than critics found it to be of only limited help. For instance, maxim (i) could potentially and incorrectly exclude true reports that conflicted with dominant false beliefs; maxim (ii) was of limited value in cases where the multiple reports derived from one another; or maxim (iii) did not account for the experience that 'too great a display of internal consistency' might be a signal 'that a polished performance had been prepared' (1994: 233). Only one maxim was never challenged: 'assent to testimony from sources of acknowledged integrity and disinterestedness' (1994: 212). On first sight, this might not sound like a helpful maxim; how were practitioners of the new science supposed to know who the disinterested reporters were?

This is where gentility and freedom became important. Shapin's major historical thesis is that 'English experimental philosophy . . . emerged partly through the purposeful relocation of the conventions, codes, and values of gentlemanly conversation into the domain of natural philosophy' (1994: p. xvii). And one of the central beliefs about the gentleman was that he was a natural truth-teller. The gentleman was thought to be a natural truth-teller primarily on the grounds of his possessing a special 'disinterestedness' (1994: 83). This disinterestedness was taken to derive in good part from the gentleman's economic circumstances. As Henry Peacham put it in his influential treatise *The Complete Gentleman* at the time: 'whosoever labour for their livelihood and gain have no share at all in nobility or gentry' (Shapin 1994: 50). In other words, to be a gentleman was to be financially independent and secure. And a life that was independent and secure in this sense was equated with a free life. The overarching thought linking freedom, gentry, and testimony together was thus as

follows: 'Gentlemen were truth-tellers because nothing could work upon them that would induce them to be otherwise' (1994: 84).

The conventions of gentlemanly experimental philosophy did not allow for anyone to openly express disbelief in a report coming from a gentleman. The situation was very different for all those who did not make the gentry grade: women, servants, 'the poor and the mean in general', merchants, Catholics, Continental gentry, Italians, and politicians. In the cases of all of these groups, their 'unreliable truthfulness . . . was pervasively referred to their constrained circumstances' (1994: 86). This obviously did not mean that no one but a gentleman was ever believed: given that much of, say, Boyle's experimental work was carried out by 'domestics' such widespread distrust would have destroyed the whole enterprise of natural philosophy. One has to see the link between freedom and truthfulness as a resource: citing the constrained circumstances of domestics—technicians was a way in which a gentleman could explain experimental failures. And whatever information the domestics produced, it became knowledge, and thus a property of the gentlemanly community of natural philosophers, only once it was vouched for by Boyle (or another gentleman) (1994: ch. 8).

I hope that this briefest of summaries of Shapin's central theme makes convincing that his *Social History of Truth* qualifies as a 'real genealogy' by Williams's criteria. After all, Shapin shows us how a conceptual link between freedom and testimony, or indeed freedom *and knowledge*, was articulated in a period that was formative of natural science as it evolved over the past 300 years. Clearly, Shapin's book—and investigations like it (e.g. Shapin and Schaffer 1985)—are relevant for our understanding of (scientific) knowledge today. Although they do not help much with solutions to the Gettier problem, or the dispute between externalism and internalism, they help us understand why, say, we have our present intuitions about the links between freedom and testimony, or why testimony lies at the heart of our thinking about knowledge.

One way to measure the interest of imaginary genealogy is to ask whether it can make a helpful contribution to a better understanding of Shapin's results. It seems clear to me that indeed it can. Consider first the suggestion that to attribute protoknowledge is to attribute honour, freedom, and social power (i.e. discretion over a social practice). Reading Shapin's observations in the light of this suggestion helpfully highlights the correct 'direction of fit' between knowledge and freedom. It would be superficial to think that gentlemen excluded women, domestics, Italians, and so on, from the category of knowledge-makers on the grounds that these latter groups were constrained in their circumstances and hence not free. This reading of the situation treats freedom and lack thereof as if they were natural states one just happened to find oneself in, natural states that have nothing to do with human actions. Combining my imaginary genealogy with the realities of a stratified or class society invites a different rendering. Precisely because knowledge attributions come with (a degree of) honour, freedom, and social power, members of an upper class or caste will

always be (*prima facie*) reluctant to ascribe knowledge to members of a lower class or caste. Denying knowledge is a way of dishonouring and keeping someone in their disempowered 'station'. Or reflect on the further proposal coming from my imaginary genealogy, according to which testifiers prefer recipients who are free, that is, who are being likely to be able, and have the means, to use the transmitted information in a way intended by the testifier. Again, it should not be thought that the gentlemen 'found'—through empirical observation of their stratified society—that domestics, women, and Italians lacked these means. On the contrary, through their actions gentlemen made sure that these and other marginalized or suppressed groups would not acquire the needed tools for knowledge making.

To sum up my discussion of *Modification Three* and the *Relevance Objection*, Craig's imaginary genealogy—enriched by a broader communitarian perspective—is important for understanding *our concept of knowledge*. One move in making it relevant is to connect it appropriately with real genealogies of our knowledge practices. Imaginary genealogy can add insight and depth to the real genealogies by providing explanations based upon the speculative material of the state of nature.

12. CONCLUSIONS AND LOOSE ENDS

In this paper I have tried to give substance to the idea of a communitarian value-driven epistemology. In order to make transparent how this project might slot into more familiar, or more mainstream, projects, I have throughout maintained a critical dialogue with Kvanvig's position. The communitarian value-driven epistemology I favour starts from the fact that human knowers are highly gregarious and deeply interdependent, and draws on traditions usually ignored by epistemologists: Craig's and Williams's epistemic genealogy and the sociology of knowledge. The core of this paper was an attempt to develop Craig's genealogical study of the value of knowledge. Four 'Kvanvigian' objections to Craig—two taken from, two inspired by, Kvanvig's work—pointed the way. The objections were that Craig's imaginary genealogy falls to the swamping argument; that 'pragmatic accounts' conflate the context of discovery with the context of the final product; that Craig's state of nature is insufficiently social; and that Craig's overall quasi-historical theory is irrelevant to an understanding of our present epistemic concepts. I sought to block all four objections by means of a number of modifications of Craig's original story: by treating the institution of testimony as a collective good underwritten by the intrinsic and interrelated values of accuracy and sincerity; by rendering protoknowledge attributions as ascriptions of honour; and by allowing attributions of protoknowledge to be intertwined with attributions of freedom. The view that emerged was that the core of our knowledge practices are institutions of testimony, and that these practices are a collective good.

I am painfully aware that this paper is no more than a beginning, or a research proposal more than a finished piece of work, and that it begs numerous important questions. I conclude by listing a few further objections that should receive more proper treatment elsewhere. I shall indicate in each case the directions in which I hope a successful answer will lie.

Objection A It is easy to see that knowledge is not a collective good. Take a familiar collective good like clean air. Clean air is a collective good insofar as its costs are carried individually while its benefits are enjoyed collectively. Clean air can exist only if most of us invest into pollution-reducing technologies. But once clean air exists, everyone—regardless of their investment, or lack thereof—is able to breathe it. This is of course why collective action—that is, the action that creates and sustains collective goods—is always threatened by free-riding. Knowledge is not like clean air. There are plenty of things each one of us can learn and know without getting help from anyone else. For example, I do not need anyone's help in order to gain perceptual, memorial, or inferential knowledge. Moreover, it is not obviously true that the benefits of knowledge are enjoyed collectively. I am unable to do much with the knowledge of stamp collectors or high-energy physicists.

Reply A reply to *Objection A* might be sought in several directions. One idea is to insist that clean air—the standard case of a collective good—is not as different from knowledge as the critic assumes. We think of clean air as a collective good despite the fact that some people, in some circumstances, can enjoy its benefits individually. Think of people with respiratory illnesses who spend long periods of time in an oxygen tent. Or consider rich people who can afford to fly off to unpolluted islands. Nor should we be overly impressed with the fact that not everyone can enjoy the benefits of knowledge. The same is true for clean air. Even when the collective good has been established, there still might be some people—the poor, for example—who are not able to fully take advantage of the high air quality. Perhaps they are forced to live next to a garbage dump that locally makes for high levels of air pollution.

A different sort of reply might try to fend off *Objection A* by restricting the *scope* of the thesis that knowledge is a collective good. One highly plausible restriction is to suggest that *scientific* knowledge is a collective good. There is plenty of evidence for this view: it is pretty much impossible to think of any new item of scientific knowledge that is not dependent on myriad previous elements, many of which have reached the knower by way of testimony (Barnes, Bloor, and Henry 1996; Hardwig 1985; 1991; Kusch 2002; Shapin 1994). Or think of the self-referential social knowledge that constitutes the social world (Barnes 1988; Searle 1995). The social world is what it is in good part because of what we believe it to be. And typically these beliefs are true precisely because of the way in which we collectively keep passing them around the community (Kusch 1999).

Objection B The argument presented in this paper is no contribution to value-driven epistemology—it changes the topic. Value-driven epistemologists are interested in the question what makes individual items of knowledge epistemically valuable. But this paper discusses the question whether *bodies* or *systems* of knowledge have value.

Reply In doing so I am only following (the earlier) Kvanvig's valuable advice to epistemologists: focus on 'bunches of people rather than isolated individuals, bodies of knowledge rather than individuated propositions' (1992: 186). If that constitutes a change of topic, then the topic ought to be changed.

Objection C I take this objection from Timothy Williamson's influential book *Knowledge and its Limits* (2000: 31). There Williamson defends the views that the concept of knowledge cannot be factored into constituents and that knowledge is the most fundamental factive mental state. Williamson sympathizes with Craig's book but finds it 'too close to the traditional programme' of analysing knowledge as true belief plus a third and fourth condition. As Williamson sees it, Craig goes wrong when he stipulates that inquirers want 'true beliefs about our environment, as though this were somehow more basic than our need for knowledge of the environment'.

Reply A discussion of Williamson's attempt to replace traditional epistemology by a metaphysics of factive mental states lies well beyond the scope of this paper. I therefore have to confine myself to a couple of very general comments. It seems to me important to distinguish between two theses. The first thesis says that *knowledge* can be factorized into components. The second thesis is the claim that the concept of knowledge, analysed along Williamson's proposal, is the product of a conceptual development that at some point involved a factorized concept. It is unclear what reason Williamson has given us for rejecting the second thesis. In other words, one might argue that *knowledge* as the non-analysable concept of a certain kind of factive mental state is itself the endpoint of a conceptual evolution that starts in the conceptual situation described in Craig's state of nature.

My second general comment picks up on recent criticisms of Williamson's proposal. Some commentators have suggested that Williamson's theory does not in and by itself undo the familiar issues of traditional epistemology (cf. Reed 2005 and Whitcomb 2005). On closer inspection, the familiar problems with accidentality all reappear. In chapters on which I did not report here, Craig pays a lot of attention to the emergence of such problems with accidentality in the case of protoknowledge. Perhaps this makes his work relevant to Williamson's concerns after all.

Objection D How can this paper claim to have identified a knowledge-specific epistemic value when it has said nothing about how Gettier cases can be blocked?

Reply I have tried to show that we can think of knowledge as a collective good without any worked-out proposal of how to deal with Gettier cases. This does not mean that an understanding of Gettier cases is irrelevant. But we do not need to wait for agreement on an anti-Gettier-clause before we can make headway with rendering explicit the social-epistemic value of knowledge. Note also that Reynolds (2002) suggests that the Gettier problem might receive a solution in terms of testimonial norms. If that is right, then this would strengthen the case made here.

Objection E Given that Craig's *protoknowledge* is not threatened by accidentality, why is it the ancestor of *knowledge* and not of *understanding*?

Reply As Craig's state of nature stands now, there is no straightforward answer to this question. But I do not see why we should not be able to come up with a genealogical-narrative explanation that would make sense of the social need to operate with both concepts.¹⁴

REFERENCES

- Barnes, B. (1988). *The Nature of Power*. Urbana and Chicago: University of Illinois Press.
 — (1995). *The Elements of Social Theory*. London: UCL Press.
 — (2000). *Understanding Agency: Social Theory and Responsible Action*. London: Sage.
 — and Bloor, D., and Henry, J. (1996). *Scientific Knowledge: A Sociological Analysis*. London: Athlone.
 Bennett, J. (2002). 'The Travels and Trials of Mr Harrison's Timekeeper', in M. N. Bourguet, C. Licoppe, and H. O. Sibum (eds.), *Instruments, Travel and Science*. London: Routledge.
 Bloor, D. (1991). *Knowledge and Social Imagery*. Chicago: University of Chicago Press.
 Code, L. (1995). *Rhetorical Spaces: Essays on Gendered Locations*. New York and London: Routledge.
 Craig, E. (1990). *Knowledge and the State of Nature: An Essay in Conceptual Synthesis*. Oxford: Clarendon Press.

¹⁴ I am indebted to an unusually large group of colleagues. Cornelius Menke and Joab Rosenberg helped me in coming to grips with Craig's work. Alexander Bird, Stefan Brenner, Samantha Evans, David Feller, Mike Finn, John Forrester, James Hyslop, Ben Irvine, Nick Jardine, Adam Kay, Jeff Kochan, Sir Geoffrey Lloyd, Paula Olmos, Flavia Padovani, Jason Rampelt, Martin Rudwick, Simon Schaffer, Thomas Soderqvist, and Jorg Tuske provided examples of testimonial metaphors in our talk of knowledge. Axel Gelfert, Jeff Kochan, Melissa Lane, Peter Lipton, Richard Raatzsch, and Koen Vermeir helped improve a first draft. Fernando Broncano, Dennis Whitcomb, and Jesus Zamora acted as commentators at the two conferences where this paper was presented; Whitcomb at the 2006 Sterling meeting on the value of knowledge, Broncano and Zamora at the 2006 annual meeting of the Spanish Association for the Philosophy of Logic and Science held in Granada. I have also given talks drawn from the paper in Bristol, Cambridge, Durham, and Tartu (Estonia). Many excellent suggestions and challenges were put to me by the audiences on these occasions.

- DeRose, K. (1997). Review of A. Stroll, *Moore and Wittgenstein on Certainty* (New York: Oxford University Press, 1994), *Philosophical Review*, 106: 238–41.
- Durkheim, E. (1983). *Pragmatism and Sociology*. tr. J. C. Whitehouse, ed., J. B. Allcock. Cambridge: Cambridge University Press.
- Elgin, C. Z. (1996). *Considered Judgment*. Princeton: Princeton University Press.
- (2005). 'Williams on Truthfulness', *Philosophical Quarterly*, 55: 343–52.
- (2009 this volume). 'Is Understanding Factive?' Accessed online 8 July 2006 at <www.philosophy.stir.ac.uk/staff/duncan-pritchard/documents/ElginOnKvanvig.pdf>.
- Frankena, W. K. (1973). *Ethics*. Englewood Cliffs: Prentice-Hall.
- Fricker, M. (1998). 'Rational Authority and Social Power: Towards a Truly Social Epistemology', *Proceedings of the Aristotelian Society*, 98: 159–77.
- and Hornsby, J. (eds.) (2000). *The Cambridge Companion to Feminism in Philosophy*. Cambridge: Cambridge University Press.
- Gagnier, R., and Dupré, J. (1998). 'Reply to Amariglio and Ruccio's "Literary/Cultural Economies"', Economic Discourse and the Question of Marxism', in M. Woodmansee and M. Osteen (eds.), *The New Economic Criticism*. Routledge, 1998.
- Goldman, A. (1999). *Knowledge in a Social World*. Princeton and Oxford: Oxford University Press.
- Greco, J. (2004). 'Knowledge as Credit for True Belief', in M. DePaul and L. Zagzebski (eds.), *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Oxford University Press.
- (2005). 'Chapter Nine. The Value Problem'. See this volume, Appendix B. Accessed online 18 July 2006 at <<http://www.philosophy.stir.ac.uk/staff/duncan-pritchard/documents/GrecoOnTheValueProblem.pdf>>
- Grice (1989). *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
- Hanfling, O. (2000). *Philosophy and Ordinary Language: The Bent and Genius of Our Tongue*. London: Routledge.
- Hardwig, J. (1985). 'Epistemic Dependence', *Journal of Philosophy*, 82: 335–49.
- (1991). 'The Role of Trust in Knowledge', *Journal of Philosophy*, 88: 693–708.
- Jones, W. E. (1997). 'Why Do We Value Knowledge?', *American Philosophical Quarterly*, 34: 423–39.
- Kant, I. (2004). *Prolegomena to Any Future Metaphysics*. Cambridge: Cambridge University Press.
- Kripke, S. (1982). *Wittgenstein on Rules and Private Language*. Oxford: Blackwell.
- Kusch, M. (1999). *Psychological Knowledge: A Social History and Philosophy*. London: Routledge.
- (2002). *Knowledge by Agreement: The Programme of Communitarian Epistemology*. Oxford: Oxford University Press.
- (2006). *A Sceptical Guide to Meaning and Rules: Defending Kripke's Wittgenstein*. Chesham: Acumen.
- Kvanvig, J. L. (1992). *The Intellectual Virtues and the Life of the Mind: On the Place of the Virtues in Epistemology*. Savage, MD: Rowman and Littlefield.
- (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge: Cambridge University Press.
- (2009 this volume). 'Responses to Critics'. Accessed online 5 July 2006 at <www.missouri.edu/~kvanvig/certain_doubts/Responses.pdf>.

- Lackey, J. (2007). 'Why We Don't Deserve Credit for Everything We Know', *Synthese*, 158: 345–61. Accessed online 3 July 2006 at <<http://www.niu.edu/phil/~lackey/Credit.pdf>>.
- Lane, M. (1999). 'States of Nature, Epistemic and Political', *Proceedings of the Aristotelian Society*, 99: 211–24.
- Lloyd, G. E. R. (1979). *Magic, Reason and Experience*. Cambridge: Cambridge University Press.
- Lynch, M. (1993). *Scientific Practice and Ordinary Action: Ethnomethodology and Social Studies of Science*. Cambridge: Cambridge University Press.
- Malcolm, N. (1949), 'Defending Common Sense', *Philosophical Review*, 58: 201–20.
- (1977). 'Moore and Wittgenstein on the Sense of "I know"', in N. Malcolm, *Thought and Knowledge*. Ithaca and London: Cornell University Press, 170–99.
- Moore, G. E. (1993). *Principia Ethica*, ed. T. Baldwin. Cambridge: Cambridge University Press.
- Neta, R. (2006). 'Epistemology Factualized: New Contractarian Foundations for Epistemology', *Synthese*, 150: 247–80. Accessed online 10 July 2006 at <philosophy.unc.edu/Epistemology%20Factualized.pdf>.
- Nietzsche, F. (1990). *Twilight of the Idols/Anti-Christ*. London: Penguin.
- Pritchard, D. (2007). 'Recent Work on Epistemic Value', *American Philosophical Quarterly*, 44: 85–110. Accessed online at <www.philosophy.stir.ac.uk/staff/duncan-pritchard/documents/RecentWorkOnEpistemicValue.pdf>.
- Reed, B. (2005). 'Accidentally Factive Mental States', *Philosophy and Phenomenological Research*, 71: 134–42.
- Reynolds, S. L. (2002). 'Testimony, Knowledge, and Epistemic Goals', *Philosophical Studies*, 110: 139–61.
- Riggs, W. D. (2002). 'Reliability and the Value of Knowledge', *Philosophy and Phenomenological Research*, 64: 79–96.
- (2007). 'The Value Turn in Epistemology', in D. Pritchard and V. Hendricks (eds.), *New Waves in Epistemology*. London: Palgrave Macmillan. Accessed online 15 July 2006 at <www.philosophy.stir.ac.uk/css/documents/NewWavesRiggs2.pdf>.
- Schaffer, S. (1994). 'Babbage's Intelligence: Calculating Engines and the Factory System', *Critical Inquiry*, 21: 201–28.
- Scheff, T. J. (1988). 'Shame and Conformity: The Deference-Emotion System', *American Sociological Review*, 53: 395–406.
- Schmitt, F. F. (1992). Review of Craig 1990, *Mind*, 101: 555–9.
- (ed.) (1994). *Socializing Knowledge: The Social Dimensions of Knowledge*. Lanham, MD: Rowman and Littlefield.
- Searle, J. R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
- (1995). *The Construction of Social Reality*. London: Penguin.
- Shapin, S. (1994). *A Social History of Truth*. Chicago: University of Chicago Press.
- and S. Schaffer (1985). *Leviathan and the Air-Pump*. Princeton: Princeton University Press.
- Sosa, E. (1991). *Knowledge in Perspective*. Cambridge: Cambridge University Press.
- Travis, C. (1991). Critical Notice of H. P. Grice, *Studies in the Way of Words*, *Mind*, 100: 237–64.

- Vendler, Z. (1979). 'Telling the Facts', in P. A. French, T. E. Uehling, and H. K. Wettstein (eds.), *Contemporary Perspectives in the Philosophy of Language*, Midwest Studies in Philosophy, 2. Minneapolis: University of Minnesota Press, 220–32.
- Weinberg, J. M. (2006). 'What's Epistemology for? The Case for Neopragmatism in Normative Metaepistemology', in S. Hetherington (ed.), *Epistemology Futures*. Oxford: Oxford University Press.
- Welbourne, M. (1993). *The Community of Knowledge*. Aldershot: Gregg Revivals.
- Whitcomb, D. (2005). 'Factivity without Safety', *Pacific Philosophical Quarterly*, 89/1: 143–9. Accessed online 4 July 2006 at <www.eden.rutgers.edu/~dporterw/fws.pdf>.
- (2006). 'Wisdom in Mind and Action'. Accessed online 4 July 2006 at <<http://www.eden.rutgers.edu/~dporterw/wma.pdf>>.
- Williams, B. (2002). *Truth and Truthfulness: An Essay in Genealogy*. Princeton and Oxford: Oxford University Press.
- Williamson, T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.
- Winch, P. (1958). *The Idea of a Social Science and its Relation to Philosophy*. London: Routledge & Kegan Paul.
- Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Blackwell.
- (1961). *Notebooks 1914–1916*, tr. E. Anscombe. New York: Harper & Row.
- (1969). *On Certainty*, ed. G. E. M. Anscombe and G. H. v. Wright. Oxford: Blackwell.
- Zagzebski, L. T. (1996). *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge: Cambridge University Press.
- Zimmerman, M. (2004). 'Intrinsic and Extrinsic Value', *Stanford Encyclopedia of Philosophy*. Accessed online 13 July 2006 at <www.plato.stanford.edu/entries/value-intrinsic-extrinsic/>.

4

The Value of Understanding

Jonathan Kvanvig

1. INTRODUCTION

Understanding has a special kind of value that other epistemic states such as knowledge do not, and this fact threatens the justification for the focus on knowledge that the history of epistemology displays. I have argued elsewhere that knowledge does not possess this special value.¹ There are a couple of lines of argument, however, that threaten to extend the denial of this special value for knowledge to a denial of a special value for understanding. Underlying all such challenges is the obvious fact that the language of knowing and the language of understanding are closely related. In the end, however, I hope to show that the kind of understanding that we prize most is immune from the concerns I have argued plague the theory of knowledge.

The kind of understanding in question I term ‘objectual understanding,’ as when one understands combustion or quantum mechanics or Republican ideology. I will begin with a brief discussion of some alternative types of understanding, but will move relatively quickly to the threats I see to the special value of objectual understanding. One threat is that, if objectual understanding is special, then so should objectual knowledge be special. Another threat is that objectual understanding can be analyzed in terms of understanding followed by a *wh*-phrase, most especially understanding *why*. Since understanding *why* and knowing *why* are intimately connected, any difficulties plaguing an account of the special value of knowledge will also plague an account of the special value of understanding. If either difficulty can be sustained, we will have to conclude that understanding has no greater chance of having the special value in question than does knowledge.

I will begin with some remarks about the notions of understanding and knowledge that will be in focus here.

¹ See Kvanvig (2003).

2. TYPES OF UNDERSTANDING AND KNOWLEDGE

When we are examining cognitive successes and achievements from a purely theoretical, as opposed to practical, point of view, two grammatical forms involving understanding stand out.² First is propositional understanding, and the grammatical form in question employs a 'that'-clause. For example, propositional understanding is involved when we say that Jason understands that 'knows' is not a gradable adjective. The other primary form is objectual understanding, and here the grammatical form takes an object, as when we say that Bas understands quantum theory or Cheney's hunting buddies understand the cost of hunting with him.

There is also the range of attributions of understanding followed by 'wh'-clauses: Cheney's buddies understand where not to stand, when to duck, what to do to avoid hospitalization, why they should prefer other hunting partners, etc. In each such case, it is tempting to adopt the thesis that such uses can be explained in terms of propositional understanding. Such is likely not the case with understanding how, but we have come to expect that epistemic terms involving a 'how'-clause are more relevant to practical concerns than theoretical ones, so those can be ignored here.

The question of the relationship between objectual and propositional understanding mirrors a similar question about the relationship between objectual and propositional knowledge. One can know that Bush was President as well as know Bush; one can know that quantum theory is counterintuitive, and also know quantum theory. Moreover, the verb 'knows' also takes 'wh'-complements, and it is plausible here as well to think of such constructions as reducible to propositional knowledge.³

The central point I want to emphasize here, however, is not one about the reductive relationships in question, but rather the difference in focus when understanding is before our minds rather than knowledge. When we think about the nature of understanding, what is foremost in our minds are the ways in which pieces of information are connected with each other. To understand is to grasp the variety of such connections. It involves seeing explanatory connections, being aware of the probabilistic interrelationships, and apprehending the logical

² For consideration of the variety of attributions of understanding beyond those of central interest to epistemology, see Kvanvig (2003: ch. 8).

³ Such a supposition mirrors a similar approach in the theory of knowledge, where the standard view is that knowledge-wh can be explained in terms of knowledge-that. See e.g. Hintikka (1975), Lewis (1982), Boer and Lycan (1986), Higginbotham (1996), and Stanley and Williamson (2001). Schaffer (2007) has criticized such an approach in the theory of knowledge, arguing for a contrastivist relationship between a given question and the reducing propositional content in question. Here I will ignore the interesting issues involved, since the emendations required on the supposition that the contrastivist position is best do not affect the value issues here and are relatively easy to insert as modifications of what I say here, presupposing the standard view.

implications of the information in question. There is, of course, an element of factivity to the notion of understanding, just as there is with the notion of knowledge. But when we move past the alethic aspect of both notions, our attention turns to diverse paths. When the question is whether one knows, the issues that are foremost in our minds are issues about evidence, reliability, reasons for belief, and, perhaps most importantly, non-accidentality regarding the connection between our grounds for belief and the truth of the belief. When the question is whether one has understanding, the issues that are foremost in our minds are issues about the extent of our grasp of the structural relationships (e.g. logical, probabilistic, and explanatory relationships) between the central items of information regarding which the question of understanding arises. The questions are those that led Plato to the method of collection and division, where the end result of the application of the method is a full grasp of a thing through a classification that sorts from genera to species in such a way that the thing in question is fully identified through the system of classification in question. Some have suggested that the application of this method is what Plato intended as an answer to the question at the end of the *Theaetetus* concerning the nature of *logos* which perhaps was taken by Plato to differentiate knowledge from true opinion.⁴ Such an account of knowledge surely over-intellectualizes it, since the reflective efforts needed to complete the method are more than is ever achieved by small children and animals and even ordinary adults in cases of common perceptual knowledge. But once we recall the dispute over whether '*episteme*' should be translated as 'understanding' rather than 'knowledge,' it is not as surprising to find such intellectualizing in the account, since the information generated by the process of division is just the sort of information one would expect when understanding is present. For when understanding comes to mind, the central elements in focus are ones concerned with structural relationships between various pieces of information grasped by the possessor of understanding, unlike the central element of non-accidentality in focus when one is reflecting on the concept of knowledge.

It is worth issuing a note of caution once we have reached this point. The cautionary note concerns the predilection in the epistemology of the last century or so to attempt to settle substantive philosophical disputes by appeal to ordinary language. Such an approach here would counsel us to consider and systematize the various ways in which the language of knowledge and understanding are used in common parlance. Though ordinary parlance might be a useful starting point when thinking about substantive philosophical issues, it is surely a mistake to think that whatever forces have led to the particular meanings encoded in ordinary language are also forces sure to encode in a way that resolves philosophical issues. We should distinguish the linguistic data regarding the use of epistemic terminology from substantive points about the phenomena in

⁴ See e.g. Sayre (1969).

question. In the present context, there is no question that much of ordinary parlance treats 'understanding' and 'knowledge' as interchangeable. Here we should remind ourselves that ordinary parlance also counts as knowledge things that are not believed, are not true, and are merely rational to believe. Philosophers, like other theoreticians, have a responsibility to develop a vocabulary adequate to their subject matter, and in developing such a vocabulary they will often counsel revising ordinary language instead of merely reporting it. My remarks about the fundamental difference between understanding and knowledge are elicitable from our ordinary concepts of the two, but I make no claim here that the distinction I'm pointing to between the two is somehow faithfully reflected in ordinary language or some facet of it thought to be especially useful for revealing proper philosophical theory. It may well be that certain uses of 'understanding' should be interpreted as synonymous with knowledge claims. What interests me is understanding itself, not the way we talk. So a cursory response to appeals to what we'd say or what would be correct to assert is this: hardly anything philosophically interesting will follow about the nature or value of understanding from such data.

To the great dissatisfaction of some, however, I won't engage in extended discussion about exactly what the standards of proper theory construction are in philosophy, since my stated goal concerns the nature and value of understanding rather than philosophical metatheory. I bring the issue up only to prevent simplistic appeals to ordinary intuitions about the conditions under which ordinary language will allow the predication of understanding in a given case. The proper response to such purported refutations should be the same as the response we give to those who claim that knowledge doesn't entail true belief on the basis of my daughter's exclamation after the 2004 election: 'I just *knew* Kerry would win!' The proper response is that when doing epistemology we are focusing on a particularly significant intellectual accomplishment, one which ordinary language makes visible to us on occasion and through a glass darkly at least. Beyond that, ordinary language is the ladder we kick away having once climbed, to put the point enigmatically and hyperbolically.

Returning to the agenda and applying the above points to it, the point of note in the above is the difference in focus between the phenomena in question. We note, on the ordinary language side, that there is understanding-that, understanding-wh, and objectual understanding. There is also knowledge-that, knowledge-wh, and objectual knowledge. We also notice that in any given case in which we use an epistemically relevant variant of one term, we could have used the other. We could say either that Joe knows baseball or that he understands it; that Bo knows jazz (though you have to be careful not to pick the star athlete as the contextually determined referent here!) or that he understands it; that Sara understands that her appointment is at 11 or that she knows it; that Zac knows biology or that he understands it. Given such interchangeability, it is very natural to be drawn to the conclusion that there is not a significant difference between knowledge and understanding, perhaps to the extent that the

terms are interchangeable because synonymous. Such a hypothesis would have to explain away the propriety of other utterances, such as 'I know Peter but I don't understand him at all'; but such a task doesn't seem to be any more insurmountable than defending the factivity of knowledge by explaining away my daughter's exclamation that she just knew Kerry would win.

I don't believe the synonymy claim can be sustained, for if it were sustainable, the remarks I made about what is brought to mind when we think about knowledge versus understanding would have to be explained away and I don't believe that this can be done in service of a synonymy claim. But suppose it can. Then the synonymy claim would be irrelevant to my purposes, for what I'm interested in is the difference between two phenomena, a difference recounted in the claims I made earlier about our focus when we think about knowledge and when we think about understanding. If the terms 'knowledge' and 'understanding' turned out to be synonymous, there would still be this difference in two phenomena, and for purposes of clarity I would need to pick different terms or invent new ones to refer to the phenomena in question to avoid misleading the reader. I won't do that here, however, since I believe the difference in focus is enough to undermine the synonymy claim, and will assume that point until shown otherwise.

The history of epistemology focuses on one of these phenomena, a phenomenon requiring a focus on elements such as quality of evidence, ruling out alternative hypotheses, not being right simply by accident, and not having gotten to the truth just in virtue of a certain kind of fortuity or luck. The other phenomenon we can begin to focus on by attending to the language of objectual understanding, and here the focus is on the grasped relatedness of the items that constitute a body of information possessed by the individual in question. Regarding this latter phenomenon, it is not enough that the explanatory connections exist or that they could be discovered easily by the individual with only a little effort or reflection. Understanding involves an already-possessed awareness of the explanatory and other connections involved in the subject matter in question, an already-mastered grasp that involves or generates the illumination of a subject we resort to the language of intelligibility and sense-making to convey. To understand first-order proof theory is to have made sense of the structure of the theory and for this structure to be intelligible and plain to one. This difference between knowledge and understanding leaves open the possibility of having lucky understanding where the kind of luck in question is the kind demonstrated to be incompatible with knowledge in the Gettier literature. Historical understanding, for example, can be achieved where one's sources and one's dyslexia combine in the right way to generate correct dates even though the sources are mistaken. In such cases, one's understanding can be displayed by one's ability to answer correctly any question put to one regarding the subject matter in question. But such examples fit the model of classic Gettier cases where the information about the inaccuracy of the sources serves to undermine the claim to know.

If these points can be sustained, then the special problems that attend any account of the value of knowledge may not plague an account of the value of understanding, for if the features that constitute the nature of understanding are factivity plus the grasping of internal, structural relationships between pieces of information, then the value of understanding can be explained without fear of refutation from the ad hocery and gerrymandering endemic to proposals designed to ward off the Gettier problem. In short, knowledge lacks the unique value under discussion here because of the Gettier problem; understanding doesn't, because one can be lucky in having understanding.

One may resist this separation of knowledge and understanding by claiming that I'm contrasting objectual understanding, not with the properly analogous objectual knowledge, but with propositional knowledge. One may then resist further by claiming that once we correct for this difficulty and compare propositional understanding with propositional knowledge and objectual understanding with objectual knowledge, the perceived differences between knowledge and understanding will disappear altogether.⁵

The answer to this objection begins by noting that, whatever we want to say about the connections between the objectual level and the propositional level, we shouldn't pull objectual knowledge so far apart from propositional knowledge that we allow Gettier intrusion to play a role regarding propositional knowledge but give no place to such concerns regarding objectual knowledge. Such a distinction would allow an easy reply to the worry that my distinction between knowledge and understanding involves a failure to control for type, since we could identify objectual knowledge with objectual understanding and claim that this phenomenon is much different from the focus on propositional knowledge that characterizes the history of epistemology. But I think separating objectual knowledge from propositional knowledge in this way is too strong a claim. We should expect some relationship between these kinds of knowledge, and we should expect some relationship between propositional and objectual understanding.

Once we grant that there should be some relationship here, we can see more clearly how the worry noted above might be developed. What matters to my thesis is that there are two quite different phenomena here, and that holding each clearly before one's mind, the difference in focus described above is ascertainable. In one case, something akin to non-accidentality (I'm happy to put the point in terms of underdetermination as well, though I haven't done so) is in focus; in the other, apprehended interrelationships dominate. But suppose we identify objectual understanding with explanatory understanding, that is, with understanding why, and then claim that understanding why and knowing why are indistinguishable. If we tie understanding to knowledge at the level of explanation in this way, and then adopt the usual view that knowledge involving *wh*-complements can be identified with propositional knowledge, we are only one step away from

⁵ See Brogaard (2006).

eliminating the distinction between knowledge and understanding. All that is left is to note, as I did above, that we should expect some relationship between propositional and objectual levels for both understanding and knowledge. So, if objectual knowledge can thus be understood in terms of some collection of propositional knowledge, the argument is complete: once we pay attention to the relationships between various kinds of understanding and knowledge, we find out that we can go from objectual understanding to explanatory understanding expressed in terms of understanding-why, and then to knowing-why, which in turn takes us to propositional knowledge and finally to objectual knowledge. Moreover the road just outlined goes both directions, and thus the purported distinction between knowledge and understanding disappears.

I think, however, that there is no such road available, and the place where the above story goes wrong concerns the relationship between objectual and propositional understanding. The central difficulty for the above view involves the possibility of indeterministic systems and our capacity to understand them. Understanding incorporates explanatory connections (when they exist), and one relevant kind of explanation depends on the presence of causation (when it exists). In indeterministic systems, things happen that are uncaused, both deterministically and probabilistically. Though some theories of probabilistic causation, such as Wesley Salmon's mark transmission account,⁶ imply the existence of causation even when the result in question is completely indeterministic, that is a defect of such theories rather than a strength. If the probability of an electron going to the left is precisely the same as that of going to the right (and there is no hidden variable to account for the difference), then whichever way it goes is the result of chance rather than of causation. Chance here is, of course, not to be reified into a further explanatory or causal factor: it is, instead, just a denial of the presence of deterministic or probabilistic causation. In such cases, the connection between causation and explanation is tight: if there is no cause of the electron going to the left rather than the right, there is no explanation why the electron went to the left either.

Given these points, here's what we must say about indeterministic systems and our understanding of them. Where *S* is some indeterministic system, we can have objectual understanding of the system even though we cannot interpret this understanding in terms of being able to understand why things happen as they do in *S*. We will have an understanding of the system itself without any explanation of why some of the truths of the system are true. We thus have no understanding why these truths are true, since there is no explanatory basis for it, and there is no explanatory basis for it because the events in question are irreducibly indeterministic in such a way that there is no causal explanation as to why the actual events occurred rather than some

⁶ See Salmon (1984). For critical discussion of Salmon's approach, see Dowe (1992; 1995) and Kitcher (1989).

other events. So objectual understanding cannot be reduced to propositional understanding via appeal to 'wh'-complement attributions of understanding or explanations.

Once this argument is granted, we can achieve a more unified conception of understanding by allowing these considerations to generalize to our understanding of deterministic systems as well. We may wish to hold that objectual understanding of such systems is constituted by some collection of propositional understandings, but we should not identify objectual understanding with propositional understanding. To do so would forgo the possibility of a unified treatment of the relationship between the two with no theoretical advantage gained by the disunity.

We should also say some of the same things on the knowledge side, concerning the relationships between propositional and objectual knowledge. When we do so, we preserve the intuitive connection between these different kinds of knowledge, and doing so allows us to see that paying attention to the type of knowledge or understanding doesn't block the distinction between the two. In the case of propositional knowledge, something akin to non-accidentality is in focus; in the case of objectual understanding, apprehended interrelationships dominate. But because the different levels are related in important ways, we should not expect non-accidentality to disappear from view when we attend to the nature of objectual knowledge.

Suppose, however, that I'm wrong about this point. Suppose, that is, that knowledge-that is a quite different animal from objectual knowledge, so different that the difference in focus I claimed existed between knowledge and understanding actually exists as well between knowledge-that and objectual knowledge. What follows is that I would need to re-label the issue I want to focus on. Instead of using the language of knowledge and understanding to point to the two different phenomena in question, I'd be better advised to use the language of propositional versus objectual achievements, whether termed 'knowledge' or 'understanding.' I would have linguistically carved the phenomena apart in a way orthogonal to what really separates them. The distinction itself would remain intact, of course, but the terminology would need to change.

I'm not convinced that such a change in terminology is appropriate, however, since I don't think we have a good argument for thinking that propositional knowledge and objectual knowledge are that far apart. So I will continue to use the language of knowledge and understanding to point to the difference between two kinds of intellectual achievement.

Once we have become clear about the distinction in question and in particular about the difference in focus accompanying each of the two phenomena in question, we are in a position to assess the claim that understanding has a special kind of value that knowledge does not. In order to understand the issue before us, we need to rehearse a bit the argument for the claim that understanding has a special kind of value that knowledge does not.

3. GETTIER INTRUSION AND THE VALUE QUESTION

The fundamental problem with accounting for the value of knowledge involves explaining how what are traditionally termed the third and fourth conditions for knowledge generate a composite more valuable than true belief or justified true belief, respectively. Regarding the third condition, there is the now well-known swamping problem that plagues standard versions of reliabilism and threatens to undermine other views as well. The problem, at its core, is that to the extent that justification is conceived in instrumental terms relative to the goal of truth, the presence of truth swamps any value that justification might add. Once we've already assumed that a belief is true, learning that it also has a property whose significance is clarified in terms of an instrumental relationship to truth fails to teach us anything that would change our evaluative perspective on the belief in question.

In response to this problem, there are two quite different avenues to pursue in attempting to explain how justified true belief has more value than true belief. Some have held that virtue theories face this problem better than other versions of reliabilism.⁷ Here I will not pursue the intricacies of this position, but will grant it. The other avenue is to characterize justification so that it is extrinsically, though not instrumentally, related to the goal of truth. This path is the one favored by most internalists.

Even if this problem can be avoided, however, there is a more debilitating one concerning the fourth condition. As I and others have argued,⁸ the better an approach to the Gettier problem is at carving cases of knowledge off from cases of non-knowledge, the more ad hoc and gerrymandered the proposal. The result is a condition which has no hope whatsoever of giving a decent answer to the question of what makes un-Gettiered justified true belief more valuable than justified true belief. Hence, the hope of defending the view that knowledge is more valuable than any proper subset of its parts is dim indeed.

Not so with understanding, however, as we have already noted. When understanding comes to mind, the central elements in focus are ones concerned with structural relationships grasped by the possessor of understanding, unlike the central element of non-accidentality in focus when one is reflecting on the concept of knowledge. With understanding, the focus is on the grasped relatedness of the items that constitute a body of information possessed by the individual in question. This difference between knowledge and understanding leaves open the possibility of having lucky understanding where the kind of luck in question is the kind claimed to be incompatible with knowledge in the Gettier literature. Historical understanding provides a rich source of examples of this

⁷ See e.g. Riggs (2002*a*; 2002*b*), Sosa (2002), and Greco (2004). See also Kvanvig (2003: ch. 4).

⁸ See especially Williamson (2000).

sort, since one's justified true beliefs can be Gettiered by faults in one's sources that are fortunately compensated for by one's own memory failures, resulting in a present noetic structure that displays lucky understanding. But such examples fit the model of classic Gettier cases, where the information about the inaccuracy of the sources serves to undermine the claim to know.

If these points can be sustained, then the special problems that attend any account of the value of knowledge may not plague an account of the value of understanding. For if the features that constitute the nature of understanding are factivity plus the grasping of internal, structural relationships between pieces of information, then the value of understanding can be explained without fear of refutation from the ad hocery and gerrymandering endemic to proposals designed to ward off the Gettier problem. In short, knowledge lacks the unique value under discussion here because of the Gettier problem; understanding doesn't, because one can be lucky in having understanding.

Some demur on the point about Gettier, however. Consider the following case and argument from DePaul and Grimm:

suppose that your source for World Cup soccer scores and analysis is a Jayson–Blair style news reporter who simply makes up all of his reports about the Cup whole-cloth. You have no particular reason to suspect this about him, moreover, so from your point of view his reports seem worthy of your default trust. Your source then claims (in a particular moment of reverie) that the United States defeated Italy 2 to 1, and that the winning goal was scored by the U.S. because the Italian goalkeeper slipped in the mud, an account that you then accept as true. In this case, moreover (what luck!), all of his claims turn out to *be* true; by chance, he has precisely described the way things actually unfolded during the game. Focusing now on the role of understanding, we can also add the following: not only do you come to accept (based on his account) that the United States scored the winning goal because the goalie slipped in the mud, but you also (in some appropriately internal sense) 'grasp' or 'see' the explanatory relationship described by the reporter; that is, you 'grasp' or 'see' that the winning goal was scored *because* the goalie slipped in the mud. This is, to your mind, *why* the winning goal was scored.

According to Kvanvig, all the ingredients required for understanding now seem to be present: there is the internal 'seeing,' for one thing, and there is also the truth of the connection seen (that is, roughly, the truth of the explanatory story). But do you now genuinely understand why the winning goal was scored? It seems not. Because the details were invented whole-cloth by the reporter, even though the connection that you 'see' or 'grasp' actually obtains, you no more understand *why* the winning goal was scored by the U.S. than you know *that* the winning goal was scored by the U.S. (DePaul and Grimm 2007: 498–514)

DePaul and Grimm claim that you don't have propositional knowledge in this case, and that this lack translates into a lack of understanding concerning why the winning goal was scored on the basis of the same Gettier-like considerations involved in the case. In short, they wish to use an identification between understanding-why and knowledge-that to undermine what I have claimed about the special value of objectual understanding.

There are two points to note about this attempt. First, it should be noted that I have not identified objectual understanding with the short list of characteristics noted in the quoted passage. I have not said that understanding is simply a matter of some internal seeing of explanatory connections that in fact obtain. I have said that when we think carefully about objectual understanding, these are the sorts of things that are in primary focus and in contrast to the notions of evidence and luck that are in focus when we think about propositional knowledge. But I have also said that justification plays a central role in objectual understanding as well. I did not stress this point extensively in my book on the subject, but it is a point that is important in this context. As I have already noted, there are at least two approaches to justification that I believe have good responses to the *Meno* problem, one a very subjective approach and the other a virtue approach. In order for the soccer game example to show that there are Gettier-like issues for justification, the case would have to be constructed so that the understanding in question involved beliefs that are subjectively rational. If that test is passed and the case is still worrisome, it would not immediately call for a revision in the claims I have made about the value of understanding, but instead would call for a more minor revision in terms of adding that the beliefs in question should be required to be intellectually virtuous as well, since such a characteristic of beliefs adds value to true belief. In the case above, it would appear to be quite simple to construct the case so that the beliefs in question are all subjectively rational, but it is not as easy to see how they count as intellectually virtuous. On the usual accounts of intellectually virtuous belief, the beliefs need to be reliable in order to be virtuous and on a natural reading of the example the beliefs in question are not reliable.

I do not wish to rest a response to the case on this point, however, so I am going to grant here that the beliefs in question are both subjectively justified and intellectually virtuous. In developing the response I want to give to the above case, I wish to begin by noting again that there is something correct and expected about the claimed identification between understanding-why and knowledge that, at least so far as ordinary language goes. To get a counterexample to the approach recommended here, however, we need more than a case in which understanding is plausibly denied because of the kinds of features central to Gettier cases in the theory of knowledge. The reason we need more than this is because the theory in question here relies not so much on some purported difference in the appropriate linguistic uses of the terms in question, but rather on a difference in the phenomena at the objectual and propositional levels. In particular, for the soccer case above to threaten the account in question, we need to add to the case that objectual understanding is present in some way. As presented, the attempted counterexample relies instead on the claim that you don't understand why the winning goal was scored, and one mark of this is that you don't know that the winning goal was scored by the US. In order to consider carefully the merits of the example, then, we must change the case so

that it attributes objectual understanding of some sort, and so that the objectual understanding in question is connected in some way to the failure to understand why the winning goal was scored.

As pointed out, the example doesn't involve any claim about objectual understanding, and it is hard to see what the object of understanding is supposed to be here. Perhaps, though, we could say the following: what is in question is your understanding of the results of the game.

If we add this claim to the example, then the force of the example depends on the connections that are involved in the case. There are three connections required for the counterexample to work against the claim I have made about understanding. The first is an implication between objectual understanding and understanding-why, based on the phenomena in question. The second is a connection between understanding-why and knowing-why, and the third is between knowing-why and knowing-that. I believe objectual understanding involves psychologically grasping or apprehending whatever explanatory connections are present, but that isn't the same claim as the claim that objectual understanding implies understanding-why. Whether that claim should be granted depends on the tightness of the relationship between knowing-why and understanding-why (one might question the reducibility of knowing-why to knowing-that, but I've already granted that point). The answer here depends, I think, on an issue we have already addressed but left open, the issue about whether synonymy, or something close to it, holds between understanding and knowledge, once the type in question is controlled for. If we insist on synonymy, then the proper response is to deny that objectual understanding implies understanding-why. If synonymy is rejected, then in cases where one's objectual understanding involves realizing why certain things happen or are true, we can endorse the claim that one thereby understands why those things happen or are true. One doesn't know why, however, because to do so would require propositional knowledge, and one lacks such knowledge precisely because of the Gettierizing feature of the case.

It is easy to imagine at this point the impatience in the critic's voice: 'Look, if the terms are synonymous, there is simply no basis whatsoever for distinguishing between knowledge and understanding. To suggest otherwise, on whatever basis, is to succumb to confusion. Either they are synonyms or they are not. Only if they are not can your account get off the ground, so you'd better argue that they are not synonymous!'

The conclusion here is false. If there is such synonymy, there is still the difference between the propositional and objectual levels, and the difference in focus which I have been referring to using 'knowledge' for the former and 'understanding' for the latter would still exist. Perhaps the critic could still complain that I should pick different terms for the distinction, and I would honor such a request as soon as it is shown that the two terms are truly synonymous—that is, if it can be shown that there is no difference in cognitive significance between knowledge claims and understanding claims, once the type

in question is fixed. I don't know how to establish that point, and I doubt that it is true, so I will continue to refer to the distinction in question in terms of the distinction between knowledge and understanding.

I must admit, however, to having theory-infected responses to ordinary language claims at this point. When the example claims that you don't understand why because you don't know, I balk. But I also admit that ordinary language dispositions may not honor my balking, and so that those such as DePaul and Grimm who talk the way this example goes are not misusing language (much as it isn't a misuse of language to use 'knows' non-factively). What I deny, however, and what I claim any interesting philosophical methodology ought to deny, is that philosophical conclusions can be read off of ordinary language in this way. The interchangeability of the terminology in question in ordinary language might indicate synonymy or it might just as easily indicate the lack of any commonly noticed significant interest or purpose that arises in ordinary life that would require distinguishing the two. In any case, whatever the correct explanation for the linguistic data here, the most anyone could legitimately claim would be that such data is defeasible evidence for certain philosophical conclusions. I don't know whether the data rises even to that level, but if it does, then the account presented here and the explanations offered constitute, I submit, just the kind of defeaters one would need to refuse to draw the conclusions these data suggest.

So I think that the example fails to show that there is no objectual understanding of the results of the game, even if it is a good example of a failure to know why the US won or to understand why the US won the game. My own idiolectal inclinations are to insist that understanding-why does exist here, even though knowledge-why doesn't, but that is because I have been thinking about the differences between understanding and knowledge and realize that there will be some connection between objectual and explanatory understanding, even if the connection isn't universal because of indeterministic systems.

This point, however, raises the possibility of pressing the example DePaul and Grimm employ in a different way. This approach grants the distinction between objectual understanding and its other forms, and doesn't commit itself on any identity between understanding and knowledge even when the type in question is held fixed. Instead, it focuses on cases of objectual understanding involving systems where deterministic assumptions are at work or where whatever indeterminacies might exist are not relevant to the understanding involved. In such cases, one might grant that objectual understanding and understanding-why are distinct, but maintain that the former is constituted by the latter. In such a case, a failure to understand why the winning goal was scored would undermine the view that one understood the outcome of the game, even if understanding the outcome were different in kind from some compilation of understandings-why.

Such a response puts pressure on the approach I am taking to insist that you do understand why the US scored the winning goal. DePaul and Grimm claim

that you no more understand why this is the case than that you know that the US scored the winning goal, and we are all agreed that you don't know that.

You believe that the goalie slipped in the mud and that this slip is responsible for the US winning, but you only believe it on the basis of a report that is wholly the invention of a reporter. DePaul and Grimm ask whether this basis is sufficient for understanding why the US won and they report their own conviction that it seems not to be. But the only attempt at argument that I see is the comparison with knowledge-that. Such a comparison is best thought of, however, not as an argument, but as an invitation to the reader to reflect on one's own reaction and to note whether the reader's inclinations accord with the authors'.

So let me record my own inclination: I'm not inclined at all to think that you don't understand why the US won, though I do understand why others will be inclined to think so. The inclination to think so results from the interchangeability of knowledge locutions with understanding locutions in ordinary language. When such interchangeability is present, attempts at distinction often face opposition.

In ordinary language, for example, it is rare to distinguish hearing from listening: I listen to sounds around me and I hear them; I was hearing a lecture yesterday and listening to it; etc. But sometimes people use these terms to distinguish two different phenomena. People sometimes say that you can hear without listening, in the sense that sounds are reaching your auditory faculty but you are not consciously trying to access the sounds in question. People also sometimes say that you are listening without hearing, in the sense that you are paying attention to the information conveyed verbally, but not taking it to heart in the appropriate way.

What should we make of these various aspects of usage? Not much, I think. The gloss on the use is more important than the terminology employed, and if someone else puts pressure on the idea that you can hear without listening by getting us to attend to interchangeability uses enough to make us balk at the description, we shouldn't attach much significance to that fact. What matters is the underlying phenomenon. In the soccer match example, DePaul and Grimm stress the way in which the information you possess is defective. In particular, you don't realize that the reporter invented the story of the winning goal. In fact, it is likely that you are assuming that the reporter is telling you what he knows to be true. In such a case, there is something you lack understanding of: you don't understand how you came to have the understanding you do have of the outcome of the game. If assumptions count as mental states and are involved in the story of understanding in the way beliefs are, we can say something stronger, and that is that your understanding of how you came to understand the outcome of the game is weak or non-existent, since it involves an important falsehood. In such a way, we can honor the existence of some defect of understanding without this defect filtering into the explanatory understanding in question in the way it filters into the related propositional knowledge in question.

Even given such a response to the particular case in question, however, there is the more general question. Even if, because of the possibility of indeterministic systems, we distinguish objectual understanding from explanatory understanding, there is still the hypothesis to consider more directly that explanatory understanding is, in the deterministic case, that which constitutes objectual understanding. I am not certain at this point whether there is such a connection, but I'm inclined to think so. The point of the above response to the soccer example is to suggest a strategy for answering purported counterexamples involving explanatory understanding on the assumption that objectual and explanatory understanding are related in this way. The strategy is to acknowledge some defect in understanding that results from the deficiencies noted, but to distinguish understanding from knowledge because the defect has direct implications for the possession of propositional and explanatory knowledge itself. The implications of the defect for understanding are less direct, as in the case above, existing instead at the metalevel.

4. CONCLUSION

The account of understanding and its value presented above is thus defensible in the face of certain types of objections to it, but it is worth stressing in closing one mark of a purely theoretical sort in its favor. In recent years, Swampman cases have abounded in philosophy for various purposes, and one such purpose to which such examples have been used is to threaten the lessons of the Gettier literature. Swampman is a fully formed human being that arises by sheer accident out of the swamp as a result of some natural event such as a lightning strike.⁹ In the epistemologically significant version of the case, Swampman arises with incredible cognitive features. He is able to answer any question put to him about any subject whatsoever. He is not guessing. He is not making things up. He is testifying in the same ordinary way that you or I testify when we are trying to answer honestly and sincerely to questions we are inclined to answer. This example is intended to lead to the conclusion that the focus of the Gettier literature in finding some condition to rule out accidentality or luck is misplaced, since in the senses of these terms relevant to epistemology, Swampman's condition is accidental and lucky.¹⁰ Instead, the difference between knowledge and true opinion may be disappear simply by adding more true beliefs.

The argument from Swampman has not induced wholesale reconception of the theory of knowledge, but no response to the example can be adequate without

⁹ See Foley (1996).

¹⁰ For an excellent discussion of the variety of notions of luck, accidentality, and chance, and the differences between these concepts, see Pritchard (2005).

explaining what is so impressive about Swampman as well as what is missing. One possibility is the simple one that he has many more true beliefs than we do, about a broad and varied number of topics and subjects. That point is correct, but there seems to be something distinctively epistemic about his condition as well, since there is quite a bit of awkwardness in asserting that we know so much more than Swampman does. My suggestion has two aspects to it. The first aspect is that Swampman has understanding even though he doesn't have knowledge. He understands all the topics and subjects about which you may care to query. That's the first part of my suggestion. The second part is that the language of understanding and knowledge is so often interchangeable in ordinary language that once we recognize the understanding possessed, it is awkward to deny him knowledge.¹¹ The phenomenon here is exactly that underlying the questions addressed above about the theory defended here. Knowledge and understanding, once type is controlled for, seem nearly interchangeable in ordinary language, leading to the conclusion that one is Gettierizable if and only if the other is.

The approach taken here provides the resources for avoiding this inclination. Careful attention to the phenomena in question should allow us both to see the temptation and avoid it. In avoiding it, we can see how to explain the admirability involved in Swampman's ability to answer questions based on information possessed. And it allows us to see Swampman's defects as well, because he can't have any understanding of how he came to be so blessed by the gods to have the first-order comprehensive understanding that he has. It also allows us to do something more. It allows us to see the special and unique value of understanding, a value that warrants expanding the conception of epistemology beyond a focus on the theory of knowledge its history displays.

REFERENCES

- Boer, S., and Lycan, W. (1986). *Knowing Who*. Cambridge, MA: Cambridge University Press.
- Brogaard, B. (2006). 'Can Virtue Reliabilism Explain the Value of Knowledge?' *Canadian Journal of Philosophy*, 36: 335–54.
- DePaul, M., and Grimm, S. (2007). 'Review Essay: Jonathan Kvanvig's *The Value of Knowledge and the Pursuit of Understanding*,' *Philosophy and Phenomenological Research*, 74: 498–514.
- Dowe, D. (1992). 'Wesley Salmon's Process Theory of Causality and the Conserved Quantity Theory,' *Philosophy of Science*, 59: 195–216.

¹¹ Though, as Alvin Goldman reminded me in conversation, it is not awkward to attribute knowledge in the weak sense of the term in which knowledge can be identified with true belief alone. For discussion of the distinction between strong and weak types of knowledge, see Goldman (1986).

- (1995). 'Causality and Conserved Quantities: A Reply to Salmon,' *Philosophy of Science*, 62: 321–33.
- Foley, R. (1996). 'Knowledge is Accurate and Comprehensive Enough True Belief,' in J. Kvanvig (ed.), *Warrant in Contemporary Epistemology: Essays in Honor of Plantinga's Theory of Knowledge*. Savage, MD: Rowman and Littlefield.
- Goldman, A. (1986). *Epistemology and Cognition*. Cambridge, MA: Harvard University Press.
- Greco, J. (2004). 'Knowledge as Credit for True Belief,' in M. DePaul and L. Zagzebski (eds.), *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Oxford University Press.
- Higginbotham, J. (1996). 'The Semantics of Questions,' in S. Lappin (ed.), *The Handbook of Contemporary Semantic Theory*. Oxford: Oxford University Press.
- Hintikka, J. (1975). 'Different Constructions in Terms of the Basic Epistemological Verbs: A Survey of Some Problems and Proposals,' in Hintikka's *The Intentions of Intentionality and Other New Models for Modalities*. Dordrecht: Reidel.
- Kitcher, P. (1989). 'Explanatory Unification and the Causal Structure of the World,' in P. Kitcher and W. Salmon (eds.), *Minnesota Studies in the Philosophy of Science*, 13: 410–505. Minneapolis: University of Minnesota Press.
- Kvanvig, J. (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge, MA: Cambridge University Press.
- Lewis, D. (1982). 'Whether Report,' in T. Pauli (ed.), *Philosophical Essays Dedicated to Lennart Aqvist on his Fiftieth Birthday*. Uppsala: University of Uppsala.
- Pritchard, D. H. (2005). *Epistemic Luck*. Oxford: Oxford University Press.
- 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.
- Salmon, W. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.
- Sayre, K. (1969). *Plato's Analytical Method*. Chicago: University of Chicago Press.
- Schaffer, J. (2007). 'Knowing the Answer,' *Philosophy and Phenomenological Research*, 75: 383–403.
- Sosa, E. (2002). 'The Place of Truth in Epistemology,' in M. DePaul and L. Zagzebski (eds.), *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Oxford University Press.
- Stanley, J., and Williamson, T. (2001). 'Knowing How,' *Journal of Philosophy*, 98: 411–44.
- Williamson, T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.

5

Ugly Analyses and Value

Michael R. DePaul

The anatomist ought never to emulate the painter; nor in his accurate dissections and portraitures of the smaller parts of the human body, pretend to give his figures any graceful and engaging attitude or expression. There is even something hideous, or at least minute in the views of things, which he presents; and 'tis necessary the objects shou'd be set more at a distance, and be more cover'd up from sight, to make them engaging to the eye and imagination.

(David Hume, *A Treatise of Human Nature*, bk. 3, pt. 3, § 6.)

1. INTRODUCTION

I'm interested in exploring, and rejecting, a form of argument that is at least strongly suggested in three prominent, recent works on epistemology. The three works I have in mind are Stephen Stich's *The Fragmentation of Reason* (1990), Timothy Williamson's *Knowledge and its Limits* (2000) and, most recently, Jonathan Kvanvig's *The Value of Knowledge and the Pursuit of Understanding* (2003).¹ In Stich this form of argument seems to constitute a part of his case that true belief is not valuable. I take Kvanvig to use the argument in the course of showing that knowledge is not *distinctively* valuable. Williamson touches on the argument most briefly of the three. He does not use it to show that truth or knowledge or anything else we might be inclined to value in fact lacks value: he turns the argument on its head to reach the conclusion that knowledge cannot be analyzed. But Williamson provides a succinct, explicit statement of the kind of reasoning that concerns me. After stressing that knowledge matters and that it is important to us he comments, "This importance would be hard to understand if the concept *knows* were the more or less ad hoc sprawl

¹ Since I began working on this topic, many people have suggested other examples where this form of argument may have been employed within epistemology as well as other areas of philosophy. But as I have not yet been able to investigate these suggestions, I refrain from making further specific allegations regarding use of the argument in question.

that analyses have had to become; why should we care so much about *that*?' (2000: 31).

If any of these arguments succeeds, the consequences for traditional epistemology would be huge. If knowledge cannot be analyzed, over the last fifty years or so many epistemologists have devoted themselves to a quixotic quest. If, on the other hand, knowledge lacks value, then efforts to analyze knowledge may not have been doomed from the start, but they seem pointless: why would anyone much care exactly what knowledge is if it has no special value? And all the energy invested into arguments for and against skepticism would seem to be wasted as well: why should we get ourselves in such a knot about whether it is possible to have something that has no distinctive value? Finally, given our strong inclination immediately to accept that truth is valuable and the prominence of the view that the distinctive feature of epistemic evaluation is that it is made from a special point of view that is defined by a concern for truth,² it is hard to imagine what might do more to alter the shape of contemporary epistemology than a successful argument that truth lacks value. Hence, even if the form of argument that concerns me constitutes only one tine of the multi-pronged attacks mounted by Stich, Kvanvig, and Williamson, it behooves us to examine this form of argument, since the stakes are potentially so high.

I want to clear up some potential misunderstandings before proceeding. Look back to the quote from Williamson. The term 'knowledge' could possibly be a source of ambiguity, but I hope it is clear that the issue regarding the value of knowledge does not concern the value of some abstract thing such as a property or universal—Knowledge, or maybe even KNOWLEDGE. We do not take the ordinary claim that gold is valuable to concern the property or universal GOLD or any other abstract thing; we take it to be about the lumps of stuff that are gold. I understand the claim that knowledge is valuable in a correspondingly humble way, as concerning the particular items that count as knowledge.

I understand the assertion that knowledge is valuable as more substantial than the claim that *some* items of knowledge have *some* sort of value. We should not conclude that scraps of paper are valuable just because some particular scrap of paper is valuable in some way, say because a significant historical figure wrote an important note upon it and it has survived into the present. Similarly, someone might value some peculiar item of knowledge, say the date and time of Einstein's birth, because of special circumstances—for example, because the person's psychology is oddly constructed so that he gets a great deal of giddy pleasure from knowing this—but we should not say that knowledge is valuable only for this reason. I understand the question of whether knowledge is valuable to ask whether items of knowledge have some value as, or in virtue of being, items of knowledge, or, to put it another way, whether items of knowledge are valuable as such. Similarly for other claims regarding value: that truth is valuable, or that

² And also, of course, a concern to avoid false beliefs.

true belief is valuable, or that gold is valuable, etc. In each case, I understand the claim that V is valuable to mean that individual V things are valuable as such.³

Some might take my initial formulation to entail intrinsic rather than extrinsic value, but that is not my intension. Gold is valuable. I understand this to mean that individual items of gold are valuable as such, but the value of gold is surely extrinsic, not intrinsic. So a claim that V is valuable might assert that V things have either intrinsic or extrinsic value. But the arguments that interest me concern intrinsic value, so I will focus on this rather than extrinsic value. I also do not intend the claim that V is valuable to indicate that the value had by individual V things in virtue of being V must be large or even the most significant value these things have. A piece of cheap costume jewelry might be gold, and have some value as such, but this value will not amount to much. And although King Tut's death mask is gold, and has some value as such, it surely has much more artistic and historical value. I do not even take the claim that V is valuable to entail that each individual V thing has a value that is on the whole positive. The value a thing has, or would have, in virtue of being V can be outweighed or possibly even defeated;⁴ for example, it might be that pleasure in the unjustified suffering of another has no positive value at all, even though pleasure is valuable.⁵

The idea that the universal KNOWLEDGE is valuable is so far-fetched that no one would interpret the claim that knowledge is valuable in this way. But the quote from Williamson may invite a modernized version of the idea that is perhaps not so far-fetched. Since he begins by focusing our attention on the 'ad hoc sprawl' that analyses of 'the concept *knows*' have become and then asks why we should care about '*that*,' it would not be outrageous to take him to be asking why we should care about or value such a *concept* of knowledge. On the other hand, he might be asking why we should value *knowledge*, understood in the way I have just explained, given that analyses of the concept of knowledge have had to become such ad hoc sprawls. I admit there are interesting and important questions about when to care about *concepts*—about the conditions under which particular *concepts* are valuable or not. And the proposal that concepts requiring analyses that are messy, ad hoc sprawls are not valuable seems plausible enough. But I am not concerned with questions about the value of *concepts* in this paper; my focus will instead be on the value of the *things* that fall in the extensions of

³ I would like what I say in this paper to apply to such concepts as 'person' or 'cat.' While it is natural to say, 'knowledge (truth, or true belief) is valuable,' one would never say, 'person (cat) is valuable.' In such cases one is forced to use a plural, thus, 'persons (cats) are valuable.' This makes it quite obvious that in these cases one means to be claiming that individual persons or cats are valuable, not that the concept or any sort of universal or abstract thing is valuable. I mean such cases to fall under my generic formulation, 'V is valuable.'

⁴ Apologies here to Jonathan Dancy (2006) and other particularists.

⁵ Perhaps I could get away with claiming that 'V is valuable' requires only that it is typically the case that individual V things have some value in virtue of being V. The obstacle would be to explain how a thing could be valuable *in virtue of being V* when it is only typically the case that V things are valuable.

concepts. Thus, to provide some mundane examples, my concern is with such questions as whether cats are valuable as such, that is, whether individual cats have some value just in virtue of being cats, not with whether the concept 'cats' is valuable, with whether jade is valuable as such, not with whether the concept 'jade' is valuable, with whether pleasure is valuable as such, not with whether the concept 'pleasure' is valuable, and so on.⁶ If it sometimes seems that I have slipped into talking about the value of concepts, please interpret this as an infelicitous choice of words on my part.⁷

In what follows, I'll first briefly explain the arguments offered by Stich and Kvanvig. (I'll talk about Williamson's argument as well, but my treatment will be brief and will not occupy a separate section.) I'll then provide a general statement of the form of argument I think they are employing. And I'll conclude by presenting my reasons for skepticism regarding this form of argument.

2. STICH ON THE VALUE OF TRUE BELIEF

In chapter 5 of *The Fragmentation of Reason* (Stich 1990), Stephen Stich argues against the claim that true belief is valuable. A significant part of his argument proceeds by examining what he thinks is one of the most promising analyses of true belief: a causal/functional theory. (Even though he is not committed to this analysis being correct, he thinks it illustrates features that will be shared by whatever turns out to be the correct analysis.) One way to get started on an analysis of true belief is to specify an interpretation function that pairs beliefs—understood as 'real psychological states' (1990: 103)—with propositions, truth conditions, contents, or something else 'semantic' (1990: 104). A belief will then be true when the proposition obtains, the conditions are satisfied, etc. The causal/functional theory is so named because it employs a causal/functional interpretation function. One of the things any interpretation function must do is pair up the right things with referring terms such as proper names (1990: 108). A causal theory of reference seems required to do the trick. Stich's discussion of this element of the theory nicely displays the basic tenor of his argument.

The problem, as Stich sees it, is that when we actually identify the kinds of causal chains that link referring terms with their referents, it turns out that the specification of these chains is determined by our intuitions regarding a series of ever more complicated examples and counterexamples, and hence that this

⁶ In an effort to keep things clear, I will continue to employ the highly original practice of using quotes to identify concepts whenever there might be some confusion about whether I mean to be talking about the concept or the things that fall under the concept.

⁷ I'm indebted to Marian David for alerting me that in earlier drafts I was not clear about whether I was concerned with the value of concepts, the things that fall under concepts, or something else such as universals or properties.

specification is extremely complex and idiosyncratic. He describes the situation as follows:

But if, as we have been assuming, it is the causal/functional interpretation function that is sanctioned by intuition, then it is not a particularly simple or natural function. Rather, it is something of a hodgepodge, built from a more or less heterogeneous family of strategies for fixing the reference of terms and another family of strategies for transmitting reference from one speaker to another. (1990: 119)

Stich emphasizes the idiosyncratic nature of this concept of reference, and hence our concept of true belief, by pointing out that a slightly different specification of the causal chains that fix and transmit reference would pair some terms with different referents. This different specification would conflict with some of our intuitive judgments about cases, but we could still use it to define a concept very much like our intuitive concept of reference, and then use the new, not quite intuitive concept of reference to define a concept that is just a little different from our ordinary concept of truth. Indeed, we could define a whole series of such concepts, which Stich labels *TRUTH**, *TRUTH***, ... Some of these truth-like concepts would be very similar to our usual, intuitive concept of truth, but others would be very different from it.⁸

Now comes the part of the argument that interests me here. Given that our concept of truth involves a messy, idiosyncratic concept of reference, Stich questions the sense of valuing beliefs that are true in this ordinary sense.⁹ And of course, given that he has explained how to construct a whole series of truth-like concepts, Stich can press the question by asking why it is ordinary truth that we value rather than one or more of the *TRUTH**s. At the beginning of his inquiry Stich explained that 'without an account of what it is for beliefs to be true, it is all but impossible to think clearly about whether we value having true beliefs' (1990: 104). Now that we have a good start on an account of the nature of true beliefs, he thinks it is clear that there is not any 'simple or natural relation between mental states and truth conditions. It is a jerry-built contraption' (1990: 126).

Stich is clearly offering up several considerations that he hopes will shake our conviction that true belief is valuable.¹⁰ One may be the idea that our intuitive

⁸ Stich explains how to define the various truth-like concepts in Stich 1990: 115–18.

⁹ Stich considers in turn the claims that true beliefs are intrinsically valuable and instrumentally valuable. He admits that, should push come to shove, a person can intrinsically value whatever he or she wants—including truth—whatever the valued object turns out to really be and however it must be analyzed. See Stich 1990: 118. Hence, it is to some extent misleading to characterize him as offering an *argument* for the proposition that true belief lacks intrinsic value. Rather, he is offering considerations he thinks have the power to lead those who begin the discussion intrinsically valuing truth to stop valuing it. That said, I do not believe I do him a very great injustice when I characterize him as offering an argument since it is quite clear he thinks that, once one sees the general shape the correct analysis of true belief must have, the proper response is not to value true belief any longer.

¹⁰ Stich begins his argument against valuing true belief by pointing out that the causal/functional theory is partial and idiosyncratic (Stich 1990: 118), and then addresses each in turn. The problems

notion does not manage to pick out a real kind of thing in the natural world. Another is clearly that we have no reason to prefer beliefs that are true in the intuitive sense over beliefs that are true in one of the TRUE* senses,¹¹ although if he wanted this consideration to bear much argumentative weight, he would face the problem of explaining why truth must be preferable to other things in the neighborhood in order to be valuable.¹² Another factor that clearly bothers him is that the interpretation function is in the end determined by nothing more than commonsense intuitions, which he believes to have a questionable origin (Stich 1990: 120). But I believe one significant consideration Stich wants to advance is that analysis reveals the concept of truth to be a messy ‘hodgepodge’ or ‘jerry-built contraption’—or in the fancy technical terminology suggested by my title: ugly.¹³ Once we understand that our concept of truth is so ugly, we are to conclude that we do not really value true beliefs as such after all.¹⁴

3. KVANVIG ON THE VALUE OF KNOWLEDGE

Kvanvig devotes the majority of *The Value of Knowledge and the Pursuit of Understanding* (Kvanvig 2003) to developing an account of the value of knowledge on which knowledge is more valuable than any proper subset of its constituents. One might quibble about the idea that knowledge, literally, has constituents, but Kvanvig’s idea is clear enough: the constituents of knowledge correspond to the conditions included in a correct analysis of knowledge.¹⁵ Epistemologists may not have constructed any analysis of knowledge that is generally regarded as correct, but nearly all of them would agree that to know a proposition one must believe it and it must be true. So Kvanvig begins with these and, rather oddly, concludes that each has value that it can contribute to knowledge.¹⁶ Epistemologists also agree that knowing requires that one’s true belief be epistemically good or appropriate in some sense. For a long time epistemologists thought that a belief was good in the relevant sense just in case the person

he attributes to the partial nature of the function are not directly relevant to my concerns in this paper, but see Appendix 1 for some discussion.

¹¹ Both these concerns are suggested immediately before the passage I quoted from Stich (1990: 119) above.

¹² Vincent van Gogh did a number of paintings of vases with sunflowers. I may have no ground for preferring one of these, say the first he painted, to the others, but it hardly follows that there is something misguided about my valuing his first sunflower painting.

¹³ I am assuming that if an adequate analysis of a concept is messy, jerry-built, ad hoc, etc.—in short, ugly—we can say that the concept is messy, jerry-built, ad hoc, and, in a word, ugly.

¹⁴ Thanks to Alan Millar for calling my attention to and stressing the significance of some of the other elements of Stich’s case against the value of true belief.

¹⁵ For some quibbling, see DePaul and Grimm (2007).

¹⁶ Kvanvig at 2003: 29–38 argues for the value of belief and for the value of truth at 2003: 38–43. See DePaul and Grimm (2007) for a critical discussion of the claims that truth, full stop, is valuable and that belief, again, full stop, is valuable.

was *justified* in holding it, but there is no longer a consensus regarding this point.¹⁷ Hence, Kvanvig examines a number of candidates for the requisite type of good believing.¹⁸ A major problem that besets some candidates, most notably reliability, is that the value of true belief ‘swamps’ whatever value these candidates might have to contribute to knowledge.¹⁹ Hence, these candidates cannot explain the *superior* value of knowledge as compared with true belief. But Kvanvig does grant that two candidates—subjective justification and virtuous belief formation—can contribute some value to knowledge that goes beyond the value of mere true belief.

At this point Kvanvig’s effort to develop an account of the value of knowledge runs into trouble. As Gettier (1963) and those who have followed him with ever more clever and occasionally complicated (or even bizarre) counterexamples have shown, even a true belief that is subjectively justified and was virtuously formed might fail to be an instance of knowledge. So the analysis of knowledge needs yet another condition, and hence knowledge must have yet another constituent. This constituent must exclude the sorts of possibilities revealed in the myriad Gettier-style counterexamples that epistemologists have cooked up in the last forty odd years. As yet, no one has formulated a condition that epistemologists agree does the job. Indeed, what agreement epistemologists have reached in this area mostly concerns the failure of various efforts to characterize this constituent of knowledge.²⁰

Nevertheless, it is fairly clear, according to Kvanvig, that the root problem behind Gettier examples is ‘luck,’ ‘accidentality,’ or ‘fortuitousness’: in Gettier cases the person may have a justified true belief, but in some sense the person gets things right only by lucky accident (2003: 113). According to this understanding of the origin of Gettier problems, it is so difficult to construct a condition that

¹⁷ I here assume justification should be understood in a traditional, internalist way. If we are willing to allow externalist understandings of justification, e.g. as reliably formed belief, then the way to make the same point would be by saying that there is no longer a consensus regarding the basic characteristics of justification.

¹⁸ To be more specific, Kvanvig is willing to count externalist views as accounts of justification, so he begins by discussing these (2003: 44–52). He then turns to more traditional internalist accounts of justification (2003: 52–75). Then in ch. 4 he examines attempts to explain the epistemic goodness required for knowledge in terms of the exercise of intellectual virtues.

¹⁹ Kvanvig’s discussion of the swamping problem begins at 2003: 45. I raised a similar concern about the major efforts to understand the value of knowledge, without using the term ‘swamping,’ in my 1993: 75–80. I would put the point in this way: If one takes truth to be the only epistemic good and takes justification (or warrant, or whatever one chooses to call the evaluative feature that distinguishes knowledge from mere true belief) only to have instrumental value, as a means to truth (as is overwhelmingly natural for reliabilism), then one has an incoherent view of the value of knowledge. The reason is that once one has a true belief one has attained all the epistemic value it is possible for that belief to have. The fact that the belief was formed in a reliable way cannot add anything. In my 1993 I used an illustration from baseball: nothing is added to the value of a game winning home run by the fact that it was hit by a reliable power hitter. All the value resides in the fact that it is a home run, and it would have the same value if it were hit by a weak singles hitter.

²⁰ Kvanvig considers efforts to answer the Gettier problem and their impact with respect to the value of knowledge in his pivotal fifth chapter.

rules out all and only Gettier-style examples because not all kinds of luck are incompatible with knowing. Any belief results from good fortune in various ways, starting with the believer's good fortune to be alive and capable of forming beliefs at all (2003: 115). Focusing on the question of the value of knowledge rather than its nature, Kvanvig claims that while it may seem obvious that it is good to get things right in a way that is not accidental or lucky, once one realizes the 'labyrinthine,' 'ad hoc and gerrymandered' complexity required in a condition that excludes just that special type of good fortune that is inconsistent with knowledge, it is not at all obvious that we value those things that satisfy such a condition (2003: 116–17).

At this point, Kvanvig digs into the specifics of the post-Gettier literature on the analysis of knowledge. He believes two approaches show some promise of succeeding as analyses, that is, of being immune to intuitive counterexample; these are defeasibility analyses and (so-called) conclusive reasons analyses.²¹ But when he considers what these analyses have to say regarding the value of knowledge, Kvanvig finds the same pattern that emerged from his more general consideration of the Gettier problem and luck. A brief examination of Kvanvig on defeasibility analyses will suffice for my purposes. The root idea of defeasibility analyses is that when one knows one's belief is not merely true, justified, and virtuously formed, it is also *undefeated* in the sense that there is no information that one might acquire that would make it the case that one is no longer justified. Kvanvig happily grants that such immunity to being overturned by further inquiry sounds pretty good.²² The problem is, an analysis of knowledge that requires absolute immunity to defeat is too strong.

Suppose you see someone you take to be Tom Grabit steal a book from the library. You know Tom well; you got a good look at him and what he was doing; so, you are justified in believing that Tom took the book. And he did take the book. But his mother, a woman of impeccable credentials, has testified to the police—some strict library!—that Tom is out of town and his twin brother Tim took the book. So there is a defeater for your belief. If you were to learn of Tom's mother's testimony, you would no longer be justified in believing Tom took the book. If that were the end of the story, you wouldn't know, but this is a *misleading* rather than a *genuine* defeater. Walking out of the police station Tom's mother, who has never before told a lie, breaks down and confesses that she fabricated her story about the twin brother to protect Tom. Her confession defeats the defeater, reinstating, so to speak, your knowledge.²³

²¹ Kvanvig considers defeasibility approaches at 2003: 125–33 and conclusive reasons approaches at 2003: 133–9.

²² He writes, 'At first glance, this approach offers much promise regarding the question of the value of knowledge, for it is valuable to have an epistemic standing that is not defeated by any additional information' (2003: 125).

²³ Kvanvig presents a version of the Grabit case and explains how it forces the defeasibility theorist to distinguish between genuine and misleading defeaters at 2003: 125–6.

You should now be able to predict how Kvanvig will finish his story, so I can skip the details and jump to the conclusion.²⁴ There are many different kinds of defeaters and defeater defeaters and what have you. Efforts to distinguish the genuine defeaters that can undermine knowledge from those that are compatible with knowledge have made progress with respect to avoiding counterexamples, but they have had to become increasingly complex. What follows with respect to the value of knowledge? Here is how Kvanvig concludes his examination of one effort, due to Peter Klein, to draw the required distinction among defeaters. It is, I think, representative.

It is hard to see this distinction as anything more than gerrymandering needed to prevent counterexamples to one's account of knowledge, and it is easy to side with Williamson in remarking 'Why should we care about *that*?' The distinction between these kinds of defeaters tracks no intuitive difference in value, leaving us with an account of the nature of knowledge incapable of helping to explain the value of knowledge. (2003: 129–30)

I would describe the argument here as follows: The distinction drawn between those defeaters that undermine knowledge and those that do not is extremely complicated, ad hoc, gerrymandered—in a word, ugly. The no defeater condition that incorporates this distinction is therefore ugly as well. Hence, there is no reason to value things that satisfy a no defeater condition for knowledge.

Kvanvig's discussion of conclusive reasons approaches has a similar shape. Here the problem is not distinguishing between defeaters, but refining or specifying one or both of the simple counterfactuals, '*If the claim in question were false, one would not believe it*' and '*In similar situations in which one believes the claim, it is true*' (2003: 133), so as to exclude all and only Gettier style counterexamples. It is no surprise that this becomes a messy, complicated business. Thus, Kvanvig believes conclusive reasons approaches run into the same problem as defeasibility approaches when it comes to the value of knowledge (2003: 138).

At the end of his examination of the Gettier problem Kvanvig does not conclude that knowledge lacks any value. Remember, in earlier chapters he found valuable constituents of knowledge, specifically, true belief, subjective justification, and virtuous belief formation. But he does conclude that knowledge lacks any special or distinctive value. All the value knowledge has can be accounted for in terms of the value of virtuous, subjectively justified, true belief. But an essential constituent of knowledge—the one that excludes Gettier-style counterexamples—lacks value, and hence makes no contribution to the value of knowledge. The long series of failed attempts to produce an analysis of knowledge that is immune to Gettier examples gives us ample reason to be confident that, whatever the successful condition is, it will be an ugly sprawl of a thing. But

²⁴ The lengthy discussion of defeasibility analyses I'm skimming over occurs at Kvanvig 2003: 125–33.

then, according to Kvanvig, there is no reason to think that any value attaches to just those things that satisfy such a condition.²⁵

4. THE UGLINESS ARGUMENT

I hope that my presentation of Stich and Kvanvig has already begun to bring the form of argument that concerns me into focus. I will call this form of argument *the ugliness argument*. The argument begins by examining efforts to provide an analysis of some concept, where we take the things to which the concept applies to have some special value. We tend to view truth, or more specifically true beliefs, as valuable, so Stich begins with our efforts to analyze the concept of true belief. We also tend to think that knowledge is valuable, indeed, distinctively valuable, and Kvanvig begins by examining our efforts to analyze knowledge.²⁶ Although we do not yet have an analysis of either true belief or knowledge in hand that clearly succeeds, the efforts to construct such analyses have a clear direction, that is, towards baroque complexity. Stich and Kvanvig therefore take themselves to have compelling evidence for thinking that successful analyses—of true belief and knowledge respectively—will be ugly. Upon consideration of such an ugly analysis, it is not apparent to us that, or why, we should value the things to which the concept applies. The argument concludes that the things to which the concept applies, that is, true beliefs or items of knowledge, are not valuable.

Stich and Kvanvig have had to offer versions of the ugliness argument that are somewhat more complex than they might have been because we do not have in hand successful analyses of true belief or knowledge. If we did have adequate analyses of these concepts in hand, Stich and Kvanvig could have simply constructed their arguments using the correct analyses. So let's examine

²⁵ To be perfectly accurate, Kvanvig concludes his discussion of the Gettier problem as follows: 'If knowledge is valuable, it is so on the basis of factors other than the value of its constituents' (2003: 139). He goes on to consider and reject efforts to understand the value of knowledge without seeing this value as derived from valuable constituents, e.g. by viewing knowledge itself as intrinsically valuable. These discussions are not relevant to my concerns here, but the quote from p. 139 is relevant in another way. While discussing the Gettier problem, Kvanvig frequently writes of explaining the value of knowledge rather than of knowledge being valuable. Does he then hold merely that we cannot explain the value of knowledge in terms of the value of the Gettier-excluding constituent rather than that this constituent is not valuable? The quoted passage suggests that Kvanvig holds the stronger position. The introductory paragraph of ch. 8 provides further confirmation. Kvanvig there claims the value of knowledge 'is exhausted by the value of a subset of its constituents' (2003: 185), which entails that some constituent of knowledge lacks value.

²⁶ My detailed description of Kvanvig's argument in the last section should have made it clear that what he actually argues is that knowledge is not *distinctly* valuable because one essential constituent of knowledge is not valuable. I'm here taking what I think is the small risk of misleading or confusing readers by attributing to him the unqualified claim that knowledge is not valuable for ease of exposition.

the argument they might have constructed, if we had adequate analyses. The viability of the arguments Stich and Kvanvig actually used obviously depend upon the viability of this simpler form of argument.

The argument begins with an adequate analysis of a concept, and for the argument that follows to be philosophically interesting, we must have a fairly strong tendency to value those things that fall within the extension of the concept.

- (1) A is an adequate analysis of concept 'V' (where we have a strong tendency to value V).²⁷

The argument then notes the nature of A.

- (2) A is ugly.

As my use of 'ugly' above should indicate, an analysis is ugly when it is complicated, messy, and seems to be ad hoc, gerrymandered, or jerry-rigged. The next step notes how we regard things satisfying A after reflecting on the nature of this analysis. The idea here is that we are to ignore our antecedent views about the value of V and just reflect on A and the question of whether we would value things that satisfy the conditions of A.

- (3) Because A is ugly, upon due consideration of A, it is not apparent to us that things satisfying the conditions of A are valuable.²⁸

The argument concludes,

- (C) V is not valuable.

The conclusion here is intended to assert that the things in question are not *intrinsically* valuable. Stich and Kvanvig are both clear about this and respectively provide separate arguments regarding the extrinsic value of true belief and knowledge. They do not claim that the ugliness of the analysis of a concept necessarily prevents the things in the extension of the concept from having any of the kinds of causal connections with other good things that would make them instrumentally good.²⁹

²⁷ Remember that I understand 'we have a strong tendency to value V' as equivalent to 'we have a strong tendency to value individual V things as such.' Similarly, I take 'V is not valuable' in (C) to mean 'individual V things are not valuable as such.'

²⁸ A slightly different form of the argument might have a third premise asserting that when we consider A it seems apparent to us that things satisfying the conditions of A are not valuable. I do not believe anything I will say turns upon which way we formulate premise (3). I formulate it as I do because it seems closer to what Stich and Kvanvig actually say and also because it makes a claim that is weaker in the sense that it is more easily true.

²⁹ Stich doesn't think the ugliness of analysis is totally unrelated to the question of extrinsic value. Having admitted that he cannot show that true belief does not contribute to the attainment of anything good, and hence that he cannot show that it is not instrumentally valuable in some way, Stich describes one way the ugliness of the analysis of the concept of true belief might lead one to question whether true beliefs are instrumentally good. As a first step, he points out that we

There is another way of understanding the arguments I'm attributing to Stich and Kvanvig. I have to this point been assuming that the philosophical analyses in question are analyses of *concepts*. What if the targets of the relevant type of analysis are not concepts but things? On this interpretation, Stich does not begin by considering an analysis of our *concept* of true belief, but by considering an analysis of true belief itself, and similarly, Kvanvig's argument begins with an analysis of knowledge itself, not an analysis of our concept of knowledge. Perhaps it will be useful to mark the distinction I want to draw here by restricting the term 'analysis' to targets that are concepts and using either 'account' or 'theory' when the target is some entity that is not a concept.

We tend to assume that the concepts of things that we employ at any given time and place are accurate in the sense that an analysis of our concept (e.g. of cats or cancer) would correspond with a correct theory (of cats or cancer). It is therefore easy to ignore the distinction between analyses of concepts and what I am calling theories of things. But it is clear that the distinction holds at least in many cases, as is shown by the fact that our concepts of many ordinary natural kinds changed over time as we came to learn more and more about the relevant things, and sometimes even learned things that were incompatible with our older concepts of these things.

We could, therefore, formulate as follows what I will call the *realist* version of the ugliness argument. I will refer to the version of the ugliness argument I first presented as the *conceptual* version.

- (1_R) T is an adequate account or theory of V (where we have a strong tendency to value V).
- (2_R) T is ugly.
- (3_R) Because T is ugly, upon due consideration of T, it is not apparent to us that things satisfying the conditions of T are valuable.
- (C) V is not valuable.

If we think of ordinary things, such as salt or tomatoes, we will need to understand 'adequate' in (1_R) differently from the way in which we understood it

need to think about what a thing that is supposed to be instrumentally good is to be compared with. Something might produce some good results, and in particular much better results than its complement, but what if there is something else in the neighborhood that produces much better results? Stich then applies this consideration to truth, falsity, and the various truth-like concepts he has identified. It may well be that true beliefs are more valuable, instrumentally, than false beliefs. But is there any reason to think that true beliefs are better, instrumentally, than beliefs that are 'true' in any of the other senses? Once we understand the ugly nature of the causal/functional interpretation function, how can we expect that beliefs that are true in the ordinary sense are going to lead to better things than beliefs that are 'true' in any of the other senses Stich characterized? He offers this argument on pp. 121–2. The argument does, however, seem vulnerable to the concern I raised above: the claim that something is good, intrinsically or instrumentally, is not a claim that it is better than anything else, even anything else in the near neighborhood.

as used in (1). The primary criterion for the adequacy of a philosophical analysis is immunity to intuitive counterexample.³⁰ Immunity to intuitive counterexample *might* somehow be involved in the adequacy of theories of ordinary things—after all, scientists do sometimes use thought experiments—but such immunity certainly is not the primary or even a very important condition for the adequacy of theories. For one thing, it is much more important that such theories be able to account for all the relevant empirical observations than that they be immune to intuitive counterexample. Indeed, such theories not uncommonly run counter to what we might intuitively expect. In the end, I would be inclined to say that theories of ordinary things would not be adequate unless they were true, or at least approximately true, with respect to those ordinary things. Not so for an adequate analysis of a concept, as is illustrated by the possibility of faulty or inadequate concepts. If a group of people have some deep but mistaken belief regarding the fundamental nature of Xs, this mistake could very well infect their concept 'X.' In such a case, an adequate analysis of their concept 'X' would include the mistake, but an adequate theory of Xs obviously would not.³¹

The fact that the adequacy conditions for analyses are weaker than those for theories provides one reason for favoring the realistic over the conceptual version of the ugliness argument. If the analysis of our concept 'V' is ugly, and this leads us to begin doubting the value of V, there is always the possibility that the problem lies with our concept 'V' rather than the V things themselves. It may be that an adequate theory of Vs would not be at all ugly, and that reflection on such a theory would not leave us wondering why anyone would value such things. There is no similar, easy retrenchment maneuver available when faced with a realistic version of the ugliness argument.

I am nevertheless going to focus on the conceptual version of the ugliness argument. The main reason for this is that in the cases of interest to us here there is no consensus about how to draw, or even if one can draw, an interesting distinction between an adequate analysis of the concept and an adequate theory of the thing or kind of thing in question. For ordinary things like tigers, water, or cell phones, it is easy enough to see how our ordinary concepts of these things could come apart from the reality, and hence how adequate analyses might not match adequate theories. But what about the concepts that interest philosophers, concepts such as knowledge or true belief? We certainly cannot distinguish adequate analyses from adequate theories in these cases by saying that adequate analyses must only be immune to intuitive counterexample while adequate theories must account for empirical observations. When working on the concepts that specially interest philosophers, it seems that all the significant

³⁰ We certainly would want to require that an adequate analysis not be circular, and we may require other things as well.

³¹ There are difficult issues regarding when a concept 'X' is so far off the mark that the proper thing to say is that there simply are no Xs rather than that there are Xs, but our concept 'X' is inadequate and should be replaced with a successor concept.

'data' is supplied by intuitive judgments. It may still be possible to draw an interesting distinction between adequate analyses and adequate theories in such cases, for example, by narrowly specifying the intuitive judgments relevant to analyses while allowing a wider range of considerations (still largely intuitive) to impinge upon theories. But such a distinction will not be nearly so sharp as it seems to be in the case of ordinary things and concepts. In the cases that interest us, the activity of developing an adequate theory will be continuous with the process of analysis. As a result, I think we can safely assume, at least for the purposes of this paper, that adequate philosophical analyses of such concepts as 'knowledge' and 'true belief' will also supply adequate theories of knowledge and true belief.³²

One passage from Stich where he explains how his argument is supposed to work suggests he might agree:

there is another kind of consideration that might be efficacious in persuading someone that she should not, or does not really, accord intrinsic value to the having of true beliefs. . . . we can try to be sure that she sees clearly the real nature of what she values—that she appreciates what having true beliefs comes to. (1990: 118)

As odd as it seems, given that he is an arch skeptic when it comes to the value of analyses that do no more than correctly account for intuitive judgments, Stich here seems to assume that an adequate analysis of a concept reveals the true nature of the things in question, at least for the kinds of concepts of special interest to philosophers, that is, in my terms, that an adequate analysis of the concept also supplies a true theory of the thing. One might think it would be most charitable to take Stich to be describing a way of thinking about what an analysis of true belief provides that is amenable to those who value true beliefs, rather than speaking for himself, and then, in effect, showing them that on their own terms true belief is not something it makes sense to value. But I doubt the charity of attributing such an argument to Stich. If one thinks a philosophical analysis does not reveal the real nature of true belief, it would seem somewhat dishonest to try to shake someone's confidence about the value of true belief by pointing to an ugly analysis and telling the person that this is what true belief really is.

I shall now consider two simplified versions of the ugliness argument. The first takes the ugliness of A to provide direct support for the conclusion that V is not valuable and drops the step where we note how we evaluate things that

³² I feel compelled to note that I have fairly serious reservations about what I have said in this paragraph. Deep down I think it is possible to distinguish between analyses and theories even for the kinds of things of most interest to philosophers. I also believe that philosophers are typically interested in theories rather than analyses, in spite of the way they tend to talk. But in my judgment it is best to avoid these complications for the purposes of this paper. For a couple of stabs at articulating reasons for doubting that philosophers are much concerned with analyses, see my 2006 and 2000.

satisfy A. The second version of the argument understands the consideration of the ugliness of an analysis as a heuristic device to help us see that we do not regard things that satisfy A as valuable. According to this understanding, all the argumentative work is actually being done by our failure, upon reflection, to regard things that satisfy A as valuable. Here then are these two versions of the argument:

The Pure Ugliness version

- (1) A is an adequate analysis of concept 'V' (where we have a strong tendency to value V).
- (2) A is ugly.
- (C) V is not valuable.

The Reflective Evaluation version

- (1) A is an adequate analysis of concept 'V' (where we have a strong tendency to value V).
- (3*) Upon due consideration of A, it is not apparent to us that things satisfying the conditions of A are valuable.
- (C) V is not valuable.

I think we can quickly dismiss the pure ugliness version. The ugliness of the analysis of a concept simply is not incompatible with the things in the extension of the concept having value. I am not alone in valuing the Costa Rican cloud forest ecosystem. An accurate account of our concept of the cloud forest ecosystem would surely be exceedingly complex, disjunctive, arbitrary, etc.—in short, quite ugly. Indeed, I'm pretty sure that the most complicated analyses of knowledge will look almost elegantly simple by comparison. Yet I feel no inclination to therefore retreat from my view that the cloud forest ecosystem is valuable. This example illustrates a second important problem with the ugliness version of the argument: what should we do if we encountered a case where reflection upon the ugly analysis, and even the ugliness of the analysis, left us convinced that things satisfying the analysis *are* valuable? It is doubtful that in such a case it would make sense to draw the conclusion dictated by the ugliness version of the argument.

A final point, one might object that the ugliness argument is supposed to apply to a philosophical analysis, but what I use in my example is a scientific theory or account. I'm sorry if I gave that impression, but that is not how I intended the example to be taken. I meant to be considering an analysis of *my concept* of the cloud forest ecosystem. Even given my mere amateur's understanding, I know enough about what defines the cloud forest ecosystem that any plausible analysis of my concept will be plenty ugly. And we should not forget that examples

like this abound. Psychological research makes it quite clear that our concepts of ordinary things such as birds or whales are such that either they cannot be analyzed with necessary and sufficient conditions at all or that such analyses will be extremely ugly. Nevertheless, I for one am happy to say I value, for example, whales, and I'm betting I'm not alone. Hence, as I said above, I think we can dismiss the ugliness version of the argument.

If at this point one objects that I am here applying the argument to ordinary and natural kind concepts rather than the kinds of concepts of special interest to philosophers, I reply as follows. While the application of the ugliness argument to concepts such as 'knowledge' and 'true belief' might be of special interest to philosophers, I did not formulate the argument so that it is restricted to such concepts. And I believe it would be improper to do so without some sort of additional argument, on the part of those who use this form of argument, for thinking that there is some special difference between the things and concepts in question. If this form of argument does not work in general, that is, when tested against ordinary cases, why should we assume it to be valid when it is employed to show that things like knowledge and true belief lack value?

Can we say, then, that the reflective evaluation version captures all the argumentative force of the original ugliness argument? I'm inclined to say it does. Just think about what happens when our reflective evaluation does not march in stride with ugliness. I considered one such case above, where, if I am right, we are inclined to value the things that satisfy an ugly analysis. In such cases there does not seem to be any reason to conclude that the things in the extension of the concept lack value. It is our reflective evaluation that determines what we should conclude. The relevant case, however, is one where an analysis is not at all ugly, but upon reflection we are not inclined to value the things that satisfy the analysis. In such a case, I do not see why a proponent of the form of argument we are considering should not be just as willing to conclude that things in the extension of the concept lack value, as in the case where the reflective evaluation is in part a response to the ugliness of the analysis. I'm having a hard time thinking of a good example where we might have been initially inclined to view the thing under analysis as valuable, but perhaps the following will do. We can give a beautifully elegant analysis of the concept of a pyramid. After all, it is a geometric object. But upon consideration of this lovely analysis I am not at all inclined to value pyramids. Hence, I am happy to conclude that pyramids have no value as such—and to do so in spite of the fact that there are some people who do think things having this shape have some special value.³³

I shall therefore proceed assuming the reflective evaluation version captures the full argumentative force of the arguments I've attributed to Stich and Kvanvig (and take Williamson to turn on its head). I realize that they do emphasize

³³ The claim I mean to make here is of course that pyramids have no value as such. This is consistent with certain pyramids, e.g. the great pyramids of Egypt, having plenty of value.

ugliness in their discussions, but I think we can safely regard this as merely a means to evoke the flat evaluative assessment that is actually doing all the work.³⁴ They would be just as happy to evoke that assessment in some other way and proceed to the conclusion. In the end, I am not sure much turns on whether we focus on the original ugliness argument or the reflective evaluation version. I believe that the critical points I make apply equally well to both versions.

5. CRITICAL EVALUATION

The first instinct of a proponent of an analysis attacked by the original ugliness argument or the pure ugliness version would likely be to deny premise (2), attempting to call our attention to all the lovely but unnoticed features of the analysis. I'll leave such responses to those who have crafted the relevant analyses with such loving care. If I'm right and the real force of the argument is captured by the reflective evaluation version, such a defense is not really to the point anyway, unless by focusing on the lovely features of the analysis we are led to make a different reflective evaluation. How then might one respond to the reflective evaluation version of the argument? Rather than considering whether it might be better to deny premise (1) or premise (3*), I would instead like to question a presumption of the transition from (1) and (3*) to (C).

Particularly given that we are so clearly fallible when it comes to making evaluations, it would be crazily optimistic to suppose that if we recognize something is valuable, then we will recognize it to be valuable under any other description we consider. The argument obviously does not presuppose this overly optimistic idea, but it does seem to presuppose a more restricted form of it, specifically, that our ability to recognize value is preserved by analysis. Here's a first crack at the presupposition that we recognize value under analysis.

(RVUA) If A is an adequate analysis of concept 'V' and person S recognizes that V is valuable (in the sense of 'recognize' that implies truth), then upon due consideration of A it will be apparent to S that things satisfying the conditions of A are valuable.³⁵

³⁴ I feel confident that proceeding this way is not unfair to Kvanvig or to the argument Williamson makes in the passage I quoted near the start of this paper. But in Stich's case I should qualify what I have said. He clearly tries to press a number of considerations against valuing true belief. I think the reflective assessment argument captures one of these, but there may be others in which the brute ugliness of the analysis does play a more significant role.

³⁵ It may have been possible to simplify the discussion by formulating the principle in this way: If A is an adequate analysis of 'V,' and V is (V things are) valuable, then upon consideration of A it will be apparent to us that things satisfying A are valuable. But this seems to imply that if we have an adequate analysis of a valuable thing, we will recognize that thing to be valuable. I did not want to exclude the possibility that there are intrinsically valuable things that we do not recognize to be valuable, and that in such cases we might not recognize that things satisfying an adequate analysis

If (RVUA) were true, it would underwrite the transition from (1) and (3*) to (C). (1) tells us we have an adequate analysis of V, but (3*) tells us that when we reflect upon the analysis it is not clear to us that things satisfying the conditions of the analysis are valuable. Hence, according to (RVUA), we could not really be recognizing that V is valuable. Of course, in the cases that interest us it has always *seemed* obvious to us that V is valuable, so the only reason we would fail actually to *recognize* that V is valuable is if V is not valuable. And so we reach (C).

Let's begin with a clarification that should avoid one potential objection to (RVUA). If A is really complicated or formulated using terminology that S is not familiar with, then S will not really understand A. And if S does not understand A, it would be very odd for S to value things that satisfy the conditions of A regardless of whether A provides an adequate analysis of 'V.' But S's failure to value things that satisfy A because of an inability to understand A surely shouldn't be taken to provide a reason for thinking that V is not valuable. I could avoid this counterexample by revising (RVUA) to require that S understand A, but I shall instead simply understand 'due consideration' to imply understanding. So as I understand (RVUA) it is not possible for a person to give an analysis due consideration unless he or she truly understands the analysis.

Even when strengthened in this way, I find it hard to see why one should accept (RVUA). The first thing that strikes me about it is that it is an *epistemic* principle rather than what we might call a material principle. It does not tell us that if A is an analysis of 'V' and something holds of V, then it will also hold of the things that satisfy the conditions of A. (RVUA) tells us what we can count on any person, S, who considers the matter to believe about the value of things that satisfy the conditions of A given only that S values V and that A analyzes 'V.'³⁶ I tend to be very skeptical about such principles. I am always impressed by the astounding capacity people have to believe things, or fail to believe things, in all sorts of circumstances. I suppose that no one can believe a proposition of the form 'P and not P' while self-consciously, occurrently considering the proposition. But stray very far from something like this, and I have strong doubts. I just do not think we can say much about what persons necessarily will, or will not, believe, recognize or know about one thing given only that they believe, recognize, or know something else and that some logical relation or necessary connection holds between the two things.

of the thing are valuable either. Thus my weaker formulation, which tells us only that when we recognize that something is valuable on its own, we will recognize that things satisfying a correct analysis are valuable as well.

³⁶ I am here assuming that if it is apparent to S that P, then S believes that P. Presumably for P to be apparent to S, it would also be necessary that S's belief that P have some fairly strong positive epistemic status. It is likely that more than this is necessary, but at a minimum it cannot be apparent to someone that something is the case unless that person believes that it is the case.

The second thing to note about (RVUA) is that it is a type of *closure principle*.³⁷ It tells us, in effect, that evaluation is closed under adequate analysis. It is structurally similar to the more familiar principles that tell us, for example, that belief or knowledge is closed under logical implication. As a rule I'm not a big fan of closure principles, and (RVUA) is no exception. Perhaps if concepts were transparent, so that anyone possessing 'V' would immediately know that A is an adequate analysis of 'V,' then something like (RVUA) might hold true. In this case, upon considering A the person would immediately recognize that it is an analysis of 'V,' so, given that the person recognizes that V is valuable, it stands to reason that it would be apparent to such a person that things satisfying A, which the person would immediately recognize to be V things, are valuable.³⁸ But of course we do not require so much of analyses. Instead, the main requirements of analyses is that they be immune to intuitive counterexample and that they not be circular. Hence, it is all too possible even for a person who understands A to fail to recognize that A provides an analysis of 'V.' (The person might, for example, need to work through a fairly large number of examples and potential counterexamples in order to see that A is an adequate analysis.) It is not at all apparent why we should think that anything as strong as (RVUA) must hold true for any analysis that meets only these rather minimal conditions on adequate analyses.

We could build more into the antecedent of (RVUA). Here's a first step:

(RVUA₁) If A is an adequate analysis of 'V,' S knows that A is an adequate analysis of 'V,' and S recognizes that V is valuable, then upon due consideration of A it will be apparent to S that things satisfying the conditions of A are valuable.

But this isn't going to be enough; S might not be thinking about the fact that A is an analysis of 'V' when S considers A, or S might be tired or distracted when considering A, or S might just be irrational. Under these circumstances it might not be apparent to S that the things that satisfy the conditions of A are valuable. And so we might be tempted to further beef up the antecedent of (RVUA₁):

(RVUA₂) If A is an adequate analysis of 'V,' S knows that A is an adequate analysis of 'V,' S is currently cognizant of the fact that A is a correct analysis of 'V,' S recognizes that V is valuable and S is fully rational, then upon due consideration of A it will be apparent to S that things satisfying the conditions of A are valuable.

³⁷ Thanks to Fritz Warfield for pointing this out.

³⁸ Indeed, this may put the matter in too round about a way: If 'V' is transparent, whenever using 'V' a person would also have A in mind. Hence, there seems to be no distance between valuing V things and valuing things meeting the conditions of A.

We may need to chisel at (RVUA₂) some more to avoid all possible counter-examples. (Alternatively, it may not be possible to get this sort of principle right.) But let's grant that (RVUA₂) is correct for the sake of argument. There are still problems with using a principle like this to undergird the reflective evaluation version of the ugliness argument. For one thing, the more we beef up the antecedent with idealizations, the greater the chances that when a person fails to recognize the value of the things that satisfy A, this will be because the person falls short of all the idealizations and not because he or she doesn't really recognize that V is valuable. I know, we are implicitly assuming that the relevant S's here are all analytic philosophers, not factory workers, delivery men, migrant farm workers, professors of English Literature, or my poor old grandmother (who could cook mean Italian food, but didn't know beans about philosophy). And of course we philosophers, especially of the analytic stripe, are uniformly clear-headed, clever, rational, attentive, undistracted, and so on. But I still think we need to worry about the possibility that we fall short of all the idealizations built into (RVUA₂) or whatever successor principle finally gets things exactly right.

There is a more important problem all the versions of (RVUA) share. In the important cases a certain kind of conflict arises: we intuitively find something valuable considered on its own terms, but when we focus our attention on the details of an analysis of the corresponding concept, we do not intuitively value the things that satisfy the analysis. (RVUA) and similar principles tell us, in effect, that in such cases of conflict, our intuitive evaluation regarding the analysis, or more precisely, our lack of a positive intuitive evaluation regarding the things that satisfy it, must always take precedence over the conflicting intuitive evaluation. But why should we think that we should always privilege the intuitive evaluations we form while focusing on the analysis?

One possible answer—the best I can come up with—is suggested by the quote from Stich I used above when explaining why I was going to explore the conceptual rather than the realistic version of the ugliness argument: an analysis reveals the true, deep, essential nature of a thing. When we consider an analysis, we come to understand the true nature of the thing analyzed. Hence, we should trust the intuitive evaluations we make while focused on the analysis over those we make regarding the unanalyzed thing. I personally find it mighty easy to doubt whether adequate philosophical analyses always manage to reveal such true, deep natures, and I am a fairly traditional philosopher who is not inclined to criticize the project of constructing philosophical analyses. But even putting aside these doubts, as is only fair given that I decided to work with the conceptual rather than the realistic version of the ugliness argument, I am still not convinced that we should always favor the evaluations we make when focusing on the analysis.

One reason I am not convinced is that there is no guarantee that an analysis—even an analysis revealing a thing's true nature—will operate on the

same level as our ordinary thinking. Indeed, one might expect that an analysis that reveals a thing's true, deep nature will not operate on an ordinary level. A correct analysis of this type need not employ the same concepts we ordinarily employ. Even supposing that we fully understand the analysis, it might well employ concepts that are quite foreign to the ordinary modes of thought within which evaluations find their home. But then we should not expect that what we find obvious about things when we think about them in our ordinary way, employing our ordinary concepts, will also strike us as obvious, or even believable, when things are described using the concepts necessary for an analysis revealing a thing's true nature. Hence, we should not read too much into the fact that we do not intuitively value the things that satisfy the conditions of an analysis. So there's one reason to suspect that such evaluations will be less reliable than the ones we ordinarily make when thinking about things in familiar ways.³⁹

A second reason for doubting the evaluations we make regarding analyses that reveal deep natures is that such analyses have what I might call a metaphysical rather than an evaluative focus. The aim of such an analysis is to find conditions satisfied by all and only the relevant things in all possible worlds. I see no reason for thinking that an analysis constructed with this aim will reveal all factors that are *evaluatively* significant. Why should we presume that the features that make a thing what it is metaphysically are the features in virtue of which that thing has intrinsic value? If these features are not the same, then when we focus our attention on the features that make a thing what it is, we will not be attending to the features in virtue of which it has value.

Finally, I'm betting that it will become increasingly easy to find counterexamples to (RVUA)-type principles as good analyses of various things are produced. For example, I value pleasure. I fully expect that when a good philosophical analysis of the concept of 'pleasure' that reveals its deep nature is produced, it will not be intuitively obvious that the things that satisfy the analysis are valuable. This does not shake my confidence in the value of pleasure one bit. Similarly for a host of other things I am absolutely confident are valuable, such as affection, the experience of listening to fine music, or human beings.⁴⁰

³⁹ At this point Alan Millar objects that while it may be true that the 'analyses' of natural kind concepts, which often involve scientific theorizing, employ concepts that are not ordinary and familiar, this is less obviously so in the case of traditional philosophical analyses. But I think the point does hold for philosophical analyses as well. The analyses of knowledge Kvanvig offers for our consideration provide excellent illustrations. The concepts of misleading and genuine defeaters, especially when they are fully elaborated, are neither familiar nor ordinary. They are technical concepts constructed by philosophers for a particular purpose. The same thing is true of the counterfactuals that would have to be formulated to make the conclusive reasons approach work.

⁴⁰ An anonymous referee suggests yet another possible version of the ugliness argument that highlights the role of explanation. See Appendix 2 for consideration of this version.

6. THE OPEN QUESTION ARGUMENT

While its application to truth and knowledge may be novel, a version of the argument based on an (RVUA)-type principle, indeed, a limit case of this argument, played a huge role in the history of twentieth-century metaethics. The argument in question is G. E. Moore's (1903) infamous open question argument against naturalistic definitions of good.⁴¹ Moore claimed that no naturalistic definition, N, proposed for 'good' can be correct because the question of whether a thing that is N is good will always be open. He held that if N really defined 'good,' then it would be obvious that a thing that is N is good. The question of whether an N thing really is good would necessarily be closed. I think I can safely use the simplest version of (RVUA) for the discussion in this section. The open question argument is a limit case of the argument grounded on (RVUA) because its crucial premise simply substitutes 'good' for 'V' in the antecedent of this principle and 'good' for 'valuable' throughout, yielding the following principle:

- (RVUA_M) If A is an adequate analysis of the concept 'good' and person S recognizes that good things are good, then upon due consideration of A it will be apparent to S that things satisfying the conditions of A are good.

As a result of these substitutions the second conjunct of the antecedent will always be satisfied.

Of course, Moore is not out to conclude that good is not in fact good, so his argument does not parallel the arguments offered by Stich and Kvanvig. Rather he turns the argument on its head in the way that Williamson does to conclude that good cannot be analyzed, at least naturalistically, just as Williamson concludes that knowledge cannot be analyzed.

The fact that Moore's open question argument also turns out to be grounded in (RVUA) might have been thought to bolster the plausibility of the reflective evaluation version of the argument—but only immediately after Moore presented his argument. The subsequent history of ethics has not been kind to the open question argument. It is widely agreed that the argument fails, most obviously because Moore relies upon an understanding of *definition* that is much too simplistic. Moore assumed, in effect, that what we know or believe about a concept will immediately be transferred over to a definition of the concept. His argument presupposes that upon considering a thing we see that a definition of a concept applies to, we will immediately take to be true of that thing whatever we

⁴¹ For a fine discussion of the role the open question argument has played in twentieth-century metaethics see Darwall et al. 1992.

immediately take to be true of things we can see that the original concept applies to directly. And this just is not true.⁴²

There are, of course, differences between the idea behind the open question argument and the argument based on (RVUA), but they are not significant. One obvious difference is that Moore constructed his argument in terms of *definitions* while (RVUA) is formulated in terms of *analyses*. But if this is not merely a stylistic variation, the connection between concepts required for a definition is even closer and more intimate than what is required for an analysis. Hence, (RVUA) should be even more dubious than the similar assumptions Moore made regarding definitions. Another difference is the one I've already mentioned. Moore's argument assumes that if a certain strong necessary connection holds between the concept of goodness and some other complex of concepts, then it should be obvious that the things to which that complex of concepts applies are good. (RVUA) in fact extends this idea. It says that if that strong necessary connection holds between two concepts (where one may be complex), and it is obvious to us that the things to which one of the concepts applies are valuable, then it will be obvious to us that the things to which the other applies are valuable. Here again, the difference works to the detriment of (RVUA). If Moore's assumption isn't true, and we cannot just assume that we will recognize the things to which a correct definition of good applies as good, why would we expect (RVUA) to hold true?

I therefore reject the form of argument being employed by Stich and Kvanvig, and also by Williamson. I see no reason at all for thinking that the fact that the analysis of something is ugly, and that we therefore do not immediately see that things satisfying that analysis are valuable, entails that the things falling in the extension of the concept are not valuable. But I do not think this is the end of the matter. The passage from Hume with which I began this paper continues:

An anatomist, however, is admirably fitted to give advice to a painter; and 'tis even impracticable to excel in the latter art, without the assistance of the former. We must have an exact knowledge of the parts, their situation and connexion, before we can design with any elegance or correctness. And thus the most abstract speculations concerning human nature, however cold and unentertaining, become subservient to practical morality; and may render this latter science more correct in its precepts, and more persuasive in its exhortations.⁴³

Hume is surely right here, as is Stich, when they suggest that our judgments regarding the intrinsic values of things are best made when informed with a clear

⁴² Kvanvig recognizes the failure of the open question argument at 2003: 77. But he only explicitly mentions that this failure shows that there can be necessary truths that are not recognizable a priori. He does not consider that the argument he offers against thinking that knowledge is good might be similar to the open question argument, and thus subject to similar objections.

⁴³ Hume (1740) bk. 3, pt. 3, last paragraph. Thanks to Geoff Sayre McCord for alerting me to this passage.

understanding of the real natures of these things. So my point here is not that there is nothing to be gained by reflecting upon adequate analyses of our concepts when we are attempting to decide questions regarding the value of the things that fall under those concepts. My point is merely that if we fail to recognize that things satisfying the conditions of such an analysis are valuable, it does not immediately follow from this fact alone that those things are not intrinsically valuable as things falling under that concept. Even if it is correct, this negative conclusion obviously falls well short of being the last word on the issues I've been considering. The obvious next question to consider is how reflection on adequate analyses, or for that matter theories, should factor into our views regarding what it makes sense to value.⁴⁴

APPENDIX 1: STICH ON VALUING TRUE BELIEF AND VALUING OTHER COGNITIVE STATES

Stephen Stich points out that the causal/functional theory's interpretation function is partial in that it only assigns truth conditions to certain mental states. The states that are not assigned truth conditions cannot be evaluated as true or false. Given this fact, what is supposed to be the problem with valuing true beliefs? Stich writes,

our current cognitive processes are a tiny island in a vast unexplored computational space, a space that we may reasonably suppose to contain riches beyond imagining. But almost all that space is beyond the reach of the causal/functional interpretation function; it is a domain in which there is neither truth nor falsity. Those who would accord intrinsic value to the holding of true beliefs may well be reluctant to explore that vast space and will resist adopting what may be found, since we know in advance that it contains no true beliefs. But theirs is a profoundly conservative normative stand. For what they value in the end products of cognitions must be semantically interpretable, and what is semantically interpretable cannot depart too radically from current patterns of reasoning or from familiar ways of causally tying mental states to extramental reality. To value true belief is to resolve that in matters cognitive, one will not venture very far from where we are now. (Stich 1990: 119)

I can see no reason to accept Stich's point. It is as if he thinks intrinsically valuing one thing prevents one from also valuing other things. I value the experience of appreciating

⁴⁴ This paper is a much expanded version of a section I wrote for DePaul and Grimm 2007. I'm indebted to Stephen Grimm for discussions of these issues and also for his comments on an earlier version of this paper. I am also indebted to Marian David and Fritz Warfield for their comments on an earlier version of this paper. I was lucky enough to have the opportunity to present a version of this paper as part of the Kline Workshop at the University of Missouri. The comments from the members of the philosophy department there, the other workshop participants, and especially Jennifer Lackey who commented on my paper, were very useful to me. Finally, the comments of one of the editors of this volume, Alan Millar, were a great help, as were the comments from two anonymous referees for OUP. I should thank one of these referees in particular, who chastized me about my wordiness. This led me to shorten and simplify many sentences, but I'm sure a better stylist could do better still.

various sorts of music. Thus, I value the cognitive processes that lead to the psychological states that are such experiences. If I cannot now process a certain type of music, such as classical Indian music or bebop jazz, so as to lead to the relevant kind of experience, I would welcome altering the way I process music so that I could produce the relevant psychological states that are now inaccessible to me. But all the cognitive processes leading to experiences of appreciating music occupy a small portion of a vast computational space. The cognitive processes involved in the production of the experiences of enjoying poetry or fine food lie in other portions of computational space. Does the fact that I value the experience of appreciating music make it more likely that I will not also explore these other regions of computational space? Perhaps to some extent, since I only have so much time and energy. But my intrinsically valuing the appreciation of music certainly does not prevent me from intrinsically valuing many other kinds of psychological states and being interested in experimenting with the cognitive processes that lead to these other kinds of states. I think the situation I have described is precisely parallel to that described by Stich. Indeed, I do not see why the psychological states and cognitive processes I've identified do not in fact occupy parts of the computational space Stich describes. So that I could have made my point by asking why we should think that intrinsically valuing truth makes a person less likely to be interested in exploring the appreciation of music, poetry, fine food, or a host of other things that are the result of cognitive or computational processes.

APPENDIX 2: EXPLANATORY VERSIONS OF THE UGLINESS ARGUMENT

The *explanatory* version of the ugliness argument runs as follows:

- (1) A is an adequate analysis of 'V' (where we have a strong tendency to value V).
- (2) A is ugly.
- (3_E) Because A is ugly, it does not explain why the things that satisfy the conditions of A are valuable.
- (C) V is not valuable.

Various passages in Kvanvig as well as the brief statement of the argument by Williamson quoted above might naturally be interpreted as presenting this version of the argument.

Considerations similar to those favoring the reflective evaluation version of the original argument over the pure ugliness version support the same sort of simplification here. The argumentative force of the explanatory version of the ugliness argument seems to derive from the claim that A does not explain why things satisfying the conditions of A have value. Some analyses might fail to provide such an explanation because they are ugly, but presumably this is not the only reason an analysis could fail to explain value. It is hard to imagine someone convinced by the explanatory version of the ugliness argument not being equally persuaded that V is not valuable by showing that a correct analysis, A, of 'V' fails to explain the value of the things that satisfy A's conditions for some reason other than the ugliness of A. On the other hand, proponents of the evaluative version of the argument probably would not be inclined to deny that V is valuable in a case where A manages to explain the value of the things that satisfy it in spite of being ugly. Hence, we

can focus on what I'll call *the pure explanatory version* of the argument, which takes the focus on ugliness as only one of a number of possible reasons why an analysis might fail to explain value.

- (1) A is an adequate analysis of 'V' (where we have a strong tendency to value V).
- (3_E*) A does not explain why the things that satisfy the conditions of A are valuable.
- (C) V is not valuable.

I do not believe we should find this version of the argument any more convincing than the other versions I have considered. It has flaws that are analogous to those of the other arguments. I'll briefly consider just one. Let's begin by granting that an adequate philosophical analysis reveals the real, deep nature of a thing. One might then argue as follows: Since the evaluative properties V things have as such supervene on their non-evaluative properties, and all the non-evaluative properties that could be relevant to the intrinsic value V things have *as such* are ultimately determined by whatever it is that makes them V things, whatever intrinsic value V things have *as such* ultimately supervenes on those properties revealed by a correct philosophical analysis. Hence, since the properties that figure in a correct philosophical analysis of 'V' determine the intrinsic value V things have as such, a correct philosophical analysis of 'V' must explain the value V things have as such. While this may seem like a plausible line of reasoning, I think it should also be apparent where the problem lies: explaining and determining are not the same thing. For determination we only need the right sort of necessary connection. For explanation we need some understanding. It seems possible to realize that there is a necessary connection between two things without understanding why that connection holds. If this is indeed possible, then a correct philosophical analysis of 'V' could fail to yield a satisfactory explanation of why V things are valuable.

REFERENCES

- Dancy, J. (2006). *Ethics without Principles*. Oxford: Oxford University Press.
- Darwall, S., Gibbard, A., and Railton, P. (1992). 'Toward *Fin de Siècle* Ethics: Some Trends,' *Philosophical Review*, 101: 115–89.
- DePaul, M. (1993). *Balance and Refinement: Beyond Coherence Methods of Moral Inquiry*. London: Routledge.
- (2000). 'Linguistics is Not a Good Model for Philosophy,' *Spindel Conference 1999, The Role of the Empirical (and of the a priori) in Epistemology, Supplement to the Southern Journal of Philosophy*, 38: 113–20.
- (2006). 'Intuitions in Moral Inquiry,' in D. Copp (ed.), *The Oxford Handbook of Ethical Theory*. Oxford: Oxford University Press.
- and Grimm, S. (2007). 'Critical Study of *The Value of Knowledge and the Pursuit of Understanding* by Jonathan Kvanvig,' *Philosophy and Phenomenological Research*, 74: 498–513.
- Gettier, E. (1963). 'Is Justified True Belief Knowledge?' *Analysis*, 23: 121–3.
- Hume, D. (1740). *A Treatise of Human Nature*. London: Thomas Longman (McMaster University Archive for the History of Economic Thought <<http://socserv2.socsci.mcmaster.ca/~econ/ugcm/3ll3/hume/treat.html>>).

- Kvanvig, J. (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge, MA: Cambridge University Press.
- Moore, G. E. (1903). *Principia Ethica*. Cambridge: Cambridge University Press.
- Stich, S. (1990). *The Fragmentation of Reason*. Cambridge, MA: Bradford Books.
- Williamson, T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.

6

The Goods and the Motivation of Believing

Ward E. Jones

1. EXPLORATION AND ITS GOODS

As epistemologists we are concerned with the processes and procedures involved in the *exploration* of ourselves and the world around us. Exploration comes in a wide range of forms, from the inattentive registering of new information into our belief systems as we pursue our projects, to the inquiries institutionalized in the natural and human sciences. Clearly, exploration is an activity that generates and embodies states and things of value, and so it should come as no surprise that the past twenty years has seen epistemologists turn to focus on exploration via the values that it embodies and generates—a change of perspective that Wayne Riggs (2006) has called ‘the value turn in epistemology’.¹ Put most generally, those involved in the value turn concentrate on the ways in which exploration contributes to the goodness of life. I suspect that this change of focus will, over time, prove useful, as it will (among other things) lead epistemologists to ask different questions about the same terrain. The present paper is intended as an illustration of this potential; I will be working towards pressing two questions in epistemology that emerge with great clarity when we look at exploration with an eye to the values that it generates.²

The starting point for my discussion is the thought that *believing* results in or embodies a range of *goods*; I will approach the values of exploration, in other words, through *doxastic goods*. Doxastic goods correspond to the ways in which believing is a valuable state, or contributes to valuable states or things. As with all starting points, this one is by no means innocent.

¹ Some of the more influential works in the Value Turn are Craig 1990; Alston 1993; Kvanvig 2003.

² It would be misleading to characterize this paper (or the value turn in general) as an exploration of the analogies and disanalogies between belief and action. Such a characterization would be motivated, I suspect, by the regrettable fact that value theorists have focused almost solely upon the values associated with voluntary action, and that, consequently, we understand them better than we do the values associated with exploration.

First, this starting point ignores stages, processes, and states involved in exploration other than believing. Inquiry also includes, for example, a phase that aims not at beliefs but at *questions*, a stage in which problems are raised and developed, and their importance discussed. This process, of recognizing our own ignorance and determining which aspects of our ignorance should be rectified, is not separate from inquiry; it is part of inquiry, with its own procedures, criteria for success, and goods. Particular inquirers engaged in the process of questioning may not see themselves as *at all* involved in the pursuit of belief. Furthermore, even if we ignore the questioning phase of inquiry, belief may not be the only or best aim of the 'later' or 'concluding' phase of inquiry. Perhaps the ultimate aim of inquiry is or should be some state distinct from belief. Bas van Fraassen has argued, for example, that the end product of theorizing should be something weaker than belief, a state he calls 'acceptance' (van Fraassen 1980).

A more significant concern about my starting point will be expressed by those who think that discussions of value should begin not with goods or benefits, but with norms and reasons. Proponents of this latter approach, who are often deontologists or Kantians, say that in many realms value should not be seen as residing in goods in the world, but should rather be understood in terms of our following or being able to follow norms and reasons. I do not wish to deny that this latter approach will, ultimately, prove to be the correct one to take regarding the values of exploration. However, as we will see, the extent and nature of doxastic reasons and norms have long been subjects of contention, and by approaching this area via the goods we gain in virtue of believing, we may find new ground for advancing the discussion.

The following is a list of some possible doxastic goods, which I will hereafter refer to as the 'List':

- the feeling of well-being or happiness that a belief can give us;
- the value of holding an emotion whose appropriateness is informed by an appropriate belief;³
- the value of being characterized by a virtue which is (partly) constituted by a belief or is appropriately informed by a belief;
- the value of a relationship which is made possible or better because of a belief;
- the value of belonging to a social group in virtue of one's beliefs;
- the value of acting successfully on the basis of an appropriate belief;
- the value of believing something simple or useful;
- the credit I obtain or deserve when I am responsible for believing a truth or holding a justified belief;

³ I use the word 'appropriate' here (and elsewhere in this list) rather than 'true', because many doxastic goods can accrue to us whether or not the belief that delivers the good is true.

- the value of believing something that will explain a wide range of phenomena;
- the value of believing something that fits with the rest of my belief system;
- the value gained by truly believing, by justifiedly believing, or by justifiedly *and* truly believing;
- the value gained by knowing.

As we can see, doxastic goods are by no means a trivial subset of the values of exploration. On the contrary, believing seems to provide or give access to an enormous range of goods. The List may look, at first glance, uncontentious. However, some philosophers would take issue with the List as I have presented it. They would want either to question whether many of the members on the List are, really, doxastic goods, or to stress a distinction between some of the goods on the List and others. The considerations that lead them to such responses are the central concern of this paper.

Perhaps the most prominent dispute regarding the List is that between so-called Evidentialists and Pragmatists—a debate going back in recent history to the work of W. K. Clifford (1887/1999) and William James (1923). This debate hinges upon whether believing in order to attain particular entries on the List are cases in which believing *per se* goes well. The Evidentialist says the members towards the top of the List are not germane to the nature of believing; believing in order to achieve happiness or group-belonging is not an instance of good believing. As a consequence, the Evidentialist says that the goods at the top of the List (if they are *doxastic* goods at all) do not provide us with reasons to believe (although they may provide us with reasons to act *so as* to believe). The Pragmatist, in contrast, denies that (at least some of) the entries towards the top of the List are not reasons to believe; for her, an agent's believing in order to gain happiness or group belonging may by that very fact be an instance of good believing.

My primary concern in this paper is not with the debate between the Evidentialist and Pragmatist, although some of what I say will have a bearing on this debate.⁴ Rather, I am here interested more generally in what should be on the table as we talk about the various goods of believing. More precisely, I will be arguing that one consideration that must not be ignored is the ability of a doxastic good to *motivate* belief.⁵ The plan of the remaining three sections of this paper is thus. In the Section 2 I argue that some doxastic goods are *surreptitious*

⁴ My sympathies with the Evidentialist position, some of which I have pressed in Jones 2004, will be evident, but defending Evidentialism is not my primary aim here.

⁵ When I speak of the 'motivation of belief', as I will throughout this paper, this phrase should not be understood in the more technical sense of 'motivated belief' that is found in work on self-deception and wishful thinking (e.g. Pears 1984). Rather, the 'motivation of belief' refers to any process in which an agent is motivated (moved, caused, led) by some consideration to adopt a belief.

motivators; that is, some doxastic goods are such that in so far as they work to bring about belief, they must do so without the agent herself being aware of their so working. The existence of surreptitious motivators raises two questions, which I address in the remaining two sections of this paper. First, I ask, 'Do surreptitious doxastic goods provide us with reasons to believe?' Recent work in value theory on the relationship between reasons and motives suggests an argument for the following claim: if a good cannot consciously motivate a believer to believe, then that good cannot provide a reason to believe. I am sympathetic to this position, and will rehearse three arguments for it. The question I turn to in Section 4 is the following, 'Is it possible that the important exploratory values in our lives are surreptitious motivators?' My discussion of this second question will involve looking at the so-called credit account of the value of knowledge, which conceives of the value of knowledge as a surreptitious motivator. Questioning whether the credit account is acceptable, I will end with a brief look at an alternative way of conceiving of knowledge and its value.

2. THE SURREPTITIOUS GOODS OF BELIEVING

The goods that we receive in virtue of believing fall into two categories, according to their ability to bring about belief *in the light of our conscious consideration of them*. Some doxastic goods are such that my consciously realizing that coming to believe that *p* will (or already does) allow me to attain them can, itself, determine my belief that *p*. Other doxastic goods, however, cannot. The latter can only work *surreptitiously* in bringing about belief; they cannot bring about belief in a way that is 'open' to the believer.

The core of the phenomenon is the following: many of the benefits that can be gained from believing are such that the agent's thinking that she believes what she does *in order to* gain that benefit is in tension with the belief itself. I may believe that my sick mother is recovering from her illness, and I may recognize that this belief makes me feel relief. However, in so far as I think that my belief is dependent upon my feeling relief—that I am believing that my mother is recovering only in order to feel relief—so far will that belief be weakened. Another example: I may believe that God exists, and I may recognize the goods that this belief gives me (e.g. alleviates my fear of death, increases my chances of immortality), but as long as I continue to believe that God exists, I will not see my belief as being wholly *dependent upon* my concern for these goods. My believing that *p* is incompatible with my seeing my doing so as determined by my concern for these kinds of gains.

This feature of believing separates it from other truth-aimed states like guessing, speculating, or conjecturing. It is true that when I guess or speculate, as when I believe, I am after a truth about the subject matter before me (Velleman 2000). However, in contrast to belief, I can be fully aware that I am guessing,

speculating, or conjecturing *right now* because of other goods that these states give me. I can in full awareness have each of the following thoughts, for example: I am guessing that *p* now because I am playing a trivia game; I am speculating that *p* now because I have a student before me and I need to give her some ideas; I am conjecturing that *p* right now because I am having a friendly conversation about politics or history, and I find it enjoyable to do so. In all of these cases I see myself as aiming at a truth, but I also see each of these states as a *pressurized*, *motivated*, or *enjoyable* aiming at truth. Belief is different. One of my beliefs *may be* a pressurized, motivated, or enjoyable aiming at truth, but I cannot see it as such. As David Owens has written, 'The guesser can [decide what to guess] by reflecting on how best to strike a balance between the goal of truth and other goals her guessing serves; the believer cannot' (Owens 2002: 395). Someone can see herself as guessing that *p* (rather than that *q*) in order to win a game or gain a prize, but she cannot see her belief this way. In acknowledging that I have a belief that *p*, I characterize myself as being in a state *only* for the purpose of being committed to a truth (or knowledge, or understanding) about the subject matter at hand.

The stronger I think that my commitment is held in the light of a concern for this kind of gain, the stronger will be my inclination to characterize this as something other than a belief. We all suspect that some of our beliefs are affected by certain surreptitious determinants; to cite just one familiar example: given my pervasive agreement with my friends, colleagues, and peers, I often find myself with the suspicion that my personal relationship with them—or feelings for them—has had an effect on my beliefs, an effect that is not dependent upon my having some reason to think that this person is more likely to have true beliefs about the subject matter upon which we agree. However, the stronger these suspicions are, the weaker is the belief itself. As Michael Ayers has written, 'In so far as we doubt that grounds wholly determine our belief, so far is our belief itself subjectively insecure' (Ayers 1991: 148).

It is a familiar feature of folk psychology that we say of other persons (or ourselves in the past) that they are (or we were) believing in order to gain relief or to avoid pain. We say that a believer is engaging in wishful thought, or that she believes that *p* because it is too painful to believe otherwise. However, it is revealing that we are very reluctant to accept that person's explaining her own belief in this way. Imagine being confronted with a friend, Jones, who tells you that he believes that it will be sunny tomorrow because he wants to have a picnic. Jones is telling you that he believes something about the weather because of his own wishes for what the weather will be like. It seems clear, however, that you would not take Jones's claims at face value. You would not accept his assertion of both the belief *and* the explanation. Instead, you would tend to re-interpret Jones's statement: either he believes that it will be sunny tomorrow and he is joking about the explanation, or he is merely confessing that he does not really believe that it is going to be sunny tomorrow at all. The best explanation for this

is that we recognize that Jones's believing something is in tension with his seeing it as an attempt to gain such a good.

Some of the goods on the List are going to be surreptitious, but not all of them. A believer can readily recognize that her believing that *p* was and is motivated by a concern to grasp a truth about a subject matter. She wants, for example, a true belief about the status of her mother's health, and she believes that her mother's health is not improving because her father (whom she trusts) told her so. She can readily acknowledge the immediate good of this belief—that of gaining the true belief about her mother's health that she sought. And she can readily acknowledge that *achieving this good* is what motivated her to hold this belief. She believes what she does because she wants a truth about her mother's health. Grasping a truth is one of the ways in which believing can go well, and one of the ways in which exploration can be successful; it is one of the things that we seek from believing, and one of the goods that results from it. However, in contrast to the other goods that have been discussed in the present section, it is not true that my thinking that I have a belief because it is true interferes with my holding the belief. To recognize that I believe that *p* in order to have a true belief about some subject matter is not in tension with believing that *p*. On the contrary, this explanation of my belief is readily acknowledgeable by me.

In sum, we can divide the members of the List into those that are *acknowledgeable* or *open* motivators, and those that must motivate *surreptitiously*. The latter cannot be recognized by the believer herself as playing a determining role in her belief. They can be admitted in third-personal doxastic explanations, but not in first-personal, present-tense doxastic explanations. I can be motivated to believe something by a concern for surreptitious goods that I might get from such a belief, but my doing so must be hidden from me.⁶

Why does it matter, to questions regarding the values of exploration, that some doxastic goods are surreptitious? This is the question that I will be answering in the remainder of this paper.

3. GOODS, REASONS, AND MOTIVATION

The relationship between goods, reasons, and motivation is both intricate and contested. I have two aims in this section: (i) to illustrate the two main sources of disagreement over the relationship among this trio of notions, and (ii) to support my sympathy for one view of the relationship between doxastic reasons and doxastic motivation, a view that entails that the surreptitiously motivating goods on the List are either not goods at all or not goods that can provide us with reasons to believe.

⁶ The claim of this section is discussed and defended at length in Jones 2002. There I also defend the claim that this is a necessary phenomenon, constitutive of believing, a claim that need not be defended here.

3.1. Reasons and Motivation, Reasons and Goods: A Rough Map

The main source of disagreement in the debate over the relationship between reasons and motivation arises with the following question: if Jones is in the presence of a reason to $\emptyset$, is he *necessarily* in the presence of a *potential motivator* of his $\emptyset$ -ing? Must, in other words, Jones's reason to $\emptyset$ also be something that can move him to $\emptyset$?

Those who answer this question negatively will claim that it is at least possible that there is a reason for Jones to $\emptyset$ even though it would in no way be possible for Jones to be motivated by that reason to $\emptyset$. Whether or not Jones's reason to $\emptyset$ can motivate him is *contingent* (e.g. upon Jones's psychological status). In a recent discussion of reasons, Allan Gibbard eloquently spells out this view. He writes,

The Marquis, suppose, has no motivation to keep from hurting me, and no motivation could be aroused by sheer information, vivid awareness, and moves toward consistency. I can still say that he has reason not to kick me: kicking me would hurt me, and that, in and of itself, is reason not to kick me. (Gibbard 2003: 291)

On this conception of reasons, a reason to $\emptyset$ is *simply* something that *favours* $\emptyset$ -ing, irrespective of whether it has the potential to motivate any particular agent to $\emptyset$.

While other writers agree that reasons to $\emptyset$ favour $\emptyset$ -ing, they think that this is only part of the story. This is only one side of reasons, their *normative* side. However, they will continue, reasons also have another side, in that they can make a difference to what we do. They are also, by their nature, that *in the light of which* we do something.⁷ Reasons, according to this position, have a *motivating* side in addition to their normative side.⁸ On this position, it is constitutive of a reason that it be a potential motivator of an agent's behaviour. Accordingly, in order to determine whether something provides a reason to $\emptyset$, we also have to determine that it has the potential to motivate; if Jones could not have acted or believed given the consideration of $\emptyset$, then $\emptyset$ cannot be considered one of Jones's reasons.⁹ Reasons are necessarily potential motivators.

The claim that reasons are necessarily potential motivators is complicated by the various ways in which we can understand what it means to be a 'potential motivator'. A vivid example of this arises in the debate between Bernard Williams (1980; 1995) and John McDowell (1995a) over 'internal and external reasons'. Both agree that reasons are necessarily potential motivators, that 'Any reason

⁷ It is important that the italicized phrase in this sentence be left imprecise, as different ways of making it more precise define the camps that are the concern of this section. A closely related (but just as loose and more awkward) phrase is 'in response to the consideration of which'.

⁸ The conception of reasons as having 'two sides' comes from Dancy's 2002: ch 1.

⁹ For explicit defences of this position, see Williams 1980 and Dancy 1994.

for action must be something that *could* explain someone's acting in the way for which it is a reason' (McDowell 1995a: 70). However, Williams argues that nothing counts as a reason for Jones to do something unless it can, *by Jones's own deliberation*, be brought within what Williams calls Jones's 'motivational set'. An agent does not have a reason to $\emptyset$ unless it is possible for that agent to come to desire $\emptyset$ -ing or its possible result by deliberation. Jones does not have reason to $\emptyset$ unless Jones can, by deliberation, come to see $\emptyset$ -ing as desirable. McDowell has a more liberal view than Williams with respect to the 'could' here, that is, with respect to what counts as a potential motivator. He suggests that something should be seen as a potential motivator—and, thus, a reason—for Jones to do something even though it could only reach Jones's 'motivational set' by some route other than deliberation.

As a consequence, Williams's list of potential reasons is smaller than McDowell's. By having a *stricter* conception of a potential motivator, Williams rules out certain reasons—those he calls 'external'. On the other hand, by being more *liberal* in the ways in which a reason can motivate behaviour, McDowell allows, we might say, for the existence of more (kinds of) reasons, including those that are external. These two categories—strict and liberal—are *relative* to each other, and not absolute; Williams has a *more* restricted attitude towards potential motivation than McDowell.

So, we have three positions on the relationship between reasons and motivation. Proponents of the position that denies a necessary relationship between reasons and motivation hold that *anything that favours* $\emptyset$ -ing gives one a *prima facie* reason to $\emptyset$, independently of whether or not it has the potential to motivate. Their opponents claim that reasons are necessarily potential motivators. Once one occupies this latter position, then one's view of what it means to be a potential motivator becomes important, and will necessarily affect one's view of whether something is or is not a reason. Figure 6.1 lays out these various positions. There are various ways in which one can have a more or less restricted view of potential motivators. Williams and McDowell show us one way—in terms of accessibility via deliberation—but we will see a different way—in terms of surreptitiousness—in a moment.

In the doxastic realm, I think that the right-hand, 'stricter' view of doxastic reasons and their motivating power is the most attractive. Before defending this, however, I want to briefly look at the two views of the relationship between goods and reasons. On the first, all goods are reasons; if something does not provide us with a reason, then it *ipso facto* is not a good. On the second position, something can be a good even if it does not provide a reason.

On the first view, a good has a straightforward and internal relationship to a reason: a good is necessarily something that gives me a reason for behaving—acting, believing, feeling—in some way or another with respect to it. If something of value is attainable, creatable, or conservable, then I *ipso facto* have a *prima facie* reason to behave in such a way that I can attain that good,



Figure 6.1.

bring it about, preserve it, or respect it. As a *prima facie* reason, it can be overridden by a conflicting good (reason), but if something is a good, it provides a reason nonetheless. Joseph Raz expresses this straightforward conception of the relationship between goods and reasons when he writes,

In general the value of what has value, and the action its value is a reason for, are intrinsically connected. We cannot understand what is of value [e.g.] in a party without understanding what [a party] is a reason for, that is, when one has reason to go to one, and how one behaves at a party. . . . What goes for parties goes for musical compositions, mountaineering, philosophy, love, and everything else. (Raz 2001: 164–5)

To see something as being of value, this passage claims, is necessarily to see it as providing one with a reason to behave in such a way as to attain, preserve, or respect it. If something does not or cannot, for some reason, provide us with a reason to act, feel or believe, then it thereby does not count as a good. Such a thing might *seem* to us, at first, to be a good, but were we to discover that it does not after all provide us with a reason, then it would be thereby revealed to be only an *apparent good*, and not a good at all.

The alternative position denies this necessary link between goods and reasons. While goods are the kind of thing that as a matter of fact provide us with reasons, something can still be a good, this position allows, even though it does not provide someone with a reason. On this view, a good or valuable thing or state of affairs need not be related to a person or persons in such a way that it can drive them to act or believe. On the contrary, a person may be *alienated* in some way (e.g. spatially, temporally, cognitively, motivationally) from a good such that it does not or cannot provide them with reasons; nonetheless, this does not disqualify it from being a good.

Because the proponent of the stricter view of reasons and motivation—recall the right-hand position of Figure 6.1—rules out certain *apparent* things or states of value as providing us with reasons, she must face a choice between these two positions on goods and reasons. The proponent of the stricter view says that something can only provide us with a reason to $\emptyset$ if it is able to motivate us in a certain way. So: if X cannot so motivate us, is it a good or not? The proponent of the first position on goods and their relationship to reasons says that if X cannot provide us with a reason to $\emptyset$, then it is not a good; the proponent of the second position says that X may still count as a good even though it cannot provide us with a reason to $\emptyset$.

3.2. Interrogating the List

The discussion of the previous few pages is meant to show that there is a range of questions to be asked in a discussion of possible goods. It is time to consider how such questions should be brought to bear on our List of possible doxastic goods from Section 1. I have been suggesting that the first question we should ask about a member of the List is the following:

(Q1) Can it provide someone with a reason to believe something?

Our answer to this question depends upon our answer to a second question:

(Q2) Are reasons necessarily potential motivators?

Those who respond in the negative will say that whether a member of the list can motivate belief is irrelevant to its status as a doxastic reason; this is the position depicted on the left-hand side of Figure 6.1. Those who claim, on the other hand, that reasons must be potential motivators will say that we must look at the motivating power of a member of the list in order to determine whether it can provide us with a doxastic reason.

We saw in Section 2 that some of the members of the List can only bring about belief in a surreptitious manner. A believer cannot be led to belief by the conscious consideration of those members towards the top of the List. So, those who think that reasons must be potential motivators now have a third question to answer:

(Q3) Must doxastic reasons be able to motivate consciously?

A lenient view of potential motivation (recall the middle position in Figure 6.1) will allow that reasons can motivate surreptitiously. If someone has believed in order to gain some benefit, then, according to this lenient position, we can see the believer as having believed for a reason (i.e. to gain the benefit) *even if* the belief-formation had to proceed unconsciously. Even the surreptitious goods on the List can be doxastic reasons. A stricter view of potential motivation (the right-hand position in Figure 6.1), on the other hand, will say that surreptitious

motivation is not good enough. If someone can only unconsciously be brought to believe in the light of considering a member of the List, then that member will not count as a doxastic reason. Only non-surreptitious motivators are potential reasons to believe.

In the next sub-section, I will rehearse three arguments for a version of this stricter view of doxastic reasons. As we saw in the previous sub-section, however, the proponent of such a strict view has a fourth question to answer:

- (Q4) Are the non-surreptitious motivators on the List only *apparent* doxastic goods, or are they *bona fide* goods that do not provide us with doxastic reasons?

I will not look at reasons that the proponent of the stricter position may give for adopting either of the positions described in (Q4). Nonetheless, even if she ends up adopting the weaker of the two options, and allows that all members of the List are bona fide goods, her answers to the first three of our questions reveals her denial of something significant to some of the members of the List, namely that they can provide us with reasons to believe anything at all.

3.3. The Case against Surreptitious Goods as Reasons

The case for a strict position on doxastic reasons can be summarized into the following argument:

The Argument

Something is or provides us with a reason to believe only if it can motivate belief in the light of the believer's conscious consideration of it. Doxastic goods that can only motivate surreptitiously, therefore, do not provide us with *reasons* to believe (although they may provide us with reasons to *act* so as to gain a belief). If they are goods at all, then they are, at best, beneficial side-effects of believing.

Analogously to Bernard Williams's dismissal of external reasons in the light of his strictures on action motivation, we here have a serious demotion of surreptitious doxastic goods in the light of strictures on doxastic motivation: either they do not exist at all, or they cannot provide us with doxastic reasons. The crux of the argument is in its restriction on what counts as a reason to believe. If my being aware of some good that I will gain by believing cannot consciously lead me to believe, then the thing I am aware of cannot provide me with a reason to believe. If the members towards the top of the List are goods, they are not goods that can provide reasons to believe. It might be *to my advantage* to believe in order to gain such goods, and such goods might provide me with reasons to *act* so as to believe (*à la* Pascal), but it cannot itself be or provide a reason to believe.

Can the Argument provide us with an argument for Evidentialism, for the view that only truth-aimed reasons are doxastic reasons? Indeed it can, and I am

aware of two recent attempts, by Jonathan Adler (2002) and Nishi Shah (2006), to use something like the Argument to establish Evidentialism. However, the Argument needs supplementing in order to provide a case for Evidentialism. First, the Evidentialist needs to defend an extra claim to get from the conclusion of the Argument to her position. The conclusion of the Argument is

- (1) Only non-surreptitious doxastic goods provide us with reasons to believe.

What the Evidentialist wants is

- (E) Only truth-related goods provide us with reasons to believe.

For all that (1) tells us, there might be a doxastic good that does not motivate surreptitiously, but which generates reasons that are not truth-related. So, the Evidentialist needs to defend the claim that

- (2) All non-surreptitious doxastic goods are truth-related.

Various recent discussions of the phenomenon of *transparency* in conscious doxastic deliberation have revealed the plausibility of (2).¹⁰ As Shah, for example, writes, ‘the [conscious] deliberative question *whether to believe that p* inevitably gives way to the factual question *whether p*’ (2006: 481). When consciously reflecting on whether to believe or continue believing that *p*, *all* that we take to be relevant are considerations that bear upon whether or not it is true that *p*. If this is right, then non-surreptitious doxastic goods—those which we can pay attention to in conscious doxastic deliberation—will be co-extensive with those goods that generate truth-related reasons.

More fundamentally, the Argument itself needs substantiation; we need a *reason* to believe its main premise, namely that doxastic reasons cannot motivate surreptitiously. We need a reason not to accept either the middle or left-hand position in Figure 6.1—either the view that allows that non-deliberating doxastic motivators can be reasons to believe, or the view that wholly ignores a good’s ability to motivate in considering it to be a reason. At issue, clearly, is the nature of a reason, an enormous topic, and one that I cannot hope to make a significant contribution to here. Nonetheless, I want to briefly gesture towards three considerations in favour of the more restricted view, looking first at Jonathan Adler’s recent defence of Evidentialism in the first chapter of his book *Belief’s Own Ethics* (Adler 2002).¹¹

Adler argues that Evidentialism can be shown to be true by a close examination of the concept of *belief*. Adler defends a claim closely related to that defended above in Section 2: while it is true that beliefs can be formed and held on the basis

¹⁰ A comprehensive recent discussion of transparency can be found in Moran (2001). Also see recent work by David Velleman and Nishi Shah.

¹¹ I will be presenting a version of Adler’s position that is weaker (and I think more plausible) than his own.

of no epistemic evidence at all, this fact cannot be recognized by the conscious believer. Adler's explanation of this phenomenon is that the concept of belief has strict norms built into it, and among them is the norm that one should believe something only if one is believing on the basis of epistemic evidence. The reason that it is impossible to consciously believe in order to gain non-epistemic goods, is that 'in first-person awareness we recognize the demands of belief' (2002: 52); in particular, we recognize both that a norm for believing is to be held only on the basis of epistemic evidence *and* that we are not following this norm.

Put in terms that I have been using in this paper, Adler's suggestion seems to be this: the fact that certain goods can only motivate surreptitiously reveals that believing in order to gain such goods is a violation of the norms of believing. The key player in Adler's argument is conscious reflection; conscious, reflective believing is necessarily going to be a case of believing in accordance with (what the believer takes to be) the basic norms of belief. Any case of believing that *must* be non-conscious and non-reflective, therefore, must be a case of believing that violates the basic norms of belief. Any good on the List that must motivate belief behind the scenes, as it were, must not be a motivation in accordance with the norms of believing. Therefore, surreptitious goods—which must work behind the scenes—are working illicitly. Believing on the basis of surreptitious goods is illicit, a violation of the basic norms of believing, and so surreptitious goods do not provide us with reasons to believe. While some of the details of Adler's argument need filling in, I find its central thought plausible: the fact that our reflective awareness blocks certain goods as being attained in believing is a function of our being aware that believing in order to gain such goods is not a case of believing going well.

A second defence of the Argument hinges, like the first, on thinking of doxastic reasons as intimately related to norms governing believing. Norms need be neither consciously considered nor consciously followed. I can adhere to a norm without ever having thought explicitly about what that norm is. Norms of language and concept-use are well-known examples. Jones's using a concept in a certain way at a certain time can be explained in virtue of her accepting and following a particular norm, even though Jones could not state the norm that she is following. Nor, *a fortiori*, need it be true that in order to have my behaviour dictated by a norm, I must be thinking about the norm at the time of use. I can follow a norm without being aware of what I am doing, without, that is, knowing what norm I am behaving in accordance with. Norms play a pervasive role in our lives, but in doing so they often remain in the background, hidden from our conscious purview of what we are doing. There is a difference, however, between claiming that a norm *need not* be consciously followed, and claiming that it *cannot* be consciously followed. The latter makes no sense. It is essential to something's being a norm that it be such that we are able, reflectively, to 'lead ourselves to follow it'. Put succinctly, even if not consciously followed, a norm must be consciously followable. Behaving on the basis of an accepted norm is

necessarily to behave on the basis of something that one *can* consciously endorse and follow.

Is that which is true of norms also true of reasons? Any answer here will be controversial, but the positive response has its adherents. Christine Korsgaard, for example, writes in *The Sources of Normativity* that 'the normative word "reason" refers to a kind of reflective success' (1996: 93). And speaking of doxastic reasoning in particular, Bill Brewer writes that: 'Epistemically productive reasoning is . . . a compulsion in thought *by reason*, and as such always involves some conscious understanding of why one is right in one's conclusions' (1995: 242).¹² If these thoughts are along the right line, that is, if reasons are like norms in possessing the necessary characteristic of being consciously followable, then it would follow that surreptitious doxastic goods cannot provide us with reasons, for a surreptitious doxastic good cannot, by its very nature, allow a conscious awareness of being led us to believe.

A third, closely related, consideration that supports the Argument can be found earlier in Korsgaard's discussion in *The Sources of Normativity*. She is concerned, at the outset, with the conditions upon an acceptable portrayal of our adherence to morality.

[A] successful normative theory must meet a condition which is sometimes called 'transparency'. . . If a theory's explanation of how morality motivates us essentially depends on the fact that the source or nature of our motives is concealed from us, or that we often act blindly or from habit, then it lacks transparency. The true nature of moral motives must not be concealed from the agent's point of view if those motives are to be efficacious. . .

Otherwise, Korsgaard claims, the theory would not be a *normative* theory at all:

A normative moral theory must be one that allows us to act in the full light of knowledge of what morality is and why we are susceptible to its influences, and at the same time to believe that our actions are justified and make sense. (Korsgaard 1996: 17)¹³

A non-transparent moral theory would ultimately amount to a portrayal of moral behaviour as motivated by considerations that are hidden from us. Similarly, a non-transparent theory of believing would portray believing as motivated by considerations that are hidden from us.

Why would this be a problem? Why would it matter if a moral or doxastic theory portrays us as motivated by hidden considerations? One thought is that in so far as it does, it would portray us as diminished agents. It is crucial to our identity as agents that we can allow reflective adherence to reasons to guide our lives. It is because it is possible for me to act *because* and *in the light of* my reflection upon what I think is the right thing to do, that I am an agent of

¹² See also Millar 2004: ch 7.

¹³ Korsgaard's use of the word 'transparency', which is here being used to characterize a *theory*, differs from the use of the same word discussed on p. 150, where it is used to characterize *believing*.

moral action. Similarly, it is because it is possible for me to believe *because* and *in the light of* my reflection upon what I think is the right thing to believe, that I am a doxastic agent. While it is clearly true that we are sometimes moved to believe on the basis of hidden motives (e.g. when we engage in wishful thinking or self-deception), in so far as we do so we are not believing as agents. My doxastic agency is essentially tied up with the fact that reasons can, in my full awareness, dictate what I do. If we are doxastic agents, goes this thought, then it must be that we are at least capable of following doxastic reasons. The realm of reason-following is the realm of agency. A theorist who allows there to be reasons that are not consciously followable would be, on this account, a theorist confused about the intimate relationship between reasons, conscious following, and agency.

These three considerations in favour of the claim that surreptitious doxastic goods do not provide us with reasons to believe seem to me to have considerable strength. However, I do not take any of them (or their conjunction) to be decisive, especially in the abbreviated version presented here. At a minimum, however, all three point towards the large issues—among them self-reflection, norms, and our doxastic agency—that need exploring as we work to decide the relevance of surreptitiousness to doxastic reasons and goods.

4. HOW SIGNIFICANT CAN A SURREPTITIOUS VALUE BE?

4.1. The Attenuated Role of Surreptitious Goods

Surreptitious doxastic goods must be all but absent in two areas of our doxastic lives. First, surreptitious doxastic goods can be of little or no positive significance in our reflective view of ourselves as believers. This is a consequence of the fact that they will not be something that we consider in our deliberations about what to believe. As Richard Foley has astutely observed, ‘when people reflect upon what reasons they have to believe something. . . they rarely even consider the practical advantages that might accrue to them by believing it’ (Foley 1987: 214–15). In my own doxastic deliberation, in considering what to believe, I do not focus on all of the goods that may accrue to me in coming to believe one thing rather than another. The reason for this, as we saw above, is that a believer cannot acknowledge that he believes out of a concern for what Foley calls ‘practical advantages’.

As we saw above, doxastic deliberation is transparent; in deliberating about what to believe, we focus not upon the belief and its features, but upon the world and what we know of it, upon what features the world is likely to have or not have. That, and not ‘practical advantage’, is the locus of doxastic reflection. As a consequence, surreptitious doxastic goods are not going to loom large in our

views of ourselves, our lives, and the direction that we want our lives to go. We are not going to rank surreptitious believing as an important feature of ourselves; we are not going to seek out surreptitious believing as a significant goal; we are not going to be proud of or prize ourselves as surreptitious believers. Again, the reason for this is that surreptitious doxastic goods are, as their name implies, hidden from us.

A second arena from which surreptitious goods will be almost wholly absent is that of doxastic discourse; surreptitious goods can play little role in *argumentation*. Again, this is something that Foley has pointed out to us: 'Likewise, when someone tries to convince another person that he has reasons to believe something, they rarely even mention the practical benefits that might result from believing it' (Foley 1987: 215). The 'practical benefits' of believing are rarely spoken of in our doxastic discourse. In attempting to convince someone of something, we almost never raise these kinds of benefits (e.g. happiness, relief, fame) that a believer will attain from doing so. Scientists, philosophers, historians, and those engaged in informal arguments at the pub do not attempt to win over their opponents by pointing out that goods of this ilk will accrue to them by converting to the other side. An argumentative move of the form 'You should believe that *p*, because it will make you happy or bring you fame', is almost wholly impotent to move belief.

I say *almost* impotent, because, while a surreptitious doxastic good cannot consciously motivate belief, it can consciously motivate *action*. In particular, such a good can motivate an action *that is meant to lead to belief*. I may act with the intention of gaining the belief that *p*, in full awareness that what I hope to gain is a surreptitious good from this belief. This, indeed, is what Blaise Pascal was attempting to get his readers to do in the *Pensées*. After arguing that we should believe in the Articles of Faith because of the pragmatic benefit it will give us, Pascal suggests engaging in behaviour that will lead to such belief. 'Learn from those who were once bound like you and who now wager all they have', goes his famous advice to his reader, 'They behaved just as if they did believe, taking holy water, having masses said, and so on. That will make you believe quite naturally, and will make you more docile' (Pascal 1966, fragment 418, Lafuma numbering).

Pascal recognized that even though a desire for salvation (or a fear of damnation) cannot in full awareness lead one to believe, it can lead one to act in order to gain the belief. It is the latter that he offers as a way to belief. He also recognized that we do not immediately respond with belief to the prospect of this kind of doxastic good; we can only seek such goods indirectly. What is more, even if Pascal's indirect argument for belief works on one of his readers, its success must, in the end, be hidden from the believer. Once I become a Pascalian believer, I can recognize that it was the desire for immortality led me to follow Pascal's advice to *act* in certain ways; however, I cannot think that my *believing* in God's existence depends upon my desire

for immortality. Pascalian belief-formation must, in a sense, cover its own tracks.¹⁴

In short, surreptitious goods must by and large drop out of our evaluative views of ourselves and discourse directed at convincing or undermining belief. As a consequence, questions arise as to what explanatory role we can give them in our accounting for the goods of exploration.¹⁵

4.2. The Value of Knowledge as Surreptitious

One of the primary objectives of those involved in the value turn in epistemology is to describe and explain the more important doxastic values. How does good believing contribute to the value of being virtuous, of relationships, of social belonging? What, if anything, is the value of holding a belief that aids explanation or that coheres with other beliefs I have? What, precisely, is the value of holding a true or justified belief, or of possessing knowledge? As we proceed in answering such questions, it will emerge that some doxastic values are more important than others, that some are more central or fundamental in our lives. For these, the question arises, 'Can an *important* value be one that only surreptitiously motivates?'

I feel a strong inclination to deny that it can, to insist that the more important a value is, the more open it must be in its effect upon us. Indeed, one might wonder how what we view to be an important value could become so viewed were it *not* open. How could a surreptitious value, one that only worked behind the scenes, come to *seem* important? However, an initially plausible account of one of the more important doxastic values—that of knowledge—appeals to a surreptitious motivator. In this final part of the paper, I turn to this account and the implications of its dealing in surreptitious currency.

One of the more elusive targets in the burgeoning discussion of epistemic values has been an explanation of the value of knowledge. Sometimes called the 'value problem', the question can be put in various ways. What value accrues to us as knowers that we do not gain from being true believers? What accounts for the value that we place on knowledge over that of *mere* true belief (i.e. true belief

¹⁴ For a more extensive discussion of Pascal on this issue, see Jones 1998.

¹⁵ In this context, it may be worth comparing surreptitious doxastic goods to surreptitious ethical virtues. Here is Bernard Williams, for example, on *generosity*: 'It is not the basic characteristic of a generous man's deliberations that they use the premise "I am a generous man" . . . Though the generous man is partly characterized by what goes into his deliberations, it is not that what goes into them are reflections on his generosity.' From Williams 1976. In the reprint in Williams 1981 the passage is at p. 48. And here is Julia Driver describing *modesty*: 'A desired feature of any account of modesty is that it explain the oddity of: (1) I am modest. (1) seems to be oddly self-defeating. If I were to utter (1), the charitable person would think that I was joking. Others would think that I was being nonsensical.' And, again: 'an asymmetry exists between the self-ascription of the virtue and the other-ascription of it. I can ascribe the virtue to another, but I cannot coherently and sincerely ascribe it to myself' (Driver 1989: 375, 380).

that does not count as knowledge)? What is our concern for knowledge or being knowers a concern *for*?

On a widespread view of the difference between knowledge and true belief, these questions do not allow for easy answers. According to this instrumentalist view, knowledge is a true belief with the property of *having been gained from a good (e.g. warranted) position or having been gained by using a good (e.g. justifying, reliable) method*. But that account—in which good belief-forming methods are seen as only a *means* to getting true beliefs—leaves us without an obvious account of the value of knowledge. This is because it leaves us without an account of the value of a true belief achieved by good means *over* the value of a true belief *not* gotten by good means. In both situations *we have all that we seek*, namely a true belief. There is nothing of value, in this story, to distinguish a true belief gained from good means over one gained from some other means. To use an analogy, if *all* you wanted to do is to have gotten across a river, then it does not matter whether you used a reliable or unreliable, safe or unsafe, method, as long as you are on the other side. In this case, *having gotten across the river using a safe method* is no more valuable a state than *having gotten across using a risky method*. In both situations, one has achieved what one wanted to achieve.¹⁶

One initially plausible response to the value problem appeals to the notion of credit. The evaluative difference between possessing knowledge and possessing a merely true belief, it is claimed, is that in the former case the believer is responsible for her possession of the true belief, and deserves credit for it. Just as we deserve credit for our successful actions, we also deserve credit for our successful beliefs. It is not that one actually *receives* credit from another person, in the form of praise or commendation; rather, it is that one has fulfilled the conditions under which praise or commendation might be appropriate. If commendation is to be given for Jones's belief, then Jones himself is the one to receive it. This account explains the value of knowledge, because, it is claimed, credit is valuable. Wayne Riggs summarizes the position thus:

A person who is causally efficacious in bringing about some positively valuable outcome is 'due' some amount of credit for having done so. . . . [W]e value coming to hold a true belief in a (sufficiently) non-accidental way because we get more epistemic credit for the true belief than we would have had we gotten it right accidentally. (Riggs 2002: 92–3)¹⁷

The value of knowledge lies in the credit that an agent deserves for accomplishing the task of gaining a true belief. Just as the state of *being a hero* is a valuable one because one deserves credit for some good deed, *being a knower* is a valuable state because one deserves credit for having gained a true belief.

Credit is something that we attain directly upon and in virtue of believing a true or warranted belief. However, like many of the doxastic goods discussed in

¹⁶ To date, the most extensive discussion of the Value Problem is Kvanvig (2003).

¹⁷ See also John Greco (2002).

Section 2, credit is a good that can only surreptitiously motivate one to believe. While my concern to gain a true or justified belief can consciously lead me to believe that *p*, my concern to gain *credit* for a justified true belief cannot. I cannot explain my believing that *p* as being done in order to gain credit for doing so, any more than I can explain my believing that *p* as being done in order to be famous or happy. Being in the state *deserving of epistemic credit* is not identical with the state *believing well enough to gain knowledge*; rather, the former state is achieved *in virtue of* the latter state. Credit is a pragmatic good that we gain *through* believing well enough to gain knowledge. So, if *achieving the good of credit* motivates a believer to believe, then it must do so surreptitiously. If the credit account is correct about the value of knowledge, then knowledge is a surreptitious doxastic good.¹⁸

The implications of surreptitiousness have already been described. Making knowledge into a surreptitious doxastic good relegates our concern for knowledge to a concern for that which can only surreptitiously motivate believing. Knowing, on the credit account, is not a good that we can openly recognize as being responsible for our believing, or which can openly lead us to believe. My conscious desire to be a knower is not itself, except indirectly, something that can lead me to be a knower; such a desire is not going to consciously work in achieving its fulfilment. In sum, the credit account must claim that my being a knower is the result of either a *hidden* concern to be a knower or an open concern for something *other than* knowledge. Consequently, on the credit account knowledge will not play a role in the interpersonal dialectical phenomena—testimony, advice, argument—in which we engage and which directly affects what beliefs we hold. Credit must play, at best, an attenuated role in these procedures. Advice or argument that runs, ‘You should believe that *p*, because in doing so you will get credit for doing so’ will not get me immediately to a belief any more than will the advice or argument, ‘You should believe that *p* because it will make you happy’. Appeals to credit in doxastic discourse will (unless presented as a Pascalian way) fall on deaf ears. Such appeals may work ‘behind the scenes’, but they will not directly and consciously affect belief as, say, an appeal to evidence would. My concern for knowledge will not, for example, lead me to trust the expert over the novice, or to respond to the good argument over the weak one; on the credit account, this doxastic behaviour is to be explained by some other concern (i.e. for true belief).

If the Argument from Section 3.3 is correct, then according to the credit theorist we have no *doxastic reason* to be a knower. That I will become a knower of *p*, or that I will gain knowledge that *p*, does not provide me with a reason to believe that *p*. In relegating the good of knowing to a surreptitious good, says

¹⁸ This is not to say that the credit view of the value of knowing cannot be combined with another view of the value of knowing. My claims here apply only to the work that appeals to credit, by themselves, can do in explaining such value.

the proponent of the Argument, the credit theorist must deny that the good of knowing can provide us with reason to believe, and—depending upon her answer to (Q4)—may be forced to deny that knowing is a good of believing *at all*.

The proponent of the credit account can, of course, recognize that a desire for knowledge can openly motivate, and provide reason for, epistemic *action*. I can *act* in order to receive credit for doing so; the man who waits until the video camera is on before jumping in the lake to save the drowning victim looks to be acting more out of concern for his own credit than for his victim. The same is true for what we might call *epistemic actions*, actions that we undertake in order to achieve beliefs about the world. I can launch an experiment, or start a philosophy paper, or work on a mathematical proof, while being well aware that my primary concern is for getting credit for what I believe after the investigation is done. However, even though credit can in full awareness lead me to these actions, it cannot be what I think has determined my belief. Knowledge may be something I actively seek, and it may be something of which I am proud when I have attained it, but on the credit account my concern for being a knower will be at one remove from believing itself.

These points should come as no surprise to the defender of credit. Credit accounts work within a framework in which the aim of exploration is true belief. The value problem arises in the face of an instrumental view of epistemic justification, and the credit account gives us an answer to the value problem *within* this instrumentalist framework. However, in doing so, it accepts one prominent feature of this framework: knowledge is relegated to a secondary aim of exploration. By not rejecting the instrumentalist framework the proponent of the credit account must see the good of knowledge as something of a by-product of exploration. This is reflected in the fact—which I have just been pressing—that credit cannot play an open role in epistemic discourse, and in the fact that credit accounts cannot acknowledge more than an indirect motivating role for believing.

Is the fact that credit is a surreptitious motivator a strike against the credit answer to the value problem? It is if you think that the following is and *must be explained to be* a platitude: ‘Even unsophisticated curiosity is a desire to *know*’ (Williamson 2000: 31). If you think that *this* must be a starting point for the exploration of epistemic value, then, contrary to the credit theorist, you will think that we must conceive of exploration as aiming at knowledge and, consequently, that a desire or concern for knowledge itself must directly motivate our epistemic behaviour. You will think that something has gone quite deeply wrong in analytic epistemology, and you will advocate embarking on a search for a quite different view of exploration, one in which not (or not only) true belief, but *knowledge itself* is a goal of inquiry, something a concern for which motivates our exploratory behaviour. Acknowledging that our concern for knowledge plays a more fundamental role in exploration requires that we throw out the framework

in which we see exploration as concerned only for truth. We need to find some way of showing that knowledge is its own goal. It is only then that we will have a way of seeing a concern for knowledge as directly motivating doxastic behaviour, and only then will we see our concern to be a knower as not being shorthand for our concern for something else.

What would an alternative framework, in which knowledge was seen as a primary concern of exploration, look like?¹⁹ We have begun to see what such a framework would look like in recent work by John McDowell and Timothy Williamson.²⁰ Both authors are fighting a 'hybrid' conception of knowledge, in which knowing is seen as a kind of true believing. Knowledge, on the hybrid account, is a true belief with a particular status. While neither McDowell nor Williamson is concerned with denying that what we know we believe, and that what we know is true, both are concerned to deny that knowing is constituted or comprised of true beliefs. On the contrary, both writers argue that knowledge is a distinct mental state. By making knowledge a mental state with all of the properties that we know knowledge must have—for example, it is true, it is intimately related to action, it has been gained in a good way—McDowell and Williamson can then make the state of knowing itself a goal of inquiry.

The McDowell/Williamson position on knowledge needs (and deserves) a good deal more scrutiny, and I do not wish to be seen as fully endorsing this position here. Rather, my concern is to point out that, as the McDowell/Williamson view straightforwardly conceives of knowing as a goal of exploration, it properly respects the platitude that inquiry aims at knowledge. This is because, in making *believing in such a way as to achieve knowledge* the goal of inquiry, it makes it the primary good to be attained from believing. The McDowell/Williamson view conceives of exploration in such a way that it does not separate—as the instrumental view does—believing well from believing truly. On this conception of exploration, we aim at both at the same time. As such, the McDowell/Williamson view has the potential to allow us to conceive of knowledge as something that we can value in ourselves, and as something the concern for which can openly—and not only surreptitiously—lead us to believe. It has, in short, the potential to allow us to see ourselves as believing in order to be knowers.

The credit account of the value of knowledge, on the other hand, does not have this potential. Starting from a position in which *true believing* is the core aim of inquiry, the credit theorist must look elsewhere for the value of knowledge. As a consequence, the credit theorist ends up telling us that the value of knowing lies

¹⁹ My own earlier response to the Value Problem (Jones 1997)—in a paper heavily influenced by the work of Edward Craig and Michael Ayers—involved spelling out an 'indicative' for knowledge. I am not convinced that this account will fare any better than the credit account in allowing us to conceive of knowledge as something that believers aim for.

²⁰ The relevant papers are McDowell 1982; 1994; 1995*b*, which together comprise part IV of his (1998*a*). Williamson's work in this area has culminated in his (2000). Their respective positions have differences that I need not discuss here.

in a *surreptitious* good of exploration, and this leaves us having to gloss over the platitude about our curiosity for knowledge. For the credit theorist, the driving motivation behind exploration—our ‘unsophisticated curiosity’—is, in reality, for truth. In so far as our concern to be knowers per se drives exploration and determines the states in which it results, this concern must work in a surreptitious and hidden manner.

Is the search for an answer to the value problem a search for an acknowledgeable value, a value the conscious consideration of which can lead us to belief? This question is only one of those that will rear their heads when we recognize, as I have pressed in this paper, two central features of the doxastic realm. First, our conception of the values that lie in the doxastic realm—that is, as goods, as providing reasons to believe—is intimately tied up with our conception of the motivating power of those values. Secondly, that many of the goods that we receive from believing have a severely limited motivating power, in that they must work surreptitiously. The theoretical use of surreptitious goods may not, ultimately, be a problem, but surreptitiousness is a feature of the doxastic realm that needs to be in the forefront of our attention if we are to take seriously the exploration of epistemic and doxastic value.²¹

REFERENCES

- Adler, J. (2002). *Belief's Own Ethics*. Cambridge, MA: MIT Press.
- Alston, W. P. (1993). ‘Epistemic Desiderata’, *Philosophy and Phenomenological Research*, 53: 527–51.
- Ayers, M. (1991). *Locke: Epistemology and Ontology*. London: Routledge.
- Brewer, B. (1995). ‘Compulsion by Reason’, *Proceedings of the Aristotelian Society*, suppl. vol. 69: 237–53.
- Clifford, W. K. (1877/1999). ‘The Ethics of Belief’, reprinted in *The Ethics of Belief and Other Essays*. New York: Prometheus Books (1999).
- Craig, E. (1990). *Knowledge and the State of Nature*. Oxford: Oxford University Press.
- Dancy, J. (1994). ‘Why There is Really No Such Thing as the Theory of Motivation’, *Proceedings of the Aristotelian Society*, 95: 1–18.
- (2002). *Practical Reality*. Oxford: Oxford University Press.
- Driver, J. L. (1989). ‘The Virtues of Ignorance’, *Journal of Philosophy*, 86: 373–84.
- Foley, R. (1987). *The Theory of Epistemic Rationality*. Cambridge, MA: Harvard University Press.
- Gibbard, A. (2003). ‘Reasons Thin and Thick’, *Journal of Philosophy*, 100: 288–304.
- Greco, J. (2002). ‘Knowledge as Credit for True Belief’, in M. DePaul and L. Zagzebski (eds.), *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Oxford University Press.

²¹ Thanks to Douglas Farland, Jane Friedman, Alan Millar, Veli Mitova, Mimi Parker, Alan Sparhawk, René van Woudenberg, Samantha Vice, and audiences at Rhodes University, the University of Stirling, and the University of KwaZulu Natal.

- James, W. (1923). 'The Will to Believe', *The Will to Believe and other Essays in Popular Philosophy*. New York: Longmans Green and Co.
- Jones, W. E. (1997). 'Why Do We Value Knowledge?', *American Philosophical Quarterly*, 34(4): 423–39.
- (1998). 'Religious Conversion, Self-Deception, and Pascal's Wager', *Journal of the History of Philosophy*, 36: 167–88.
- (2002). 'Explaining Our Own Beliefs: Nonepistemic Believing and Doxastic Instability', *Philosophical Studies*, 111: 217–49.
- (2004). 'The Pragmatic Explanation of Believing', *Crítica*, 36: 3–36.
- Korsgaard, C. (1996). *The Sources of Normativity*. Cambridge: Cambridge University Press.
- Kvanvig, J. (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge: Cambridge University Press.
- McDowell, J. (1982) 'Criteria, Defeasibility, and Knowledge', *Proceedings of the British Academy*, 68, 455–79. Reprinted in McDowell 1998a.
- (1994) 'Knowledge by Hearsay', in B. K. Matilal and A. Chakrabarti (eds.), *Knowing From Words*. Dordrecht: Kluwe 1994. Reprinted in McDowell (1998a).
- (1995a). 'Might There Be External Reasons?', in J. E. J. Altham and R. Harrison (eds.), *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams*. Cambridge, MA: Cambridge University Press. Page references are to this volume. The article is reprinted in McDowell 1998b.
- (1995b) 'Knowledge and the Internal', *Philosophy and Phenomenological Research*, 55: 877–93. Reprinted in McDowell 1998a.
- (1998a). *Meaning, Knowledge, and Reality*. Cambridge, MA: Harvard University Press.
- (1998b). *Mind, Value, Reality*. Cambridge, MA.: Harvard University Press.
- Millar, A. (2004). *Understanding People*. Oxford: Oxford University Press.
- Moran, R. (2001). *Authority and Estrangement*. Princeton: Princeton University Press.
- Owens, D. (2002). 'Epistemic Akrasia', *The Monist*, 85: 381–97.
- Pascal, B. (1966). *Pensées*, tr. A. J. Krailsheimer. Harmondsworth: Penguin.
- Pears, D. (1984). *Motivated Irrationality*. Oxford: Oxford University Press.
- Raz, J. (2001). *Value, Respect, and Attachment*. Cambridge, MA: Cambridge University Press.
- Riggs, W. (2002). 'Reliability and the Value of Knowledge', *Philosophy and Phenomenological Research*, 64: 79–96.
- (2006) 'The Value Turn in Epistemology', in V. Hendricks and D. H. Pritchard (eds.), *New Waves in Epistemology*. Aldershot: Ashgate.
- Shah, N. (2006). 'A New Argument for Evidentialism', *Philosophical Quarterly*, 56: 481–98.
- van Fraassen, B. (1980). *The Scientific Image*. Oxford: Oxford University Press.
- Velleman, J. D. (2000). 'On the Aim of Belief', *The Possibility of Practical Reason*. Oxford: Oxford University Press, 244–81.
- Williams, B. (1976). 'Utilitarianism and Moral Self-indulgence', in H. D. Lewis (ed.), *Contemporary British Philosophy*, 4th series. London: Allen and Unwin. Page references are to the reprint in Williams 1981.
- (1980) 'Internal and External Reasons', in R. Harrison (ed.), *Rational Action*. Cambridge: Cambridge University Press. Reprinted in Williams 1981.

- Williams, B. (1981) *Moral Luck*. Cambridge: Cambridge University Press.
- (1995) 'Internal Reasons and the Obscurity of Blame', in Williams's collection *Making Sense of Humanity and Other Philosophical Papers*. Cambridge: Cambridge University Press, 1995.
- Williamson, T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.

Practical Reasoning and the Concept of Knowledge

Matthew Weiner

Epistemologists have devoted much more attention to the question ‘What is knowledge?’ than to the question ‘Why does knowledge matter?’ There is a tremendous amount of work that attempts to determine which beliefs count as knowledge, less that argues that we should care whether our beliefs count as knowledge. Yet without a positive answer to the second question, the first question lacks interest.

The problem is particularly acute in the post-Gettier area. In the *Meno* Plato addresses the question of why knowledge is more valuable than mere true belief. But after Gettier, an account of the value of knowledge must explain something stronger: why knowledge is more valuable than mere *justified* true belief. As Kvanvig puts it, the question ‘is whether and how knowledge has a value exceeding the sum of its parts’ (Kvanvig 2003: p. x).

In this essay I will consider a version of the view that knowledge is valuable because of its role in practical reasoning. This view is based on a conception of knowledge presented by John Hawthorne (2004), on which it is constitutive of knowledge that what is known can serve as a premise for practical reasoning. Hawthorne’s view seems to provide a simple account about why we should care about whether we know ordinary statements (as opposed to probabilities, which I will discuss in Section 5). I will argue, however, that adverting to practical reasoning will not establish that knowledge is valuable in itself. This may seem a minor result, refuting a single argument for the significance of knowledge (and not one that Hawthorne himself puts forward). Yet I will argue that considering this view illuminates the true nature and significance of knowledge. When we examine which beliefs are suitable practical premises, we see that knowledge is not valuable in itself over and above the value of its components. Nevertheless the concept of knowledge is valuable, for it gives us an economical way of summing up many properties of beliefs, each property valuable in itself.

In Section 1 I explain the conception of knowledge that I will argue for in terms of a Swiss Army Knife, which has no value in itself over and above the

value of its components, but which is nevertheless valuable in that it provides an economical way to carry its components around. In Section 2 I begin to consider the argument that knowledge is valuable because the propositions that are known are the propositions that are acceptable premises for practical reasoning. I take up different standpoints on practical reasoning in order to show that there is no single standpoint from which the acceptable premises for practical reasoning are exactly those propositions that are known; from one standpoint the acceptable premises are exactly those propositions that are true, from another standpoint the acceptable premises are exactly those that are justified. Sections 3 and 4 consider further standpoints; though, from these standpoints, acceptable premises have some property that may go beyond mere justification or mere truth, still there is no single standpoint from which the acceptable premises are exactly those that are known. From each standpoint a different property of beliefs is shown to be valuable. Section 5 considers reasoning from probabilistic premises; it may be that the acceptable probabilistic premises are exactly those probabilistic propositions that are known, but this will not establish the value of knowledge in any way that will satisfy proponents of the traditional conception of knowledge, because on the conception of knowledge motivated by this view of practical reasoning we only have knowledge of probabilities. Section 6 returns to the Swiss Army Knife metaphor, in order to explain the utility of a conception of knowledge that integrates the properties of beliefs that have been shown to be valuable from the various standpoints on practical reasoning, and to account for why we might have developed a conception of knowledge on which Gettier cases do not count as knowledge.

1

On the conception of knowledge Hawthorne (2004) proposes, it is unacceptable to use p as a premise in your practical reasoning if and only if you do not know that p .¹ If this conception holds, knowledge is obviously important. Few things are more important than whether a belief is a suitable premise for practical reasoning, and on Hawthorne's analysis that question is the question of whether the belief amounts to knowledge.

Put another way: If we solve for X in 'A belief is a suitable premise for practical reasoning iff the belief has property X ,' then X is certainly a valuable epistemic property. But is X knowledge? I will argue that we may take several

¹ The Practical Environment Constraint (2004: 176) provides one direction of the biconditional; Hawthorne provides the other direction on 2004: 30. Hawthorne does not offer this conception up as motivating the importance of knowledge—he takes it for granted that the puzzles concerning knowledge are of interest in themselves (2004: 21 n. 49)—but it obviously could be converted into an argument for the importance of knowledge. (Thanks to Mylan Engel for pointing out Hawthorne's explicit statement of both directions of the biconditional.)

different standpoints on practical reasoning. Each of these standpoints yields a different solution for X, and each different X is thereby a property that is epistemically valuable in itself. But none of these Xs is knowledge; there is no single standpoint from which the suitable practical premises are those that are known.

This is not to say that the concept of knowledge has no value. Talk of knowledge is useful because of knowledge's relation to these epistemically important concepts. The concept of knowledge is analogous to a Swiss Army Knife, which is never needed *qua* Swiss Army Knife, but which is useful because it contains many individual blades that may be needed for their own sake. As the Swiss Army Knife provides an economical way of carrying blades that are valuable in themselves, so the concept of knowledge provides an economical way of expressing several other concepts that are valuable in themselves.

I will explain the Swiss Army Knife analogy with yet another analogy, which shows how a concept may be useful as an economical way of expressing other concepts. Let us consider two different scenarios in which an auto magazine might rate off-road vehicles as to whether they are Colorado-Rally-Worthy (my own imaginary term). In each scenario Colorado-Rally-Worthiness requires a certain mileage per tank, a certain horsepower, a certain cargo capacity, and a certain clearance off the ground. But only in the first scenario is Colorado-Rally-Worthiness valuable in itself; in the second scenario it is valuable only because of the value of its components.

The first scenario is this: There is one major road rally in Colorado. To win the Colorado Rally a driver must drive a certain number of miles without refueling, carrying a certain payload, going up mountainsides that require a certain horsepower, and over roads that will destroy your undercarriage if you don't have a certain clearance. Then Colorado-Rally-Worthiness is important in itself if you want to compete in the Colorado Rally. If your off-road vehicle falls short of Colorado-Rally-Worthiness in any respect, you might as well not run the race. There is a particular purpose for which Colorado-Rally-Worthiness is important as such.

The second scenario is this: Suppose that there is no single such Colorado Rally, but many different road rallies. In different rallies it is important to have a vehicle that gets many miles per tank, or that has a good cargo capacity, or many horsepower, or a high undercarriage. Then there will be no single purpose for which you require Colorado-Rally-Worthiness as such. Sometimes you may be able to do without that many horsepower; other times you may not need such a high undercarriage. (And there are so many rallies, possibly indefinitely many, that it is not practical to look for a vehicle that is suited to every single one.) But we might still care about whether our vehicle is one that the magazine calls Colorado-Rally-Worthy. We might seek out vehicles that are so designated because the designation is a quick way of summing up a lot of things we do care about. We'd like to have a vehicle that has each of these positive characteristics

to a certain degree. If we ask, 'Is it Colorado-Rally-Worthy?' that one question tells us four things we want to know.

On this second scenario, Colorado-Rally-Worthiness is what I'll call a Swiss Army Concept. There's no particular task that requires a Swiss Army Knife. Tasks require knife blades, screwdrivers, corkscrews, bottle openers, scissors, etc. Some tasks may require more than one tool, but no task requires that the various tools be in the form of a Swiss Army Knife. Nevertheless, Swiss Army Knives are quite useful. There is a reason why we have Swiss Army Knives instead of carrying around separate tiny knives, screwdrivers, etc.: it's much easier to carry them all around in one package. Similarly, there's a reason why we might care about Colorado-Rally-Worthiness even if it isn't necessary for any particular task that might be accomplished. It's easier to answer 'Is this Colorado-Rally-Worthy?' than 'What is its horsepower, cargo capacity, undercarriage, and mileage per tank?' The Swiss Army Concept is a concept that is not important in itself, but that provides an economical way of summing up several other concepts that are important in themselves. So my claim will be that knowledge is a Swiss Army Concept, at least with respect to value for practical reasoning; it provides an economical summation of the concepts that turn out to be important for practical reasoning from the various standpoints.²

2

To illuminate the multiple standpoints on practical reasoning, let us look at Hawthorne's account of the lottery problem. The lottery problem is this. We are generally unwilling to ascribe advance knowledge that a particular ticket in a fair lottery will not win, but we may be willing to ascribe knowledge of propositions that entail that this ticket will not win. We may be willing to say that you know that you will not be able to afford to go on an African safari next year even though you own a ticket for a lottery whose prize is more than the cost of a safari.³ The lottery problem can stand in for much reasoning about the not quite certain future or present. To use some examples of Vogel's (1990), we may be willing to say that you know where your car is but unwilling to say that you know it is not one of the few cars stolen each

² I suspect that knowledge will turn out to be a Swiss Army Concept for other purposes as well. For instance, Kaplan (1985) argues that knowledge is not important to inquiry, because once we have determined that a belief is justified we have settled that the believer's inquiry is exempt from criticism. We might take this as one standpoint on inquiry. From another standpoint, we might care about whether the inquiry is successful; and then we will care about whether the knowledge results in true belief. These will be seen to resemble the first two standpoints on practical reasoning that we will discuss, with knowledge providing an economical combination of truth and justification.

³ Compare Lewis's (1996) example of Bill, who we know will never be rich because he spends all his money on lottery tickets.

day; we may be willing to say that you know where you will be next week but unwilling to say that you will not be one of the few apparently healthy people who will suddenly drop dead before then. Hawthorne's view is that we can explain these ascriptions by defining knowledge in terms of suitability for practical reasoning.

The idea is this. Define knowledge so that a belief that *p* does not amount to knowledge in a certain practical environment iff it is not acceptable to use the belief as a premise for practical reasoning in that environment. Then, in the practical environment in which it is relevant, you will know that you won't be able to afford the safari; and you won't know that your ticket won't win the lottery in the practical environment in which *that* belief is relevant. In fact, in such an environment you wouldn't know that you can't afford a safari either. Thus Hawthorne's account seems to explain our judgments of knowledge while providing an important role for knowledge in practical reasoning.

To flesh out Hawthorne's argument, consider a practical environment in which you might want to use as a premise the belief that your lottery ticket won't win. Someone offers you a ticket for a 10,000-ticket lottery with a \$5,000 prize, at the price of one penny. Let us suppose that, in fact, it will turn out that this ticket loses. Still, you shouldn't reason as follows:

Argument A

(1) If I buy this lottery ticket, it will lose.

So,

(2) I'll be out a penny.

So,

(3) I should not buy this ticket.

This is terrible reasoning. You shouldn't dismiss out of hand the possibility that the ticket will win when you're considering whether to buy it; that undermines the whole point of deliberating about whether to buy the ticket. What makes the reasoning terrible is not simply that in this case it leads to a conclusion that does not maximize expected utility, but that it rules out a possibility that ought to be taken into consideration in this deliberation.⁴

Similarly, Hawthorne points out that it would be 'intuitively awful' to reason as follows:

Argument B

(4) I will not have enough money to go on an African safari next year.

So,

(1) If I buy this lottery ticket it will lose.

⁴ Thanks to a referee for a clarification here.

So,

- (3) I should not buy this ticket (2004: 174 (my numbering)).

Accordingly (1) and (4) are not acceptable premises in this practical environment, and in this environment you know neither that the ticket won't win nor that you won't be able to afford the safari.

Note that if there is something wrong with these arguments, it is with the premises, (1) or (4) respectively. For Argument A, given that the ticket costs a penny and that it is worse to lose a penny than not to, the conclusion follows as inexorably as a practical conclusion can follow. For Argument B, given the auxiliary premise that a winning lottery ticket yields enough money to go on an African safari, the conclusion is similarly inexorable. In each case, we may assume that the auxiliary premises are beyond question.⁵ For the most part, I will consider only syllogistic arguments whose only possible flaw is in the acceptability of a certain premise, the better to focus on the epistemic properties of those premises.⁶ For convenience's sake, I will call such arguments *proper* arguments. If an argument is proper and the premise in question is acceptable, then the conclusion follows. (In Section 5, I will briefly consider probabilistic practical arguments and the idea that practical reasoning should not take the form of a practical syllogism at all.)

Consider now a practical environment in which you might want to exploit your belief that you won't be able to afford a safari in a more natural way. You have bought the lottery ticket, and you are now in a bookstore buying a guidebook for next year's vacation. Hawthorne argues that it is acceptable to reason as follows:

Argument C

- (4) I will not have enough money to go on an African safari next year.

So,

- (5) I will have no use for a guidebook to Africa.

⁵ Thanks to Alan Millar for pointing out the necessity of these auxiliary premises.

⁶ It might be thought that Argument B has another flaw, in that it would be an odd way of arriving at the decision not to buy the lottery ticket. No one will actually start with the premise that she will not be able to afford the safari and conclude that her ticket won't win. But this does not amount to a flaw in the argument itself, any more than a theoretical deduction becomes invalid merely because it is strange to carry it out in particular circumstances.

Possibly one could give an account on which (4) is an acceptable premise and Argument B is proper, but on which nevertheless the action prescribed by the argument should not be carried out. On such an account one would not always be able to exploit the conclusions of proper practical reasoning with acceptable premises. This would be the practical analogue of denying deductive closure for knowledge, so that one cannot always gain knowledge by performing valid deductions from true premises. However, as an explanation of the problem with Argument B, this approach seems less promising than an account on which (4) is unacceptable.

So

(6) I should buy the local destination guide (2004: 177).⁷

Accordingly, on Hawthorne's account, you do know (4) on this occasion. You can know propositions about the future without ruling out lottery-like alternatives, so long as the decisions you are making do not require you to take those alternatives into account. It is not merely that winning the lottery is improbable enough that the local guide has a higher expected utility than the guidebook to Africa; in this case, we might think, winning the lottery is a remote enough possibility that you need not consider it at all in your deliberations. (See Section 5 for a consideration of the probabilistic version of this deliberation.)

So Hawthorne argues that our intuitive judgments of knowledge line up reasonably well with cases in which the subject's belief is an acceptable premise for practical reasoning. (There are many complications to this view, but we can leave them aside.) The question remains, however: What is it for a premise to be acceptable? When we look more closely at this question, we will see that there is no single way of asking it such that the acceptable premises are exactly the known ones.

Here is one possible answer: We care about whether practical reasoning will turn out well for us. So proper practical reasoning from acceptable premises should turn out well for the reasoner. But in the practical environment in which you have been offered the lottery ticket, the reasoning that will in fact turn out the best for you is the reasoning that leads you to decline the ticket. *Ex hypothesi* the ticket will lose, and if you bought it you would have been out a penny. This produces the uncomfortable result that Arguments A and B are both acceptable arguments, and (1) and (4) are both acceptable premises. You can reason from the premise that your ticket will lose or that you will not be able to afford a safari. In fact, the premises that actually yield the best results given proper practical syllogisms are all and only the true premises. This line of thinking shows that for practical reasoning it is important to have true beliefs.

There is an obvious way to avoid the uncomfortable result that Arguments A and B are acceptable. We can say that when we ask whether practical reasoning is acceptable, we are not asking about the practical reasoning that will in fact lead to the best outcome. From this standpoint, we view acceptable practical

⁷ Hawthorne may hold that Argument C is proper only if you have forgotten that you have a lottery ticket; if you are thinking about your ticket then you are not in a position to know (4). On the other hand, Hawthorne opposes an account on which invoking (1) in the argument from (4) to (3) destroys knowledge of the premise by making a new possibility salient; he argues that the possibility might not become salient for someone who is sufficiently dogmatic. So Hawthorne might not think that knowledge of (4), if you do have it, is destroyed when you remember the existence of the ticket. Many other ins and outs of this debate are discussed in his 2004. In my text we may stipulate that you are not thinking about your lottery ticket.

reasoning as reasoning that is not vulnerable to criticism, that is not feckless or rash or overcautious.⁸

From this standpoint, Arguments A and B clearly are vulnerable to criticism. The payoff for winning the lottery is so much higher than the cost of the ticket that you are not entitled to ignore the small chance that the ticket will win. So (1) and similarly (4) are not acceptable premises in this practical environment. Arguments A and B may be criticized even if they in fact turn out to save a penny. In the practical environment of the bookstore, however, you are entitled to use (4) as a premise. It would be feckless to refuse to buy the local guidebook because you claimed not to know that you wouldn't be able to afford a safari; this is not the sort of decision that should be thrown into doubt because of a lottery ticket. So this standpoint yields the result that Hawthorne desires: (4) is an acceptable premise in the practical environment of the bookstore but not of the lottery purchase.

The problem is that from this standpoint (4) is *always* an acceptable premise in the practical environment of the bookstore. It is acceptable even when it is false. Suppose that, in the bookstore, you refuse to follow Argument C because of the remote chance that you might win the lottery, and then you do go on to win the lottery. Failing to follow Argument C would be as feckless as ever; it would be through luck that your faulty reasoning produced the best outcome for you. Conversely, suppose you reason as in Argument C, buy the local guidebook, and go on to win the lottery. Was your original reasoning acceptable? From this standpoint, yes. If Argument C is beyond criticism in the case in which you don't win the lottery, it is beyond criticism in the case in which you do. You were not being feckless or dogmatic in thinking that you would not be able to afford a safari. That the right reasoning did not lead to the best outcome in this case is simply epistemic bad luck (though financial good luck).

From the standpoint that concerns itself with whether your practical reasoning can be criticized, what is important for practical reasoning is how well justified your beliefs are.⁹ The practical environment matters here: It determines how

⁸ Compare Hawthorne's contrast of the reading of 'should' on which it is obvious that a premise like (1) should not be used in practical reasoning with the possible reading of 'should' on which what you should have done is what would in fact have led to the best outcome (2004: 175 n. 33).

⁹ Exactly what conception of justification is at issue will depend on what notion of criticism of the practical reasoning is at issue. If criticism is warranted only when the believer has made some sort of culpable mistake in evaluating the evidence, then the relevant notion of justification will consist in not having made a culpable mistake in evaluation; if criticism of the reasoning is warranted whenever one's premise is not adequately supported by the evidence, then the relevant notion of justification will consist in the evidence's support for a proposition, without regard to the procedure through which one arrived at the proposition. For instance, if one blamelessly makes a mistake in calculation, the resulting practical reasoning can be criticized in the second sense but not in the first, since the reasoning is based on a premise that the evidence does not support, but the mistake made in evaluating the evidence was not culpable. And the premise one arrives at is justified in the second sense but not the first, since it is not supported by the evidence but there is

much justification you need for your belief to be acceptable. Nevertheless, this standpoint does not establish the importance of a factive property of beliefs. Unless the practical environment calls for absolute certainty, it will be the case that acceptable reasoning may proceed from false premises. *A fortiori*, acceptable reasoning may proceed from premises that are not known. So whether a belief counts as knowledge is not important in itself from either standpoint.

3

The argument concerning the lottery case can be applied to any practical reasoning that calls for an instantaneous decision. If it is important whether the subject's reasoning in fact leads to the best outcome, we should be concerned about whether her premises are true. If it is important whether the subject's reasoning is beyond criticism, we should be concerned about whether her premises are well enough justified given her practical situation.

Most decisions, however, are not instantaneous. To accomplish anything we need to be able to make a plan and carry it out over an extended period. In such a case success requires more than just having a true belief at any one point. So when we look at practical reasoning over an extended period of time, properties of the belief other than its truth and justification may be important.

Consider this example of Williamson's (2000: 62): a burglar is ransacking a house looking for a diamond. He knows that there is a diamond in the house, so he continues to look all night even when he fails to find it. If, on the other hand, he had a Gettiered belief that there was a diamond in the house, he might not continue to look all night. Suppose that he inferred that there was a diamond in the house because he had been told that there was one under the bed, when in fact the diamond was in the drawer. Then he would give up after failing to find the diamond under the bed. He has a justified true belief when he sets out to look for the diamond, but it will not be enough to keep him looking long enough to have a good chance of finding it.

This example might be taken to show that in some cases knowledge *is* the important concept for evaluating practical reasoning.¹⁰ But looking at extended plans will still not reveal any one standpoint from which the acceptable premises for practical reasoning are exactly the things we know. There is still a split

no culpable mistake in the procedure used to arrive at it. (Thanks to a referee for helping to clarify this point.)

¹⁰ Williamson himself immediately uses the example to argue that 'The burglar knew there was a diamond in the house' has more explanatory power than 'The burglar had a true belief that there was a diamond in the house.' Later (Williamson 2000: 9), he argues that knowledge is more stable than true belief, with reference to the *Meno* question about the value of knowledge.

between premises that will actually lead to successful reasoning and premises that lead to reasoning that is beyond criticism, and from neither perspective is knowledge what we need.

Consider what it takes for a belief to serve as a premise for practical reasoning that will produce a plan that actually succeeds. The burglar might reason as follows:

Argument D

(7) There is a diamond somewhere in this house.

(8) If I burgle the house, I'll get the diamond.

So,

(9) I should burgle the house.

For Argument D to be proper, it must be right for the burglar to want to find the diamond (so we must discount the wrongness of burglary itself!), and the value of the diamond must outweigh the costs of burgling the house. Then, for (9) to follow, both (7) and (8) must be acceptable premises. We can suppose that (8) entails that the burglar won't get caught (if you like, make it 'I'll get the diamond and won't get caught'), so among other things this entails that it is an acceptable premise for the burglar that he won't get caught while he's looking for the diamond. Let us focus on the acceptability of (7), assuming for the sake of argument that (8) is acceptable from whatever standpoint.

Consider two burglars, Moriarty and Raffles. Both initially have a true belief in (7), so both initially follow Argument D. This initial reasoning, however, is not enough to ensure that they get the diamond. The successful burglar must preserve his belief in (7) until he actually finds the diamond. So if Moriarty happens to give up his belief in (7) once he has looked under the bed while Raffles will preserve his belief in (7) until he finds the diamond, then Raffles will find the diamond and Moriarty will not. What is important is not only truth but stability of belief.

Note, however, that Raffles will find the diamond if his belief is stable, no matter how unjustified he may be in preserving his belief. Suppose that Raffles's informant told him that the diamond was under the bed, but Raffles persists in believing that there was diamond somewhere even after he has found that it is not under the bed. Or suppose that Raffles never had any evidence for believing there is a diamond in the house, but has nevertheless got the idea into his head and will not give it up until he has thoroughly searched the house. It will not matter from the standpoint of actual success that Raffles does not know (7). So long as Raffles's belief is true and persistent, he will find the diamond.¹¹

¹¹ Compare Kvanvig's objection to Williamson in terms of beliefs that are fixed by nonevidential factors (Kvanvig 2003: 15).

From the other standpoint, suppose that we are concerned with whether the burglars' reasoning can be criticized. To consider whether it can be criticized at the outset is no more than to consider how well justified their beliefs are, as discussed in the previous section. To take the extended view, we should ask whether their reasoning yields a plan that can be completed without exposing them to criticism. This will be so if they not only are justified in believing (7) at the outset, but will remain justified in believing (7) for the duration of the plan. Suppose that Raffles has been told that there is a diamond in the house, and Moriarty has been told that there is a diamond in the house and it is definitely under the bed.¹² Then Raffles ought to follow Argument D and stick to his plan of searching the house until he finds the diamond. To do otherwise would be to give up too easily. Moriarty, on the other hand, ought to follow Argument D at first; but when he finds no diamond under the bed, he ought to abandon it. His reason to believe its premise (7) has been undercut. To keep searching for the diamond would be stubborn.

From this standpoint, then, an acceptable premise is one that is justifiably believed and that is likely to stay justifiable as new evidence comes in. This rules out some Gettier cases, as in the Moriarty case just described. It does not, however, rule out all cases of false belief. Suppose that there is no diamond in the house. For Raffles, premise (7) will remain justified for as long as it takes to search the house; so his reasoning will remain beyond criticism until he completes it. Then Raffles will be seen to have failed, through bad luck. But from the standpoint of criticism an acceptable premise is not necessarily one that, combined with a proper argument, will lead to a successful plan. This standpoint on practical reasoning shows stable justification, not knowledge, to be important in itself.

4

There is one standpoint from which the acceptable premises for practical reasoning always have both truth and a justification-like property. We can ask: Would this successful practical reasoning still have succeeded if circumstances had been different? From this standpoint, an acceptable premise is not only true but counterfactually true.¹³ Hence, when we are concerned with the counterfactual

¹² The idea is that once Moriarty finds that there is no diamond under the bed he should come to doubt his original information. Williamson's original description of the case does not seem to rule out that the burglar, finding no diamond under the bed, should conclude that it was moved and look elsewhere in the house. We can stipulate that in the current situation that is not the case.

¹³ On the most plausible reading of the questions that this standpoint is concerned with, what matters is not the counterfactual success of reasoning from the exact same premise but counterfactual success of reasoning from whatever premises the agent would have come to believe in different

success of practical reasoning, the important property of belief is safe truth, where a belief that p is safe iff in nearby possible worlds the agent believes that p only if p is true (Pritchard 2005: 71).¹⁴ Besides Pritchard, Williamson (2000: 123 ff.) and Sosa (2000) have posited safe truth as a requirement on knowledge. Nevertheless, it does not seem to me that this standpoint establishes the importance of knowledge in itself for practical reasoning.

For one thing, safe truth is arguably neither necessary nor sufficient for knowledge. Against sufficiency, knowledge also needs some sort of internal justification, for the beliefs of BonJour's clairvoyant (BonJour 1985: ch. 3) are safe. In every nearby world in which the clairvoyant arrives at his belief by the same method, it is true.¹⁵ Similarly, there may be cases of safe true belief that are nevertheless Gettier cases. Suppose that, in the barn-façade case (Goldman 1976), although the county is full of barn-façades, some law or physical fact makes it impossible to build a barn-façade in Yoder's field. Yoder's field is so swampy that a barn-façade could not stand up without the additional support provided by a real barn's other walls. Then the belief that Yoder's field contains a barn is arguably safe; in the nearby worlds in which I believe there is a barn in Yoder's field, there is a barn in Yoder's field. Still, the nearby barn-façades keep us from knowing that there is a barn in Yoder's field.

This may not be a definitive refutation of the sufficiency of safety; perhaps safety can be defined so that the possibilities of error are near enough to make the beliefs unsafe. We might say that 'There is a barn in Yoder's field' belongs to a class of closely related propositions ('There is a barn in Stoltzfuss's field,' etc.), such that in nearby possible worlds I consider these propositions

circumstances, by the same method that she could come to believe the premise in question in the actual circumstances. So if success depends on the agent's successfully naming a random animal that is produced for her, and she reasons 'That's a duck, so I'll say it's a duck,' her premise 'That's a duck' could easily have been false (if a chicken had been produced instead), but she would still have been reasoning from a true premise (because then she would have reasoned from 'That's a chicken'). Hence 'counterfactual truth' should be taken to mean counterfactual truth of the premise arrived at by the same method.

¹⁴ Sosa (2000: 14) defines safety in terms of the counterfactual $B(p) \rightarrow p$, where $B(p)$ stands for belief that p . Since the antecedent of the counterfactual is true in the actual world, this is more clearly understood in terms of nearby possible worlds. If we define 'counterfactual success' so as to mean 'success in every world up to the nearest world in which the current reasoning would not be successful,' then instead of safe truth the important property will turn out to be sensitive truth, as in the analysis of knowledge in (DeRose 1995): If p were false, then the agent would not have believed that p . But it is hard to see why that particular definition of counterfactual truth would be of practical interest, unless we already assume that practical premises should be epistemically sensitive. And if we make that assumption, we have not succeeded in grounding the importance of knowledge in its role in practical reasoning.

¹⁵ The clairvoyant's belief is modally unstable in this way: If he were not clairvoyant but merely believed things that popped into his head, his beliefs would be false even though things would seem the same to him. But if worlds in which an actually reliable faculty fails count as close for the purpose of safety, then it seems as though safety must be a matter of internal justification; otherwise worlds in which perceptual faculties fail could also count as nearby, and perceptual knowledge would be considered unsafe.

and believe them falsely. Yet in other cases, it does not destroy knowledge if in nearby possible worlds we falsely believe closely related propositions. If I always identify geese as ducks, I can still know that there is a swan in a lake, even though in a nearby world in which the swan is replaced by a goose I would falsely believe the related proposition 'There is a duck in the lake.' To analyze knowledge as safe true belief, we would need a principled reason that one of these alternative propositions should count against safety and the other not. The prospects for such a reason seem poor, as do the prospects of analyzing knowledge as safe true belief, and thus of establishing the importance of knowledge from the standpoint from which we care about the safe success of practical reasoning.

Against the necessity of safety for knowledge, see Comesaña (2005) and Neta and Rohrbaugh (2004). These authors present interestingly different diagnoses of why knowledge can fail to be safe. Comesaña argues that safety requires reliable reliability, whereas knowledge merely requires reliability. This suggests that safety provides additional value over and above whatever values are folded into knowledge ascriptions; for whatever value mere reliability has, reliable reliability has more of it. Neta and Rohrbaugh (2004: 404) argue that knowledge is an important cognitive achievement, and that earned achievements are in general not safe; an achievement that is earned despite the possibility of failure is nevertheless earned.

This leads to the more important problem with the attempt to establish the practical importance of knowledge qua safe true belief. It is not clear why safe true belief would be more practically valuable than true belief alone. An achievement is earned even if it is not earned safely, so it may not matter overmuch whether our reasoning is not only successful but counterfactually successful. Surely what we care about most is success in the actual world. If we do point out a flaw in a premise that actually led to success, it is more relevant if the agent could have been criticized for relying on the premise than if the premise could have failed in some counterfactual situation. (Of course these often go together.)

In addition, the safety of the premises does not seem to have anything to do with the practical environment. Premise (4) seems just as safe in the context of Argument B as in the context of Argument C. Whether the decision at issue concerns buying a lottery ticket or a guidebook, the possibility that you can afford the safari is just as remote. Hence even if there is a standpoint from which we value safe success in our practical reasoning, the practical premises that are acceptable from this standpoint will not be just the known ones, unless we do know that our lottery tickets will lose. We might escape this consequence by defining remoteness and safety in terms of the possibilities that you ought to take into account, given the practical decision you are making. But this definition builds in the importance of your reasoning's being immune from criticism, which is motivated from a different standpoint from the one that might motivate safe truth.

Another way to try to bring truth and justification together is to argue that, in order to be immune from criticism, practical reasoning must proceed from true premises. Practical reasoning from false premises, or even from premises that are not known, is as such open to criticism. Williamson makes an argument that can easily be extended to yield this conclusion: Our evidence is identical to our knowledge, and rationality requires respecting the evidence (2000: ch. 9, revising Williamson 1997). If this were true, then reasoning from premises that were not known would always be irrational. Williamson bolsters this position by arguing that we are not necessarily in a position to know what it is rational for us to believe. If I see a table before me, it is rational for me to believe in its existence, because it is part of my evidence that I see (and know) that there is a table before me. If I am hallucinating a table, then it is not part of my evidence that I see and know that there is a table, so it is not rational for me to believe in a table. Yet the hallucinator may be internally indistinguishable from the person who sees the table. This argument can easily be extended from theoretical to practical rationality.

Even granting Williamson's analysis of the hallucination and similar cases, it will not follow that rational reasoning always proceeds from known or even true premises if we are not antecedently committed to the importance of knowledge. Consider Argument C, the bookstore argument, as it is made by two exact duplicates, one holding a winning ticket and one holding a losing ticket. Is there a sense in which the loser is deliberating rationally and the winner is not? The only difference between them concerns the result of the lottery drawing, an event that will take place after they have acted on their deliberations. In the hallucination case we could argue that the person who saw the table has direct access to the existence of a table, which the hallucinator lacks, and that this distinguishes their reasoning. But in the lottery case it is implausible that the loser has any direct access to the fact that she will not be able to afford a safari.¹⁶ There needs to be some other relevant difference if we are to conclude that the loser deliberated rationally and the winner did not.

It is true that the loser's deliberation begins from a true premise and the winner's deliberation from a false one. But this will not provide a basis for criticizing the winner's deliberation and not the loser's. If we criticize the winner for reasoning from a false premise, she may say, 'Yes, but I had every reason to believe it was true. Should I instead have reasoned from the true premise that I would be able to afford a safari? That would have led to a better outcome, but it would have been bad reasoning.' The actual truth of

¹⁶ Though I do not have space to explore the possibility here, this case also calls into question Williamson's argument that all our knowledge serves as evidence; if we have inferential knowledge about the future, that will be a good candidate for knowledge that is not itself evidence. (See Brian Weatherson's discussion of inductive knowledge as a counterexample to the knowledge-as-evidence thesis, at <http://tar.weatherson.org/2003/08/26/evidence-and-knowledge/> [accessed Sept. 9, 2006].)

her premise is irrelevant from the standpoint of criticizing her deliberation; to make it relevant, we must adopt the standpoint from which we care about the actual success of her deliberation. And, as we have seen, from that standpoint justification does not matter. Similarly, to insist that deliberation is not rational if the premises are not known is to assume the value of knowledge for practical reasoning. It will not help us use practical reasoning to establish the importance of knowledge.

5

So far the model of practical reasoning discussed has been the practical syllogism. The premises of these syllogisms are categorical statements, and the conclusion is an action that is called for given those premises. But this is not the only way to approach practical reasoning. Perhaps correct practical reasoning proceeds from premises that assign probabilities to certain outcomes to conclusions that maximize expected utility given those probabilities (or respond to those probabilities in some other way). Hence when presented with the lottery ticket one could reason thus:

Argument E

- (10) The chance that this ticket will win is 1 in 10,000.
 - (11) The ticket costs a cent and is worth \$5,000 if it wins. The expected value of the ticket is 50 cents (\$5,000 divided by 10,000).
- So,
- (12) if I buy the ticket I'll have an expected gain of 49 cents (the ticket's expected value minus its cost).
 - (13) So I should buy the ticket.

And in the bookstore one should reason as follows:

Argument F

- (14) I'll only be able to afford an African safari if this ticket wins the lottery.
 - (10) The chance that this ticket will win is 1 in 10,000.
- So,
- (15) there's only a 1 in 10,000 chance that the African destination guide will be any use, and its expected value is quite small.
 - (16) There's a 9,999 in 10,000 chance that the local destination guide will be of use, and its expected value is reasonably large.
- So
- (17) I should buy the local destination guide.

The question now is, what property must the critical premise (10) have in order to be an acceptable premise for this practical reasoning?

From the standpoint of actual success, (10) is *never* the best premise. Your practical reasoning will actually succeed if you assign probability 1 to whatever turns out to be true. Assigning any lower probability will sometimes lead you to forgo some course of action that would have turned out to be successful. We've already seen this; in the case in which your ticket loses, Argument E leads you to lose a penny, whereas if you assign probability 1 to 'This ticket won't win' you won't buy the ticket and you'll save a penny.

What about the standpoint that is concerned with whether your reasoning can be criticized? Here it matters how we interpret probability. If (10) means that there is an objective chance of 1 in 10,000, then it will not be the right kind of premise to guarantee immunity from criticism. You may be immune from criticism for employing Argument E or F even if the lottery is rigged so that there is no objective chance that your ticket will win (or lose, respectively), so long as your evidence indicates that the lottery is fair. And you may be criticized for using these arguments even if the objective chance that your ticket will win is 1 in 10,000, if your evidence indicates that the odds are different (for instance, if an ordinarily reliable source tells you that there are more or fewer tickets in the lottery).

If probabilities are interpreted more subjectively, however, then it may be that reasoning from true probabilistic premises will be immune from criticism. I do not mean the view of subjective probabilities on which any assignment of credences to propositions is permissible so long as it is consistent with the axioms of probability. (On such a view, the question of which premises about probabilities were acceptable would not arise.) Rather, I mean a conception on which the available evidence determines some probability for a proposition, which will be the credence that a believer ought to give that proposition based on that evidence.¹⁷ Suppose we interpret (10) to mean that on the available evidence the appropriate credence in 'This ticket will win' is 1 in 10,000. If you follow Argument E or F because the evidence indicates that there is a 1 in 10,000 chance that your ticket will win, you will be beyond criticism, for then you are conforming your actions to the evidence. The 'because' is necessary here; if, without evaluating the evidence, you guess that your ticket has a 1 in 10,000 chance of winning, you can be criticized even if the evidence does support that. Believing (10) because the evidence supports it means being justified in believing it, so it appears that (10) is an acceptable practical premise whenever it is true and believed with justification.

¹⁷ Other subjective interpretations of probability are possible. However, if on these interpretations (10) does not amount to 'On the available evidence, it is appropriate to give a credence of 1 in 10,000 to this ticket's winning,' then reasoning from (10) will not be immune from criticism. The interpretation discussed in the text is the friendliest to the idea that knowledge of probabilities is valuable for practical reasoning.

This may even be a standpoint from which the acceptable premises for practical reasoning are all and only the (probabilistic) propositions that are known. On the subjective interpretation, if (10) is true and believed because it is true, then we cannot criticize any proper practical reasoning that uses it as a premise. Arguably in all of these cases (10) will be known. If the conception of justification at issue is that a belief is justified whenever the evidence supports it, and the conception of probability at issue is that a proposition's probability depends on the degree to which the evidence supports the proposition, then whenever the odds of the ticket's winning are 1 in 10,000 the evidence will support belief in those odds, and thus belief in (10) will be justified.¹⁸ Since this justification for believing (10) entails the truth of (10), any belief in (10) based on this justification will amount to knowledge. Thus from this point of view the acceptable practical premises might be only those propositions about probability that are known.

However, this would only be a small victory for proponents of the intrinsic value of knowledge. To begin with, from this standpoint knowledge is not uniquely valuable, since justification, knowledge, and truth are so intimately connected. Not only does having a justified belief about a probability entail knowing what the probability is, but whenever a probability statement is true a justification is available for believing it. You need only properly evaluate the evidence, and you will with justification believe that the probability is what it actually is.

More important, this argument restricts the domain of knowledge. If knowledge is important because known premises can be used in probabilistic practical reasoning, then only probabilities can be known. Very rarely does the evidence make it appropriate to reason as though a proposition had probability 1. If our concept of knowledge is the one that is important for this kind of practical reasoning, we will almost never empirically know categorical propositions. Certainly we will not know anything about the not quite certain future or present, as in the lottery case or Vogel's car theft case. In addition, this knowledge is essentially knowledge of justification, since it has to do with the credence that the evidence supports. Probabilistic practical reasoning will not establish the importance of knowledge of ordinary facts. The critical concept here is much closer to the concept of justification that we have already seen to be important in itself.

¹⁸ On other conceptions of justification, justified belief in (10) can come apart from its truth. For instance, if justification means arriving at belief responsibly and blamelessly, it may be possible to responsibly and blamelessly arrive at an incorrect belief concerning the extent to which the evidence supports a proposition, for instance by an error in calculation. (See n. 9 above; thanks to a referee for raising this point.) In this case there will be no single standpoint from which acceptable practical premises need be both justified and true, and thus the acceptable premises will not be those propositions that are known. As with different interpretations of probability (see n. 17 above), the interpretation of justification discussed in the text is the friendliest to the idea that knowledge of probabilities is in itself valuable for practical reasoning.

6

We have seen that there is no single standpoint on practical reasoning from which knowledge of categorical propositions is important. Instead, from various standpoints, it is important that one's premises be true, justified, persistent, stably justified, and safe. We might wonder, then: If knowledge is not important for practical reasoning, why do we talk about knowledge at all when we are concerned with practical matters? Why not simply talk about the important things? Yet in fact people talk about what people know much more often than about, for instance, what they are justified in believing.¹⁹

To answer, think of the Swiss Army Knife metaphor. A Swiss Army Knife is useful to carry around when you do not know exactly what task you will be faced with. If you are faced with a task that requires a knife, a screwdriver, a corkscrew, or a bottle opener, you will have what you need; and you will not face the awkwardness of having to carry around four separate tools. Analogously, when evaluating someone's epistemic situation you may not want to know which standpoint you will eventually want to take on their practical reasoning. If you say 'S believed truly that p ,' and it becomes important to figure out whether S's reasoning should be criticized, then you won't have said anything helpful. So it will be convenient to have a quick way of expressing all these different concepts that may be important from the different standpoints on S's practical reasoning. If you say 'S knew that p ,' your audience knows that S should not be criticized for reasoning from p , and that any proper argument S made with p as a premise succeeded, and (if applicable) that S was in a position to retain her belief that p long enough for her plans to succeed, etc. Even though the concept of knowledge is not needed for answering any one of these questions, it provides an efficient way of expressing an answer to all of them.

This provides a little bit of progress toward the question about why we use a concept of knowledge that rules out Gettier cases. In some Gettier cases, a justified true belief does not count as knowledge because the belief or justification is unstable. If, like Williamson's burglar, you are about to discover countervailing but misleading evidence, your belief will not remain immune from criticism long enough for your plan to remain immune from criticism throughout its execution. If you are about to forget a belief, you may not be able to carry out any plans

¹⁹ One of the criticisms that may be made of the argument of Kaplan (1985) is that it leaves it mysterious why anyone would ever have thought knowledge important (in its contemporary guise, in which it can be based on fallible evidence).

based on it.²⁰ When we call a belief knowledge, we guarantee it to be satisfactory from these standpoints.

But this is not quite enough as an account of why we do not count Gettier cases as knowledge. The most important standpoints on practical reasoning are surely whether it leads to actual success and whether it is immune from criticism. So the most important properties of beliefs from a practical standpoint will be truth and justification. If the importance of knowledge derives mostly from the importance of justification and truth, why are there so many cases in which we judge that justified true belief is not knowledge?²¹

My answer comes from extending the Swiss Army Knife metaphor. If we become used to Swiss Army Knives as the way to carry around the tools we need, we may come to see them as valuable in themselves (even though they are not). Then someone who has a knife, a screwdriver, a corkscrew, etc. may still be seen as lacking something important. We will have brought ourselves to care about not only the individual tools, which are what we really need, but also about how they are connected. Separate tools will not seem as satisfactory as the same tools in a single package.

Analogously, when we say 'S knows that *p*,' knowledge seems to be a unified concept that may be important for its own sake. Even if knowledge is important primarily because of the importance of truth and justification, it seems as though what is important is that the truth and justification be combined in the right way. Typically a justified true belief is one in which whatever makes it justified is also whatever makes it true. So this will seem to be characteristic of knowledge. A belief that lacks this characteristic, in which justification and truth are somehow mismatched, will be seen as lacking the organic unity that typical knowledge has. Even if justification and truth are independently important, from different standpoints, when we use a single word to ascribe them together what seems important is that they come together in the right way. Hence when justification and truth are mismatched we will have a Gettier case, where we are reluctant to ascribe knowledge.

But in fact there is no standpoint from which the mismatch of knowledge and justification is particularly important for practical reasoning, except insofar as it subverts the temporal or modal stability of belief, truth, or justification. Truth, justification, and stability will be important in themselves for various ways of looking at practical reason. Knowledge is important for practical reasoning only

²⁰ Marc Moffett has devised examples where temporal instability might make us reluctant to ascribe knowledge; see http://rationalhunter.typepad.com/close_range/2004/05/a_thought_exper.html (accessed Sept. 11, 2006).

²¹ Weinberg, Nichols, and Stich (2001) have cast doubt on the universality of Gettier intuitions, but what requires explanation is why anyone at all has strong Gettier intuitions.

insofar as it combines these other qualities; and when we demand that these qualities be brought together in an organic whole, this demand does not yield anything that we need for a belief to be a good practical premise.²²

REFERENCES

- BonJour, L. (1985). *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.
- Comesaña, J. (2005). 'Unsafe Knowledge,' *Synthese*, 146: 395–404.
- DeRose, K. (1995). 'Solving the Skeptical Problem,' *Philosophical Review*, 104: 1–52.
- Goldman, A. (1976). 'Discrimination and Perceptual Knowledge,' *Journal of Philosophy*, 73: 771–91.
- Hawthorne, J. (2004). *Knowledge and Lotteries*. Oxford: Oxford University Press.
- Kaplan, M. (1985). 'It's Not What You Know that Counts,' *Journal of Philosophy*, 82: 350–63.
- Kvanvig, J. (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge, MA: Cambridge University Press.
- Lewis, D. (1996). 'Elusive Knowledge,' *Australasian Journal of Philosophy*, 74: 549–67.
- Neta, R., and Rohrbaugh, G. (2004). 'Luminosity and the Safety of Knowledge,' *Pacific Philosophical Quarterly*, 85: 396–406.
- Pollock, J., and Cruz, J. (1999). *Contemporary Theories of Knowledge*. Totowa, NJ: Rowman and Littlefield.
- Pritchard, D. (2005). *Epistemic Luck*. Oxford: Oxford University Press.
- Sosa, E. (2000). 'Skepticism and Contextualism,' in J. Tomberlin (ed.), *Philosophical Issues*, 10: 1–18.
- Vogel, J. (1990). 'Are There Counterexamples to the Closure Principle?' in M. Roth and G. Ross (eds.) *Doubting: Contemporary Perspectives on Skepticism*. Dordrecht: Kluwer.
- Weinberg, J., Nichols, S., and Stich, S. (2001). 'Normativity and Epistemic Intuitions,' *Philosophical Topics*, 29: 429–60.
- Williamson, T. (1997). 'Knowledge as Evidence,' *Mind*, 106: 717–41.
- (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.

²² Earlier versions were presented at the *Central Division Meeting of the American Philosophical Association* in April 2005 in Chicago, with commentary by Mylan Engel, and at the *Epistemic Value Conference* at the University of Stirling in August 2006, with commentary by Igor Douven. Thanks to the commentators and to members of the audience on those occasions, particularly to Berit Brogaard, Christian Pillar, Dennis Whitcomb, Stephen Green, and Duncan Pritchard. Thanks also to Ram Neta, Elijah Milgram, Matt McGrath, Keith DeRose, A. P. Martinich, Alan Millar, and two anonymous referees for helpful comments on this and related matters; apologies to anyone I have neglected to include. This research was supported in part by a Texas Tech New Humanities Faculty Grant.

Pragmatic Encroachment and Epistemic Value

Pascal Engel

1. PRAGMATIC ENCROACHMENT AND THE VALUE OF KNOWLEDGE

When in the *Meno* (97a–c) Socrates asks whether knowledge is more valuable than true belief, the notion of value which he has in mind seems clearly to be that of *practical value*. In 97a he asks whether ‘good men’ or ‘men of value’ (*agathoi androi*) will be ‘useful’ (*ophelimoí*) and when in 97c he raises his famous question about the difference between having a true belief about the road to Larissa and knowing the road to Larissa, this question is directly couched in terms of what is correct with respect to practice (*pros orthothèta praxeôs*) and the example is clearly meant to ask something about the respective roles of belief and knowledge in guiding our actions. Contemporary approaches to the problem of the value of knowledge, however, have investigated other senses of the notion of value or worth of knowledge, in terms of various notions of epistemic virtue (Sosa 2007; Greco 2002; Zagzebski 1996), or in terms of moral appraisal (Brady 2006). Although many of these accounts involve the idea that the worth in which knowledge consists has to do with some kind of practical achievement or success, most of them accept the traditional view that knowledge is an epistemic good, and that its value is mostly of a theoretical, not of a practical nature.

Let us, following Duncan Pritchard (2007), distinguish the *primary value* problem for knowledge—the *Meno* problem of whether and why knowledge is more valuable than true belief—from the *secondary value* problem—the problem of whether knowledge is more valuable than any proper subset of its parts, on the assumption that the components of knowledge are true belief and justification, that is, whether knowledge is more valuable than justification, or true belief, for instance. The claim that some subparts of knowledge can be valuable because of their role in action surfaces in the writings of some virtue epistemologists. Thus Jonathan Kvanvig writes: ‘Belief is valuable because it is action guiding’ and he tells us that true beliefs are valuable because they lead to actions that ‘are successful in satisfying desires and in achieving purposes’ (Kvanvig 2003: 30). All

such claims, however, like parallel claims about the value of knowledge, imply that knowledge, or some subpart of it, is valuable because it is *instrumental* in bringing some good for action (true belief). In other words, they take knowledge, or some subpart of it, to be *extrinsically* related, as a means to an end which is valuable, and to get its value from the nature of this end. This is also the case when one takes knowledge, or some subpart of it, to be a means towards an intellectual or theoretical value.

A quite distinct conception surfaces in Timothy Williamson's (2000), John Hawthorne's (2004), and Jason Stanley's (2005) claim that knowledge, and not simply belief, is action guiding.¹ Williamson (2000: 62) insists that knowledge, and not simply true belief, plays a role in the explanation of action and, in particular, that action may not be as well explained in terms of a true belief as it would be in terms of knowledge. We ask why a burglar spends the whole night ransacking a house, risking discovery by staying so long, and we get a much better answer if we say that he *knew* that there was a diamond in the house rather than that he believed it. Very often, when we evaluate an action we do it in the light of the agent's knowing something rather than simply in terms of his having the corresponding true belief. For instance, if you drive your car while drunk and cause an accident, you will elicit almost automatically the reaction: 'You knew that could very well happen.' Hawthorne (2004: 30) uses a different argument, through lottery considerations, to the effect that in a lottery situation a subject does not know whether he will win, and should not base his actions on this lack of knowledge. Such arguments seem to show not only that knowledge is practically relevant, but also that it is valuable *because* it is practically relevant.² These intuitions about the relevance of knowledge in practical reasoning and in rational action are meant to suggest that knowledge 'matters' more than any of its subparts, in particular more than rational belief or justification. The intuitions, however, can also be extended to the subparts of knowledge. Belief, justified belief, and other epistemic states can be said to 'matter' because of the role that they play in action.

It is one thing to say that knowledge or other epistemic states are relevant to practical rationality. It is another to say that practical rationality determines the nature of epistemic states. Recently a number of writers, in particular Fantl and McGrath (2002) and Stanley (2005), have defended the view that there is such a 'pragmatic encroachment' (the phrase is Kvanvig's) not only on knowledge, but also on rational belief, justification, and evidence. The thesis that there is such a phenomenon of 'pragmatic encroachment' is in general based on the view that knowledge, and other epistemic states, are not 'purely epistemic'. Fantl and

¹ See also Hawthorne and Stanley (2008).

² Williamson 2000: 78 comments upon the Larissa road example in the *Meno*, but he is mostly concerned with whether knowledge explains action and does not address specifically the issue of the value of knowledge in this context.

McGrath (2002) call ‘purism’ the view that epistemic states are not pragmatically encroached, and Stanley (2005) calls the same view ‘intellectualism’. On the view that pragmatic encroachment is a genuine and significant phenomenon, whether a belief constitutes knowledge depends, in an essential way, upon factors which are not truth-conducive or purely epistemic, but ‘practical’. As we shall see below (§ 5), this view is distinct from, and stronger than, Williamson’s view that knowledge is relevant to action. It says that epistemic states by themselves have a pragmatic dimension and are in part ‘practical’. This looks like a form of pragmatism, although it is not clear in what sense it resembles the more traditional forms of this doctrine, according to which epistemic reasons and justifications—such as our reasons for thinking that a belief is true or well confirmed—are either determined or can be overridden by practical or prudential reasons or justifications. In its strongest sense, pragmatism is the view that epistemic reasons are reducible to practical reasons. I deal with this strong view only in my discussion of Horwich in the following section, and otherwise I deal only with claims of pragmatic encroachment. Such claims elicit two kinds of questions:

- (1) To what extent is there a pragmatic encroachment on epistemic notions?
- (2) To what extent does this pragmatic encroachment show anything about the value of knowledge or of other epistemic states?

It has to be noted at the outset that many of the writers who give a positive answer to the first question do not, explicitly at least, pretend to give an answer to the second question. They are not interested in the value problem about knowledge, but in the problem whether knowledge is more important or more fundamental than rational belief or justification. In other words, they raise a question about the *epistemic* priority of certain notions. The second question is problematic, for it is by no means evident that the practical relevance of knowledge explains its value. After all a traditional philosophical theme is that the value of knowledge does *not* reside in its relevance to action or in its consequences for actions. Moreover, not all actions are valuable: will knowledge which guides agents to perform actions lacking in value be itself valuable?³ But one can envisage an argument purporting to show that knowledge is valuable because of its particular relevance to rational action.

In this essay I try to show that one has to give a negative answer to both questions. I do not deny the phenomenon of pragmatic encroachment upon knowledge and upon other epistemic phenomena, but I disagree about the extent of this encroachment and about its significance. So I defend a kind of purism. There are actually various kinds of pragmatic encroachment, on belief, on justified belief, and on knowledge, which we have to examine in turn. I

³ This doubt is raised by DePaul and Grimm (2007) in a review of Kvanvig 2003.

shall conclude by noting that, even if we agreed on the existence of pragmatic encroachment upon knowledge, there is no legitimate step from it to a claim about the value of knowledge.

2. PRAGMATIC ENCROACHMENT ON BELIEF AND TRUTH

In order to understand pragmatic encroachment, it is necessary to distinguish several forms of it along two dimensions: first depending upon the kind of epistemic notion with which we are dealing, and second depending upon the strength of the encroachment, that is, upon the degree of dependence of an epistemic notion upon pragmatic factors.

Let us, for the time being, start from a very general notion of pragmatic encroachment. Let us say that there is pragmatic encroachment on an epistemic state (belief, rational belief, knowledge) if the practical consequences of being in that state are relevant to the epistemic evaluation of that state. Since belief is the simplest epistemic state, and since it is at least part of more complex states such as justified belief and knowledge, let us consider it first.

The idea that belief is, at least in part, determined by factors having to do with their relevance to rational action is not new. It is part and parcel of the dispositional theory of belief, according to which to believe that *P* is to be disposed to act as if *P*. It is incorporated in what Stalnaker (1984: 15) calls the 'pragmatic picture' of belief: to believe that *P* is to be disposed to act in ways that would tend to satisfy one's desires, in a world in which *P*. The picture is itself part and parcel of the functionalist conception of belief which is often formulated as a preference conception of belief:

(B) S believes that *P* if and only if S prefers as if *P*.

(B) can itself be translated in terms of degrees of belief and desirabilities within a Bayesian framework. On this view of belief, degrees of belief are a function of degrees of desires and of preferences, as they are revealed in an agent's action. So the epistemic notion of a degree of belief (of subjective probability) is itself a function of degrees of desires or utilities. When he first introduced this picture into contemporary philosophy, Frank Ramsey called it a form of 'pragmatism', which he took to be the view that 'the meaning of a sentence is to be defined by reference to the actions to which asserting it would lead, or more vaguely still by its possible causes and effects' and that 'the degree of a belief is a causal property of it which we can express vaguely as the extent to which we are prepared to act on it' (Ramsey 1927: 51; 1926: 65).

It seems clear, on the functionalist-Bayesian notion of belief, that degree of belief or subjective probability is a function of the role that it plays in practical

reasoning and in preferences towards certain actions. But does that imply that the very notion of degree of belief or subjective probability is determined by pragmatic factors such as the desirability or utility of action, and the preference towards action manifested by an agent? Can we say that an agent's degree of belief could be in some sense dependent upon its utilities and preferences? The answer is negative. In Bayesian decision theory, beliefs depend upon desires and actions, but this dependence does not imply that degrees of belief should be fixed in function of degrees of desire, or that both could be in some sense merged. On the contrary they are independent. I take it that this is what David Lewis (1988; 1996) has shown against the so-called 'belief as desire' thesis, according to which there could be 'besires', states which would be a mixture of beliefs and desires and which could motivate our actions. Value and credence cannot collide without decision theory being crippled. So there cannot be pragmatic encroachment on the notion of degree of evidence or degree of belief which is used by the standard Bayesian framework, if we keep this notion apart from utilities.

Now it is true that standard Bayesian decision theory holds that a rational agent uses a utility function defined over consequences of practical acts, and a credence function defined over these consequences conditional upon performances of the acts with respect to which his acts conform to expected utility. In this sense there is indeed a pragmatic justification of the concept of rationality which is used by standard Bayesian theory. Dutch book arguments *are* pragmatic proofs or rationality, since they show that violating the laws of probability in one's degrees of belief would lead one to be less well off. So it is the norms of prudential rationality, and not those of epistemic rationality, that justify the conformity of the degrees of belief to the laws of probability. But this does not show that the notion of degree of belief is *dependent* upon that of utility. In fact there are ways of justifying the notion of degree of belief which do not appeal to prudential rationality but to epistemic rationality only.⁴

So no functionalist conception of belief along the lines of (B) can yield a positive reply to question (1) in § 1, because the functionalist conception does not affect the *epistemic* character of belief. Can such a functionalist theory give us an answer to question (2)? In other terms, can the pragmatic involvement of

⁴ Joyce's (1998) 'non-pragmatic vindication of probabilism' is precisely meant to integrate a purely cognitive, and nonpragmatic, ideal or goal into the Bayesian framework, through the formulation of a 'Norm of Gradational Accuracy' instead of a Norm of Truth: An epistemically rational agent must evaluate partial beliefs on the basis of their gradational accuracy, and she must strive to hold a system of partial beliefs that, in her best judgement, is likely to have an overall level of gradational accuracy at least as high as that of any alternative system she might adopt (1998: 579). We have to ensure that in some sense there is *coincidence* between beliefs that elicit a high degree of credence and beliefs which elicit a high expected utility. In order to show that we can aim for a purely epistemic goal while at the same time ensuring that our degrees of belief are measured by probability functions, we must restrict the class of functions in a certain way. And Joyce shows that there are such functions, the Brier rules. Percival's (2002) 'cognitive decision theory' is meant to have similar effects.

belief in action give us an account of the *practical value* of beliefs? In order to have such an account, we would have to show that some epistemic values at least are practical values, in other words we would have to show that they owe their epistemic standing to their relevance to action. Truth is indeed a property of our beliefs that we value. Can we say that it is a value because, and only because, of the role it plays in our action? This would, of course, amount to some form of pragmatism, in the sense of a reduction of theoretical and epistemic reasons (and values) to practical reasons (and values).

Horwich (1999: 256–8; 2006) proposes just this: to explain the value of true belief simply in terms of the action theory of beliefs (B). He tells us that ‘True belief is valuable because it *pays*; it has practical benefits; you are more likely to get what you want if you base your deliberations and actions on true beliefs than if you base them on false ones’ (2006: 348). Horwich’s point is the following, familiar one. Beliefs have the potential to guide action aimed at satisfying our desires. Successful action is action that satisfies our desires. Beliefs are more likely to lead to successful action if they are true than if they are false. So there is reason to value believing what is true and only what is true.

Horwich’s key line of thought is as follows :

Focus, to begin with, on directly action-guiding beliefs of the form, ‘If I perform A, then X will occur’. It will clearly benefit me if I have many such beliefs and if they are all true. Because when I want a given thing, and believe that a certain action will get it, then, very often, I will perform that action. And in that case, *if my belief is true* this desire will be satisfied; whereas if it isn’t true no such result is ensured. So true beliefs of the ‘directly action-guiding’ form will indeed tend to benefit me. And the more such true beliefs I have the broader the spectrum of desires that will be easy for me to satisfy in this way. Moreover, these special beliefs are the results of inferences that tend to preserve truth; so it will benefit me for the *premises* of those inferences to be true. And there is no proposition that might not someday serve as such a premise. Therefore it will indeed be good for me—at least, that’s what it’s reasonable for me to expect—if I believe every true proposition and if every proposition I believe is true. (Horwich 2006: 350; see also Horwich 1999: 256–7)

So Horwich captures this conclusion in the following claim: it will be desirable, all things considered, to have action beliefs of this kind. Horwich generalizes this to all kinds of beliefs.

(D) Other things being equal, it is good to believe that p if and only if p.

(D) does not mention truth at all, but from it Horwich derives the value of truth, encapsulated in the principles:

(VT1) It is desirable to believe what is true and only what is true.

(VT2) If something is true, then it is undesirable to disbelieve it and desirable to believe it.

Horwich admits that (VT1) and (VT2) do not explain the norm of truth—why it is correct to have true beliefs—nor the specific epistemic norms, but he claims that this can very well explain why we acknowledge such norms, and why we value truth. This is a sort of explanation of norms and values of the same kind as that given by Gibbard (1991): although we cannot directly explain the nature of values, we can explain at least why we accept certain values, and, if there is no more to the values than our accepting them, they are explained.

Horwich's idea has some bite: he wants to say that truth is incorporated in the pragmatic conception of belief and that the value, or the norm of truth,⁵ is explained simply in terms of the value of successful action. For instance imagine that I have to choose between two boxes, one on the left which contains a bomb which will explode when I open it, and one in which there is a million dollars. Suppose I do not know which is which. If I believe that the bomb is in the box on the left, I aim to avoid it by taking the box on the right. Suppose that, in such a case, I want to believe the truth according to (VT1) and (VT2). If I believe that the bomb is in the box on the right, I shall satisfy (VT1) and (VT2) by believing that the million is on the left. But the aim expressed by (VT1) and (VT2) is such that, no matter what I believe, I have the aim of believing the truth. In other words there is no more to the value of a true belief than that it leads me to act successfully. Could we say more by saying that to aim at truth is to aim at being guided by the evidence that we have for the truth of our beliefs? But what is it to consider a given datum as evidence? It is to respond to the data in such a way that one's actions are successful. Such is the deflationist's line on the value of truth. But this deflationist line cannot explain the value of truth.

According to the deflationist the property that accounts for the value of truth is the property attaching to each true belief of being such that, other things being equal, it is good to believe it. If this is understood in terms of the practical value of true belief, then the question arises as to whether the account captures the normative character of truth. Clearly, it does not. The normative character of truth is captured by the truth norm for belief: a belief is correct if and only if it is true. The fact that it is a good from a practical point of view to have true beliefs does not explain why true beliefs and only true beliefs are correct. The deflationist might answer that its being good to believe that *p* when and only when *p* is precisely meant to capture the normative character of truth: surely we have *good reasons* to believe that *p* if so believing would enable us to act successfully. But this will not do since the reasons in question are not reasons to believe that *p* and thus believe that it is true that *p*. They are at best reasons for action aimed at bringing it about that one believes that *p*. The upshot is that we need the notion of truth to explain (VT2). There may be various ways in which it is desirable to believe a proposition or desirable to disbelieve it. Horwich

⁵ All these should of course be distinguished, but we do not need here to distinguish them. For an account of the difference, see Shah 2003 and Engel 2005.

focuses on desirability from a practical point of view. But, from a perspective on which we are interested in the correctness of belief, the desirability that matters is related to its being good to believe something *because it is correct to do so and correct because true*. And there is no way to account for the relevant kind of desirability without adverting to the norm for correctness and thus bring into play the concept of truth. So, we need the notion of truth to account for the relevant dimension of desirability—the dimension that is tied to correctness.⁶

The deflationist agrees that (VT1) and (VT2) can be understood as bearing on truth as a theoretical value, intrinsic but not instrumental, and that there is a long way from the acknowledgement of the pragmatic *relevance* of true beliefs in deliberation to the reduction of the value of truth to a practical value.

So I conclude that neither a functionalist theory of belief nor a decision-theoretic view can justify the claim that, as an epistemic notion, belief is pragmatically encroached in the sense examined in this section. Neither can the action theory of belief give us the appropriate explanation of the value of true belief.

3. PRAGMATIC ENCROACHMENT ON JUSTIFIED BELIEF

Pragmatic encroachment is primarily a condition on epistemic justification, hence not simply on belief, but on justified belief. Jeremy Fantl and Matthew McGrath argue that there is a ‘pragmatic necessary condition on epistemic justification’ (2002: 70):

(PC) S is justified in believing that *p* only if S is rational to prefer as if *p*.

(PC) is meant to be opposed to the thesis which they call ‘purism’ and which they define as the following view:

Purism

For any two possible subjects S and S', if S and S' are alike with respect to the strength of their epistemic position regarding a true proposition *p*, then S and S' are alike with respect to being in a position to know that *p*. (Fantl and McGrath 2007: 558)

Now strength of epistemic position is basically the evidence that one has in favour of a belief. *Evidentialism* is in general the view that a belief that *p* is epistemically justified for a subject S at a certain time *t* if and only if it fits the evidence that

⁶ Even though Horwich recognizes that true belief has non-instrumental value, he nevertheless holds that there is an important relationship between truth and instrumental value. As he puts it, ‘it is presumably *because* most truths are useful in practical inference . . . that our society, simplifying for the sake of effectiveness, inculcates a *general concern for truth for its own sake*’ (2006: 351, my emphasis).

S has for p at t (Conee and Feldman 2005: 83). Fantl and McGrath, however, address a somewhat different version of evidentialism which they characterize thus:

Evidentialism

For any two subjects S and S' , necessarily, if S and S' have the same evidence for/against p , then S is justified in believing that p iff S' is, too. (Fantl and McGrath 2002: 68)

Fantl and McGrath's argument against Evidentialism consists first in registering our intuitions from examples similar to those familiar in the literature on contextualism in epistemology, such as the airport example used by Cohen (1999) or DeRose's (1996) bank cases.

Case 1. You want to know whether the approaching train from Boston to Providence stops in the intermediate stop in Foxboro or whether it is an express. It does not matter very much to you whether it is the express or not, although you mildly prefer it was. You ask someone waiting for the train, and he says that the train does stop in Foxboro. You believe him.

Case 2. You have the same evidence as in Case 1. But here it is very important for you to be in Foxboro, since you are going to have there an interview on which your career depends. You overhear someone say that the train stops in Foxboro. But you think 'that information might be wrong, I have to check further.'

According to Fantl and McGrath our intuition is that in Case 1 your evidence is good enough for you to know that the train stops at Foxboro, whereas in Case 2 your evidence is not good enough for you to know that the train stops at Foxboro. The difference is due to the fact that not much is at stake in Case 1, whereas a lot is at stake in Case 2, *although in both cases you have the same evidence*.

In Case 1, you are rational to prefer (A), boarding the train without enquiring further, to (B), boarding the train after enquiring first, given that the train stops at Foxboro (p), because B would involve a cost (spending time to enquire and risking to miss the train). In Case 2, the preference is the reverse.

Fantl and McGrath are guided by the plausibility of the following inference (2002).

- (1) S knows that p .
 - (2) S is rational to prefer A to B given p ($A \ \& \ p$ to $B \ \& \ p$)
- Therefore,
- (3) S is rational to prefer A to B in fact.

This seems intuitive enough. It leads them to formulate a necessary condition on knowledge.

S knows that p only if, for any state of affairs A and B , if S is rational to prefer A to B , given p , then S is rational to prefer A to B , in fact. (2002: 76)

What you know determines what you should rationally prefer. Here the reasoning relies on the plausible principle, also advocated by Hawthorne (2004: 29), Williamson (2000), and Stanley (2005), that what one knows is directly relevant to, and can be used as a premise in, practical reasoning:

- (PR) If S knows that p , S is justified in using p as a premise in his practical reasoning.

Fantl and McGrath defend, however, a pragmatic condition on justification. In this they are guided by the plausibility of the following inference (2002: 76).

- (1*) S is justified in believing that p .
 (2*) S is rational to prefer A & p to B & p .
 Therefore,
 (3*) S is rational to prefer A to B in fact.

According to Fantl and McGrath, our intuitions in the train cases justify the following reasoning:

Suppose a subject S, is justified in believing that p , but does not know that p . Suppose further that S is rational to prefer A to B, given p . Compare S to a second subject S', who has the same evidence and fundamental preferences as S but who does not know that p . S' is rational to prefer A to B. What one is rational to prefer is determined by one's evidence and fundamental preferences. Since S and S' have the same evidence and fundamental preferences, they will be rational to prefer the same states of affairs. Thus, S, too, is rational to prefer A to B. Whatever it is rational to prefer for a knower is also rational for an otherwise identical subject who is merely justified in believing to prefer. (2002: 76)

The upshot is that if the inference (1)–(3) concerning knowledge is valid so is the inference (1*)–(3*). This leads to the principle

- (JBP*) S is justified in believing that p only if, for any states of affairs A and B, if S is rational to prefer A & p to B & p , then S is rational to prefer A to B in fact.

In the train cases, the principle is used in the following way. In Case 1, since what it is rational for you to prefer in fact coincides with what it is rational for you to prefer conditionally on the truth of p , you satisfy the pragmatic condition on justification (JBP*). So, as they say, 'you may have enough evidence for justification' (2002: 80). That is, the testimony you have for believing that the train stops in Foxboro may well justify you in believing that the train stops in Foxboro. In Case 2, you have the same evidence as in Case 1 and the same fundamental preferences. But in Case 2, the stakes are high, not low as in Case 1. Your career hangs on whether the train stops in Foxboro. In this Case, as in 1, you are rational to prefer A (boarding the train without enquiring further) and p

(the train stops in Foxboro) to B (boarding the train after further enquiry) and p (the train stops in Foxboro). But, given the stakes that are involved, you are not rational to prefer A to B in fact. You need to enquire further to make sure that the train will stop in Foxboro. So in Case 2 what it is rational for you to prefer in fact diverges from what it is rational for you to prefer conditionally on the truth of p . Hence you do not satisfy (JBP*). Hence you are not justified by your evidence in believing that p . If you are justified by your evidence in believing that p in Case 1 (in which (JBP*) is satisfied), then evidentialism is false. The difference in stakes engenders a difference in justification.

What do Fantl's and McGrath's analyses actually show? In the first place, we should note that it is surprising that examples such as the train cases, which have been initially used by contextualists to show that there are variations in the sense of our ascriptions of knowledge (and here of justification as well) are used to show that there is an impact of pragmatic factors on knowledge and justification. I shall come back to this. I must say that I do not share Fantl's and McGrath's intuitions (nor, for that matter, Stanley's intuitions about similar examples; see below). My own intuitive response to the train cases is that the individual in Case 1 and in Case 2 has exactly the same justification because they have the same evidence (the testimony of another traveller in the station to the effect that the train stops at Foxboro). The fact that more is at stake in the second case does not seem to me to show that the individual is less epistemically justified in the second case than in the first. Now, one may reply to this that although the two individuals in each case are equally justified, they are not justified *in the sense required for their knowledge*.⁷ Their point is that the weight imposed by the stakes elevates the amount of epistemic justification which is required for their knowledge. Justification in this sense does not simply rest upon the evidence that one has at a given time, but also upon the evidence that one *might* have at a later time, if the stakes were higher and one had to reconsider one's earlier evidence. Let us call this kind of justification *justification** —or *enlarged* justification—and let us call *counterfactual evidence* or *evidence** the evidence that one *would* need to have if the stakes were higher. The reasoning above then purports to show that although justification in the evidentialist sense of (E) can be exactly the same, *justification** is not. Now one can certainly grant this, and if one understands the intuitions voiced in the train cases as showing a difference in *justification**, one can certainly agree that there is pragmatic encroachment in *this* sense. But the problem is precisely whether this is a pragmatic encroachment on (evidential) justification or upon *justification**. And it seems to simply beg the question to assume that the two notions of justification (and of evidence) are equivalent. It seems to me that what happens in the train cases is that the cost of information is more salient in Case 2 than in Case 1, and that it is more important to be justified about p in Case 2 than in Case 1. But the fact that it is more important, practically

⁷ I thank Jeremy Fantl for having pointed this out to me.

or pragmatically or prudentially, to be justified does not mean that we are more or less justified depending upon the stakes. The fact that we can say that one is not justified* to believe that p when the stakes on believing that p affect one's preferences importantly and the consequences of one's actions does not imply that one is not justified *simpliciter* in the straightforward evidentialist sense.⁸

Intuitions are not foolproof, either in philosophy or elsewhere. But even if our intuitions were correct, and if we recognized that the individual in Case 1 is justified to believe that the train stops in Foxboro but not justified to do so in the second case, it would remain to be shown that the difference in justification is due to the 'pragmatic' factor in question. It is not. Up to now I have not considered the formulation of the problem in terms of degrees of belief. If we suppose that degrees of belief reflect the degree of confidence that one can have towards a proposition, given a certain amount of evidence, by hypothesis in the train Cases 1 and 2 there is no reason to suppose that the individual has a different degree of belief in situation 2 than he or she has in situation 1. By definition she has the same evidence, and she should therefore believe exactly to the same degree that the train stops at Foxboro. So what explains that one can have the intuition that our subject is 'more justified' in Case 1 than in Case 2? What explains our intuition that there is a difference is of course that the evidence that is needed for believing that the trains stop at Foxboro is greater in Case 2 than in Case 1, because of what is practically at stake.⁹

⁸ In a review of Stanley 2005 (Pritchard 2006), Duncan Pritchard expresses similar intuitions. Stanley (2005: 12) envisages that one reaction we can have to his bank cases is to 'challenge the claim that these are the intuitions we have in these cases', but he does not attempt to discuss these intuitions or to counterbalance them by others.

⁹ As Julien Dutant has pointed out to me, one can translate the situation in decision theoretic terms. The agent has two options in Case 2: Act (A) on the basis of her best judgement or enquire further (E) to get an additional information i . Let $P(R)$ be the probability of being right, $U(R)$ the utility of being right and $C(T)$ the cost of being wrong. If $U(R)$ is 5 then the utility of being wrong, $U(T)$, is -5 . One supposes that if the subject is not right, he is wrong. Let $C(I)$ the cost of acquiring the information i , and $P(R|i)$ the probability of having got the right information after having acquired i . The expected utility of A is:

$$G(A) = P(R).U(R) - (1 - P(R)).C(T)$$

and the expected utility of enquiring further is:

$$G(E) = P(R|i).U(R) - (1 - P(R|i)).C(T) - C(i).$$

The agent maximizes his expected utility in enquiring iff i :

$$G(E) > G(A)$$

$$P(R|i).U(R) - (1 - P(R|i)).C(T) - C(i) > P(R).U(R) - (1 - P(R)).C(T)$$

$$[P(R|i) - P(R)].U(R) - [(1 - P(R|i)) - (1 - P(R))].C(T) > C(i)$$

$$[P(R|i) - P(R)].[U(R) - C(T)] > C(i)$$

Let the increase of the probability of being right on the basis of i : $d = [P(R|i) - P(R)]$. Let D the difference in utility between being right and being wrong: $D = [U(R) - C(T)]$. And let the cost of information be I . The subject maximizes his expected utility by enquiring iff $dD > I$. This

The difference between justification (in the strict evidentialist sense) and justification* or enlarged justification in what we may call the pragmatic sense is that in this second we are much more reluctant to attribute *knowledge*. What is happening in fact in the train cases is that in the second case we consider that although the subject has justification he does not have knowledge. In Stanley's terms (2005: 88), 'the greater the practical investment one has in a belief, the stronger one's evidence must be in order to *know* it [my italics]'. Stanley, as we shall see in the next section, talks explicitly of knowledge in such cases, and not of justification. The problem with the reasoning (1*)–(3*) above is that when the stakes are high, we lose grip on how much justification is needed for having knowledge. We are thus in a familiar situation: on the one hand, if we raise too high the requirements for justification in order for it to amount to knowledge, we run the risk of scepticism, and, on the other hand, if we accept fallibilism, we do not ensure knowledge. So the argument (1)–(3) above is much more credible than the reasoning (1*)–(3*) which ascribes justification, precisely because the state ascribed to the subject is the state of knowing.¹⁰ If what is ascribed is justification* (i.e. enough justification to yield knowledge), the reasoning (1*)–(3*) is much more appealing. But we should not equivocate between the two notions of justification and the two notions of evidence. In any case, from the fact that one's evidence* must be stronger, or that one needs to look for more counterfactual evidence than one has in Case 2 than in Case 1, it does not follow that one *has* more evidence in the first case than in the second.

What seems to happen, in the cases about justification* that Fantl and McGrath discuss is this. The practical significance of the truth of a proposition does *not* affect the (evidential) justification of the subject, although it does affect our *attribution* of knowledge, and our judgement about the amount of justification needed for knowledge. There are pragmatic limitations on the collection of evidence, but it does not imply that evidentialism is false. This does not amount to a pragmatic conception of justification. David Owens makes the point in describing the pragmatic limitations of collecting evidence:

[One can] agree with the pragmatist that practical considerations determine whether we should form a view about the truth of *p* but given that we want a belief on the matter, evidence alone determines whether we ought to believe *p* or believe not *p*. . . . Hence non-evidential considerations are needed to rationalise and motivate belief formation. But [this does not imply the pragmatist's further claim] that this makes belief subject to practical norms: the non-evidential considerations which help to fix the rationality of belief are not the practical considerations which govern the rationality of action. (Owens 2000: 31)

validates our intuition that if the cost of error is important, and the information is not too costly, it is important to try to augment one's probability of being right by enquiring further. For further elaboration of the situation in decision theoretic terms, see Weatherston 2005.

¹⁰ Thanks to Jeremy Fantl for having attracted my attention to this point. I would thus accept closure when (1)–(3) mentions knowledge, but not when it mentions justification.

Practical considerations determine whether we should form a view about the truth of p but, given that we want a belief on the matter, evidence alone determines whether we ought to believe p or believe not p . Pragmatism about justification has nothing to do with whether *evidence* counts as justification of a belief. It has to do with whether we *form* a belief or not depending on the cost of doing so. In the train case, the subject has exactly the same evidence, but in the second case unlike in the first he cares very much about how he should form his belief. This has nothing to do with the falsity of evidentialism. It has to do with whether he should form an opinion or not, or make a corresponding assertion.¹¹

There is a second reason why Fantl's and McGrath's argument is not convincing. By definition the individual's evidence in Case 1 and in Case 2 for the proposition p is the same. But his preferences are different. How can this be, given that the preferences are themselves determined by the amount of evidence that he has? That seems an incoherent description of the situation. But perhaps there is some equivocation on 'evidence' here. Perhaps the thought is that the individual in Cases 1 and 2 has the same evidence *for proposition p* in both cases, but the global evidence is distinct in Case 2, because it includes the fact that he does not care very much about whether the train is an express or not, whereas it matters in the second case. But we could express the same idea as above by saying that what Fantl and McGrath have described is not a case where one's evidence is affected *as such* but a case where one's attitude towards the evidence is affected.¹² But doesn't the evidence change if our attitude towards the evidence changes, and is not evidentialism threatened just by this fact? Indeed, if we feel the need to gather more evidence than we actually have, then it is possible that our evidence will change. But does that show that the evidence that one has at t is affected by what is practically at stake? No, the fact that one feels that one would need to have more evidence if there are high stakes does not change the relationship which exists between a belief and the evidence which justifies it. As I have proposed, we can use the concept of evidence*, understood in the same manner as the concept of justification*, in order to designate the evidence which the subject would feel to be required in a given situation. And there is no doubt that the corresponding doctrine of evidentialism* would be false. But it does not imply that evidentialism in the sense of (E) is false.¹³

¹¹ In the light of what Owens actually says, it is surprising that he should be recruited among the pragmatists by Fantl and McGrath. Schaffer (2006) says that it is rather the relevance of the possibility of error which is driving our intuitions in the cases. This is plausible. But that does not in any way disprove evidentialism.

¹² Weatherson (2005) has a similar diagnosis, but his specific point is that the difference between the two cases is that 'interests matter not because they affect the degree of confidence that an agent can reasonably have in a proposition's truth. (That is, not because they matter to epistemology.) Rather, interests matter because they affect whether those reasonable degrees of confidence amount to belief. (That is, because they matter to philosophy of mind.)'

¹³ For a similar reaction to Fantl and McGrath's case, see Conee and Feldman 2005: 103–4.

All of this is not meant to deny the phenomenon which Fantl and McGrath are pointing to, only their interpretation of it. Let us call *pragmatic relevance* the fact that the conditions under which we are led to form our beliefs are often associated with practical interests and with factors which are extra-epistemic, such as economy or gain of time and resources. Such factors are present in the formation of belief, in knowledge, and in our judgements about justification. But they do not imply that knowledge and justification are pragmatic notions. Pragmatic relevance does not entail pragmatic encroachment of the epistemic by the non-epistemic.

4. PRAGMATIC ENCROACHMENT ON KNOWLEDGE

Having found reasons for doubting the existence of pragmatic encroachment on justification, we can now turn to its application to knowledge. As we saw Fantl's and Mc Grath's in reasoning (1)–(3) about knowledge, it has much more plausibility than the reasoning (1*)–(3*) about justification. Stanley (2005) applies similar considerations directly to knowledge, on the basis of examples such as DeRose's bank cases. Stanley is concerned with a defence of the view which he calls 'interest-relative invariantism' (IRI) directed against—in particular—contextualism. The main difference between subject-sensitive invariantists like Hawthorne or interest-relative invariantists like Stanley, on the one hand, and contextualists, on the other hand, is that the latter claim that attributions of knowledge are relative to the ascriber, whereas the former claim that the meaning of 'know' is constant across contexts (and thereby count as invariantists), even though our attributions of knowledge can be true in some contexts, and false in others. Stanley gives a series of examples which are variations on the train cases. I have already noted above that the oddity is that such examples had been used by contextualists to defend a form of purism or intellectualism. Contextualists use the same examples but where interest-relative invariantists say that epistemic notions are pragmatically encroached but their meaning is invariant, contextualists stick to purism and evidentialism, but claim that the meaning of words such as 'knows' varies across contexts. It is not my purpose here, in defending purism, to defend a form of contextualism. On the contrary, though I accept invariantism about knowledge, it is not my concern here to argue for this view. I am not going to detail the examples and the discussion that Stanley directs at contextualism, because my concern here is not which theory of our knowledge claims is correct, but whether this view threatens purism or evidentialism (or intellectualism).

When he describes IRI at the beginning of his book, Stanley makes quite strong claims to the effect that 'the factors that make true belief into knowledge include elements from practical rationality' and 'what makes true belief into knowledge is not entirely an epistemic matter' (2005: 2), if 'epistemic' is understood to concern

only truth-conduciveness.¹⁴ Describing the case of his characters, Hannah and Sarah, who wonder whether the bank is open on Saturdays, and in differing situations have high or low stakes (2005: 3–5), he voices the same intuitions as those which we have elicited from the train cases. But his point is that when the stakes are high the subject *does not know* that *p*, whereas when they are low he or she knows that *p*. Unlike Fantl and McGrath, he directly addresses the issue of a pragmatic encroachment on knowledge. For him IRI is a view about knowledge, and not necessarily a view about justified belief (although it can presumably be extended to it). Why is this so? Because, as we saw above, he formulates his view not as a pragmatist conception of justification in the sense of (JBP) or (JBP*) above, but as the idea that the greater one's practical interests are, the stronger one's evidence must be. But, as we saw, this kind of pragmatic encroachment is innocuous and does not threaten the purist's notion of evidence. It does not say that pragmatic factors *enter* or determine the amount of our justification, but just that pragmatic factors impinge upon our need of more evidence.

All of Stanley's examples have to do not with whether evidence supports a proposition and is able to determine our knowledge of it but with the issue of whether believing a given proposition is 'a serious practical question' (2005: 91 ff.), or whether it is legitimate to take it into account or to ignore it. For instance, a number of propositions which have no relevance for my present plans, such as whether Christine Todd Whitman cut her toenails on 1 September 2003 or that a large asteroid might hit the earth in 10 billion years have no relevance to my present cognitive interests. But if I were to learn that the asteroid will hit the earth next week, it would affect my plans. Stanley argues that 'the fact that the negation of a proposition is an epistemic possibility for an agent in a situation prevents the subject from knowing that proposition in this situation' (2005: 96). But he immediately adds:

However the fact that a proposition is a serious practical question for an agent at a time does not automatically undermine the agent's knowledge of that proposition. It would only undermine the subject's knowledge of a proposition if, given her evidence, the probability of the negation of that proposition is not sufficiently low. (ibid.)

But that in no way prevents us from accepting evidentialism. If evidentialism is defined as above—as the thesis that evidential twins have the same justification for their beliefs—the view is not touched at all by the pragmatic factors involved in belief formation and the seriousness of epistemic possibilities. It certainly does not show that 'all normal epistemic notions are interest relative' and that 'evidence is interest relative' (Stanley 2005: 124).

¹⁴ These claims would not be so ambitious if, in the spirit of some of his declarations, Stanley considered IRI as mostly a theory of knowledge *attributions*, and not as a theory of knowledge on its own. He seems to say this: 'My purpose is to establish that knowledge is conceptually connected to practical interests. The point is compatible with many different approaches to the nature of knowledge' (2005: 89), but he actually thinks of his views as bearing on the nature of knowledge itself.

Like Fantl and McGrath, Stanley makes much of the following principle:

(PR) If one knows that p , then p is apt to figure in practical reasoning.

Hawthorne (2004: 30) gives examples such as the following which illustrate (PR). Imagine that someone is offered a penny for a lottery ticket and reasons thus:

I will lose the lottery

If I keep my ticket, I'll get nothing

If I sell my ticket I'll get 1 dollar

So I ought to sell my ticket.

Everyone agrees that this reasoning is unacceptable because the agent is not entitled to assert the first premise, since he does not *know* that it is true before the winning ticket is drawn. By contrast if the person has just heard the winning number, seen it was not his, and sold the ticket to someone else unaware of the announcement, we would not criticize the reasoning. According to Hawthorne, this shows that 'if the question whether p is practically relevant, it is acceptable to use the premise that p in one's deliberations if one knows it and (at least in very many cases) unacceptable if one does not know it' (ibid.).

(PR) is in my view correct, although it has been contested by those who claim that practical reasoning does not need to be premised on knowledge, only on rational belief (Kaplan 1985). Although there is a lot of discussion of this principle, I shall not put it into question, because it does not seem to me to imply any pragmatic encroachment on the notion of knowledge. The fact that knowledge of a premise is needed in practical reasoning no more shows that the notion of knowledge involves a pragmatic factor than the fact that a proposition is a serious possibility for an agent shows that evidence is an 'interest-relative notion'. Fantl and McGrath (2007) derive from their reasoning (1)–(3) a pragmatic notion of knowledge, according to which

(PCK) S knows that p only if S is rational to act as if p .

But here we should note that there is an important difference between (PCK) and the principle (PR) that knowledge enters as premise in a practical reasoning. What (PCK) says is that when it is not rational to act as if p —for instance because the stakes are too high to run the risk of acting as if p were true—then the subject does not know that p . This *is* pragmatic encroachment on knowledge, since it says that a necessary condition on knowledge is that one's actions and preferences should be practically rational. But the correct lesson of (PCK) goes in the opposite direction: (PCK), properly understood, actually says that, when one does not know that p , it is irrational to act as if p were true.¹⁵ Again, let

¹⁵ Hawthorne and Stanley (2008) note this in relation to Fantl's and McGrath's principle, and they say that it takes the order of explanation wrongly. I approve, and claim that this shows that

us put aside the issue whether (PR) is correct. It is quite a different thing to say that knowledge is a necessary condition of rational action and to say that acting rationally is a condition on knowledge. The second is indeed a form of pragmatism, but the first bears all the marks of what is usually considered to be an *intellectualist* view: in order to act well you have to be a knower. Indeed, to take up the formulation used at the end of the preceding section, (PR) says that knowledge is *relevant*, or ‘matters’ for practical reasoning and action. But that is not to say that what makes something knowledge somehow *depends* upon the rationality of our actions. So I conclude that (PR), true or not, cannot serve as a basis for pragmatic encroachment understood as the view that practical constraints weigh on epistemic justification and knowledge.

5. CAN PRAGMATIC ENCROACHMENT CONFER VALUE TO KNOWLEDGE?

If the foregoing is correct, none of the arguments for pragmatic encroachment of epistemic notions such as truth, evidence, justification, and knowledge show that purism, or intellectualism, defined as the view ‘that knowledge is not a matter of practical facts’ (Stanley 2005: 6) work. Contrary to what they announce, they do not show that evidentialism is false.¹⁶ This answers question (1) from Section 1 in the negative. But what about question (2): could pragmatic encroachment explain, or at least could it give us some hint as to why knowledge is valuable, and more valuable than true belief?

Of course if the phenomenon of pragmatic encroachment reduces, as I have claimed, to that of pragmatic relevance, question (2) has to be answered in the negative too. But this does not prevent us from asking: if there *were* pragmatic encroachment on knowledge, would that show anything about the value of knowledge? In fact none of the writers who defend the view that there is a pragmatic encroachment on epistemic concepts deals explicitly with the issue of the value of knowledge. They are mostly concerned with the question whether truth, justification, and knowledge are to be ascribed on the basis of epistemic considerations alone. They are interested in whether knowledge ‘matters’ with respect to other epistemic notions, in particular justification.¹⁷ So they are mostly interested in whether one can analyse knowledge in terms of belief, truth, or

there is not pragmatic encroachment on knowledge with PCK, since it is, literally, knowledge which encroaches on rational action, and not the other way round.

¹⁶ I have not, however, tried here to defend evidentialism as such. There are other difficulties with the view which I have not discussed.

¹⁷ This is perhaps to be qualified by the last sentence of Hawthorne and Stanley’s (2008) article ‘Knowledge and Action’, where they say, after addressing many of the issues about pragmatic encroachment that ‘[t]he value of knowledge is due in part to its role as a norm for action’. They do not explain, however, in what part, nor why it is so.

justification, and whether these notions are pure. But in order to deal with the value problem, in its primary or in its secondary form, we need also to ask whether a proper subpart of knowledge confers value to it, or whether it has more value than any of its subparts. In other words, it does not follow from an analysis of knowledge and justification in pragmatic terms that knowledge is valuable, either intrinsically or relatively.

But even if knowledge matters in this sense, and I agree with them that it does, that still does not give us any answer to the secondary value problem: is knowledge more valuable than any of its subparts? We would have the beginning of such an answer if it could be shown, for instance in the reliabilist way, that knowledge is apt to produce more true beliefs than sheer luck or absence of method, or if the way in which knowledge matters could be associated to some specific dispositions of knowers, as virtue epistemology proposes. But the very fact that our judgements about knowledge are relevant to our evaluation of actions, or that they are relevant for practical reasoning, to repeat, shows nothing. In particular, the kind of argument defended by Williamson about the role of knowledge in the explanation of action does not in any way show that there is pragmatic encroachment on knowledge, for it is quite open to someone to hold that knowledge is relevant to the explanation of action while denying that whether one knows that *p* turns on practical matters. The latter is indeed the position which I myself accept.

I have denied that pragmatic factors affect the *epistemic evaluation* of beliefs or the attribution of knowledge, because I take, with the orthodoxy, the question whether one is justified or knows as a question to be decided on epistemic grounds alone. But I have not denied that, as the pragmatic encroachers say, interest-relative or practical 'factors' are *relevant* to knowledge. But how are they relevant? They are relevant *within the context of enquiry*. What is 'at stake' affects whether one should, in a given circumstance, form, or maintain, a belief, or what amount of evidence one decides to take. In this sense, there is real pragmatic 'sensitivity'. Now, in so far as these pragmatic factors affect enquiry, the decisions that an agent takes to suspend belief or go ahead, it can affect his actions. In so far as these are part of his reliability as an agent, or his character, this sort of fact should be important for virtue epistemologists, and hence for the assessment of the value of knowledge as such. But, if I am correct, this fact has no bearing on the very epistemic evaluation of knowledge and belief, which depends as much on evidence as it ever did.¹⁸

¹⁸ An earlier version of this article was read at the Epistemic Value conference in Stirling. I thank for their careful comments my commentators Joe Salerno, Berit Brogaard, Duncan Pritchard, and Scott Sturgeon. I thank also Jeremy Fantl, Matthew McGrath, and Jason Stanley for their remarks on the initial version and their help in trying to correct my misunderstandings about their views. I hope that there are not too many left. Another version was read at the SOPHA conference in Aix en Provence in September 2006, and in a workshop in Lausanne in 2007. Thanks to the participants on both occasions, and especially Julien Dutant for his accurate comments and discussion. I owe much to Adrian Haddock for his help with the last version, as well as two anonymous referees.

REFERENCES

- Brady, M. S. (2006). 'Appropriate Attitudes and the Value Problem', *American Philosophical Quarterly*, 43: 91–9.
- Cohen, S. (1999). 'Contextualism, Skepticism, and the Structure of Reasons', *Philosophical Perspectives* 13, *Epistemology*: 57–89.
- Conee, E., and Feldman, R. (2005). *Evidentialism*. Oxford: Oxford University Press.
- DePaul, M., and Grimm, S. (2007). Review of Kvanvig 2003, *Philosophy and Phenomenological Research*, 74: 498–514.
- DeRose, K. (1992) 'Contextualism and Knowledge Attributions', *Philosophy and Phenomenological Research*, 52: 913–29.
- (2007). Review of Stanley 2005, *Mind*, 116: 486–9.
- Engel, P. (2005). 'Truth and the Aim of Belief', in D. Gillies (ed.), *Models in Science*. London: King's College Publications.
- Fantl, J., and McGrath, M. (2002). 'Evidence, Pragmatics and Justification', *Philosophical Review*, 111: 67–94.
- (2007). 'On Pragmatic Encroachment in Epistemology', *Philosophy and Phenomenological Research*, 75: 558–89.
- Gibbard, A. (1991). *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press.
- (2007). 'Rational Credence and the Value of Truth', in T. Gendler and J. Hawthorne (eds.), *Oxford Studies in Epistemology*, ii. Oxford: Oxford University Press.
- Greco, J. (2002). 'Knowledge as Credit for True Belief', in M. DePaul and L. Zagzebski (eds.), *Intellectual Virtue Perspectives from Ethics and Epistemology*. Oxford: Oxford University Press.
- Hawthorne, J. (2004). *Knowledge and Lotteries*. Oxford: Oxford University Press.
- and Stanley, J. (2008). 'Knowledge and Action', *Journal of Philosophy*, 105: 571–90.
- Horwich. P. (1998). *Truth*. Oxford: Oxford University Press.
- (1999). 'The Minimalist Conception of Truth', in S. Blackburn and K. Simmons, (eds.), *Truth*. Oxford: Oxford University Press (reprinted with slight revision from Horwich 1998).
- (2006). 'The Value of Truth', *Nous*, 40(2): 347–60.
- Joyce, J. (1998). 'A Nonpragmatic Vindication of Probabilism', *Philosophy of Science*, 65: 575–603.
- Kaplan, M. (1985). 'It's Not What You Know that Counts', *Journal of Philosophy*, 82: 350–63.
- Kvanvig, J. (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge: Cambridge University Press.
- Lewis, D. (1988). 'Desire as Belief', *Mind*, 97: 323–32; repr. in Lewis 2000.
- (1996). 'Desire as Belief II', *Mind*, 105(418): 303–13; repr. in Lewis 2000 .
- (2000). *Philosophical Papers, 3, Papers in Ethics and Social Philosophy*. Cambridge, MA: Cambridge University Press.
- Owens, D. (2000). *Reason without Freedom*. London: Routledge.
- Percival, P. (2002). 'Epistemic Consequentialism', *Proceedings of the Aristotelian Society*, 76, suppl. vol., 121–51.
- Pritchard, D. (2006). Review of Stanley 2005, *Notre Dame Reviews in Philosophy*, 26.06.18. <<http://ndpr.nd.edu/review.cfm?id=6885>>.

- (2007). 'Recent Work on Epistemic Value', *American Philosophical Quarterly*, 44: 85–110.
- Ramsey, F. (1926). 'Facts and Propositions', in Ramsey 1991.
- (1927). 'Belief and Probability', in Ramsey 1991.
- (1991). *Philosophical Papers*, ed. D. H. Mellor. Cambridge: Cambridge University Press.
- Shaffer, J. (2006). 'The Irrelevance of the Subject: Against Subject Sensitive Invariantism', *Philosophical Studies*, 127: 87–107.
- Shah, N. (2003). 'How Truth Governs Belief', *Philosophical Review*, 112: 447–82.
- Sosa, E. (2007). *A Virtue Epistemology: Apt Belief and Reflective Knowledge*. Oxford: Oxford University Press.
- Stalnaker, R. (1984). *Inquiry*. Cambridge, MA: MIT Press.
- Stanley, J. (2005). *Knowledge and Practical Interests*. Oxford: Oxford University Press.
- Weatherson, B. (2005). 'Can We Do without Pragmatic Encroachment?', *Philosophical Perspectives*, 19: 417–43.
- Williamson, T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.
- Zagzebski, L. (1996). *Virtues of the Mind*. Cambridge, MA: Cambridge University Press.

9

Luck, Knowledge, and Control

Wayne D. Riggs

1. INTRODUCTION

As a proponent and practitioner of value-driven epistemology,¹ I am very gratified that this collection of papers on epistemic value has been put together. I believe that the recent emphasis on epistemic value within epistemology has already borne fruit, with the promise of much more to come. One reason for this promise is that a value-driven approach to epistemology invites one to ask kinds of questions that, while certainly not prohibited by more traditional epistemological method, do not naturally arise. Twentieth-century analytic epistemology became rather myopically focused on getting the conditions under which one has knowledge just right, leaving aside for the most part questions about the role and value of knowledge in our lives more broadly. After all, while humans are knowers, we are many other things besides. Getting clear on precisely why we care about knowledge, whatever exactly it is, helps us understand how our cognitive lives are entwined with our moral lives, our prudential lives, etc.

To understand how concerns about luck intersect this broader value-oriented approach, one must approach the theory of knowledge from an unusual direction. It has been noted by countless epistemologists that luck, in some guises, undermines knowledge. Lucky guesses have been the paradigm examples for decades for why we need another condition on knowledge besides truth and belief. A lucky guess simply cannot count as knowledge, precisely because it is lucky. And, of course, Gettier cases are often described as cases in which one gets to the truth only as a matter of luck. The question I suggest we ask is, 'why are we so opposed to being right by luck?' If avoiding luck is really so central to our understanding of knowledge, then an investigation of luck, at least of the kind of luck that undermines knowledge, is an investigation of the conditions of knowledge as well. Consequently, if we can figure out why we find it so abhorrent

¹ This term was introduced, I believe, by Jonathan Kvanvig on his blog, *Certain Doubts*.

to come by true beliefs in a fortuitous fashion, we will have come a long way toward understanding why we value knowledge.

One thing this requires, of course, is a detailed account of the kind of luck that undermines knowledge. My goal in this paper is to articulate just such an account in terms of control. Very roughly, E is lucky for S to the extent that S had little or no control over bringing E about. Other accounts of such luck have been proposed, of course, so I will begin by arguing against these alternative views. An interesting and heretofore unnoted wrinkle in the control account will require a substantial revision of the view, and highlight some interesting features of luck and control. While it is beyond the scope of this paper to go on to argue that avoiding this kind of luck is valuable in the way necessary to account for the special value of knowledge, I will conclude with some suggestions of the direction I think such a defense should take.

As with many important philosophical notions we are interested in investigating, the ordinary term we use to express the notion 'luck' is multiply ambiguous, and in practice we use a variety of pragmatic conventions to disambiguate them and make our meaning clear. Indeed, we are quite good at this, so that while ambiguous terms do occasionally cause problems in everyday conversation, this is the exception rather than the rule. But all this complicates a philosophical analysis of such a term, since the various senses of it have to be distinguished in the absence of clear or crisp semantic differences in our various uses of the term. Such is our situation with regard to 'luck.' I want to delineate several common senses of this term here at the beginning of the paper, and try to get tolerably clear on the fairly narrow sense of 'luck' that I wish to investigate.

I have undertaken a similar task in a previous paper with regard to the related term 'chance.' This represented an early attempt on my part to carve out two rather narrow senses of 'luck' or 'chance' from the penumbra of everyday meanings of these terms.

Both notions fall into the family of notions I shall call notions of 'chance.' Terms that designate members of this family include 'lucky,' 'fortuitous,' 'accidental,' 'undesigned,' 'incidental,' 'unexpected,' 'random,' and so on.²

As I pointed out in that paper, these notions clearly bear on one another, and yet equally clearly are not synonymous. One could have a design to bring something about, yet expect it to fail, and thus one's success would be designed yet unexpected. An accidental occurrence with bad consequences for S would not thereby be fortuitous for S. And very little that happens in the world is genuinely random, presumably, yet many things are incidental, lucky, and so on. So any attempt to carve out a specific sense of the term 'luck' or 'lucky' needs to carefully spell out which (or which combination) of these related notions is intended to be implied. Moreover, one needs also to make clear the point of distinguishing such

² See Riggs (1998: 452–72).

a sense of the term. After all, I claim to be initiating a 'defense' of my account of luck. Arguably, one could carve up the family of 'chance' notions in a great many ways. Why not just let 'a thousand flowers bloom'?

The answer is that I intend to put my narrowed notion of 'luck' to philosophical use in a very specific way. So it needs to have the right properties to do the job I want it to do. Specifically, it needs to be conceptually related to 'knowledge' in the right sort of way. It is one of the rare points of near unanimous agreement among epistemologists that the presence of luck in one's coming to believe, truthfully, that *p*, tends to undermine what would otherwise be knowledge that *p*. I am attempting to give an account of luck that preserves this judgment about the connection between luck and knowledge. However, a successful account will have to do more than merely have the consequence that its presence is typically incompatible with knowledge. After all, our intuitions about the incompatibility of luck and knowledge do not come with a detailed theory of luck. That is, our intuitions rely on a rough and ready understanding of luck, and so the correct account of knowledge-precluding luck must be a fairly 'core' notion within that family of luck-notions mentioned earlier—it must categorize paradigm cases of luck and non-luck correctly. Otherwise, it will appear to be no more than an interesting but esoteric notion dreamed up in an ad hoc fashion to solve a specific epistemological problem.

So what are the contenders, then, for such a notion of luck? Andrew Latus has done a nice job distinguishing two well-established, though not necessarily well-articulated, alternative theories of luck (Latus 2003: 460–75). There is the 'value + chance' model attributable to Nicholas Rescher's work, and the 'lack of control' model which Latus traces back to Thomas Nagel's famous work on moral luck (1993). According to Latus (2003: 465), 'Rescher views luck as a property of events that varies inversely with the likelihood of the event and proportionally to the value of the event.' Nagel, on the other hand, says, 'Where a significant aspect of what someone does depends on factors beyond his control, yet we continue to treat him in that respect as an object of moral judgment, it can be called moral luck' (1993: 59). These two brief quotes are sufficient to give the gist of the two views. Yet a third account of luck is offered by Duncan Pritchard, in terms of the counterfactual 'safety' of an event.

Herein, I will defend a version of the 'lack of control' (or 'control' for short) theory of luck. Of the two alternatives mentioned above, I will address only Pritchard's 'safety' theory of luck, leaving Rescher's account to the side. As both Pritchard and Latus point out, Rescher's view is subject to some fairly serious objections which, at the very least, require that the core notion of 'chance' be spelled out in considerably more detail before a more thorough assessment can be made.

The rest of this paper, then, will constitute a two-part defense of the 'control' account of luck. The first part will consist of a critique of Pritchard's 'safety'

theory of luck. I will argue that his modal definition of luck falls to decisive counterexamples, both as a core notion of luck, and as a crucial component in a theory of knowledge. The second part of the defense will be to respond on behalf of the 'control account' of luck to an important objection to it, and to show that not only does the objection fail, but its failure offers us some interesting insights into the nature of luck.

2. PRITCHARD'S SAFETY ACCOUNT OF LUCK

My goal in this section is to show that both Pritchard's account of luck in general, and his account of epistemic luck in particular, are mistaken. Pritchard's general account of luck consists of two principles. They are as follows:

- (L1) If event E is lucky, then it is an event that occurs in the actual world but which does not occur in a wide class of the nearest possible worlds where the relevant initial conditions are the same.
- (L2) If an event is lucky, then it is an event that is significant to the agent concerned (or would be significant, were the agent to be availed of the relevant facts). (Pritchard 2005: 128 ff.)

While (L2) raises some interesting issues of its own, I will not have anything to say about it for the moment. I will concentrate, as other recent commentators have, on (L1).

I intend to show that the conditions offered are neither necessary nor sufficient, as evidenced by various counterexamples. I think the case against the necessity of Pritchard's two principles is quite powerful, and it has been made by, among others, Jennifer Lackey. I will briefly recount her argument for this in a moment. As for the sufficiency of Pritchard's principles, I am skeptical here too, though the arguments for this position are not nearly as potent. Nonetheless, I will make the best case I can against the sufficiency of Pritchard's account because, even if one does not find the argument decisive, it still adds some weight to the overall argument against the safety analysis of luck. (I'm calling it the 'safety' analysis because of the similarity of its modal structure to safety accounts of knowledge, one of which he goes on to develop himself.)

Before moving on to these objections, let me add an important caveat. Pritchard has made it plain that he does not intend that his conditions either for luck or for knowledge be taken to be jointly sufficient.³ Rather, he hopes to have elucidated some necessary conditions that encompass the 'core' of these notions, but that might need to be supplemented in some minor ways to be fully adequate. Thus, while counterexamples to the necessity of his conditions are obviously a

³ In personal correspondence.

threat to his view, the mere existence of counterexamples to their sufficiency does not refute the view. One would need to show that the case described in the counterexample is such a central one that any adequate account of luck must count it as such. With this caveat in mind, I will proceed with the objections.

In a very penetrating review of Pritchard's book *Epistemic Luck*, Jennifer Lackey (2006) makes a very strong case against the necessity of Pritchard's two conditions on luck. To do this, one needs to show that an event can be (intuitively) lucky even though it occurs in all or nearly all nearby possible worlds. She does this by way of a counterexample, which she calls BURIED TREASURE. I shall briefly recount the important details of it so as to see the force of her argument.

BURIED TREASURE: Sophie, knowing that she had very little time left to live, wanted to bury a chest filled with all her earthly treasures on an island she inhabited. A suitable site had to be located on the northwest corner of the island—where she had spent many of her fondest moments in life—and it had to be a spot where rose bushes could flourish—since these were her favorite flowers. As it happens, there was only one spot on the island that met both these criteria. Sophie, being excellent at detecting rose-friendly soil, immediately located this spot and buried her treasure there. One month later, Vincent, a distant neighbor of Sophie's, was driving on the NW corner of the island—which was also his most beloved place to visit—and was looking for a place to plant a rose bush in memory of his mother who had died ten years previously—since these were her favorite flowers. Being excellent at spotting rose-friendly soil, he immediately located the very spot that Sophie had buried her treasure one month earlier. As he began digging a hole to plant the rose bush in, he was astonished to find buried treasure there.⁴

Lackey points out that Vincent's discovery is both clearly lucky, and yet also clearly happens in all nearby possible worlds. This is because all the features of the case that led both to Sophie's burying the treasure where she did as well as to Vincent's digging where he did are quite central features of the actual world. Thus, one has to go to a quite distant world in order to find one where either Sophie failed to bury her treasure in that spot or where Vincent failed to dig in that spot. Thus, only in a far distant world does Vincent fail to discover buried treasure, despite his doing so being a paradigm instance of luck.

Lackey claims that any number of such counterexamples can be constructed against Pritchard's view. She even offers us a formula for constructing them:

All one has to do is first choose a paradigm instance of luck, such as winning *Jeopardy* through a purely lucky guess . . . secondly, construct a case involving such an event in which both its central aspects are counterfactually robust, though there is no deliberate or

⁴ CF. Lackey (2006: 285).

otherwise relevant connection between them; thirdly, if there are any residual doubts that the event fails [Pritchard's first condition on luck, (L1)], add further features to guarantee counterfactual robustness across nearby possible worlds. Voilà: one has a counterexample to [Pritchard's account of luck]. (Lackey 2006: 286)

Counterexamples are useful, of course, in showing that there is something wrong with a view, but they are unhelpful in making further progress unless they also give us some idea of how to diagnose the failure. Lackey's commentary on her own counterexample points us in the direction of such a diagnosis:

The fundamental problem with such modal accounts is that counterfactual robustness can be ensured through a combination of features that is entirely fortuitous. For instance, an event that appears in both the actual world and all of the relevant nearby worlds can none the less be lucky because the relevant counterfactual robustness is achieved purely through a lucky combination of external events. (2006: 289)

Lackey notes that the problem lies in the fact that the counterfactual robustness of the lucky event is itself fortuitous in the problem cases she describes. There is no 'deliberate or otherwise relevant connection' between the two elements of the case whose conjunction yields the luck (e.g. Sophie burying the treasure in spot X and Vincent's digging at spot X). So, presumably, for an account of luck to get these Lackey-type cases right, it must differentiate between situations in which there is such a relevant connection between the elements of luck in the case, and situations in which there is not. Pritchard's safety account does not do this.

Let me turn now to a consideration of the sufficiency of Pritchard's conditions on luck. To show that these conditions are not sufficient to account for luck, one must describe an event which (1) obtains in the actual world, (2) fails to obtain in a wide class of nearby possible worlds, and yet (3) is nevertheless not lucky. It is hard (perhaps impossible!) to find cases like this where the non-luckiness of the event is intuitively very clear. But let me try anyway. Consider the following case:

Smarty is the valedictorian of her high school class who is about to take her computer-delivered college entrance exams. Despite her formidable intelligence, she decides to prepare for the upcoming exam by studying diligently and taking many practice exams. The night before the exam, she gets a good night's sleep, and awakens fresh, sharp, and ready to excel. She takes the exam and scores very highly. Unbeknownst to Smarty, however, a fiendishly clever hacker with debilitating test anxiety had decided prior to the exam to wreak vengeance on all the clever students about to take it. Driven mad by his inability to get into a good college because of his poor test scores, he has vowed that all those smarty-pants test-takers will suffer just as he has had to suffer. He compiles a list of all the high school valedictorians for that year, and hacks into the exam program. For the valedictorians, he replaces the usual questions with questions from an advanced college physics exam. As it happens, he accidentally skips Smarty's name on the list (despite being very careful not to

miss anybody), and so she gets the usual questions. As a result, she is the only valedictorian who did well on the exam.

The question to be settled is this: was Smarty's high score on the exam a matter of luck? For my own part, I find that I am pulled in two directions here. I am willing to say that Smarty is lucky that her questions were not changed, but that seems to me not to amount to saying that she is lucky to have done well on the exam.

To bring out this point more clearly, suppose that she had gotten the advanced physics questions. Forced to simply guess in every case, she nevertheless chooses the correct answer in virtually every case. Here I think it is clear that her doing well on the exam is due to luck. And yet, there seems to be no obvious difference between the two cases in terms of how they fare according to Pritchard's safety condition. In each case there is a large class of nearby possible worlds in which Smarty does not get many correct answers on the exam. In the original case, this is due to the fact that in most nearby possible worlds, the hacker does not make the mistake of missing Smarty's name. In the revised case, it is due to the fact that Smarty guessed differently in many of the nearby possible worlds. So even if it is not clear that in the original case Smarty is not lucky to have done well on the exam, there still seems to be a difference in degree between the luck involved there and the luck involved in the revised case, though there is no obvious difference even of degree in how 'unsafe' they are in the two situations. What this suggests to me is that the safety account of luck is not looking in the right place for the appropriate conditions for luck.

3. EPISTEMIC LUCK AND KNOWLEDGE

The first thing to point out about Pritchard's theory of epistemic luck is that he introduces a distinction that does not arise in his more general discussions of luck. He distinguishes between what he calls *veritic* epistemic luck and *reflective* epistemic luck (2005: 146 ff.):

Veritic epistemic luck: it is a matter of luck that the agent's belief is true.

Reflective epistemic luck: given only what the agent is able to know by reflection alone, it is a matter of luck that her belief is true.

Veritic luck is simply the application of his modal definition of luck to the specific event of coming to hold a true belief. Reflective luck, by contrast, is a slightly more convoluted application of the general principle. While there are some interesting issues surrounding Pritchard's understanding of reflective luck, like most of his commentators to date, I will focus primarily on veritic luck. The main reason for this is that the definition of knowledge that Pritchard endorses takes account only of veritic luck.

I hope to have cast doubt on his general modal approach to luck, but I want to show how this approach fails specifically in an account of knowledge. Pritchard defines knowledge as, essentially, non-veritically lucky true belief. So, putting his account of luck in place as the element that turns true belief into knowledge, we should get the following: S knows that p iff S believes p, p is true, and there is no wide class of nearby possible worlds in which S believes p when p is false. This is, in fact, more or less the penultimate version of Pritchard's safety account of knowledge. However, he is forced to revise it to make the constraints on knowledge much stricter by requiring that there be *no* nearby possible worlds *at all* in which S believes p when p is false. He does this under pressure from what he takes to be the correct analysis of lottery-type cases. Here is a paraphrase of Pritchard's account of luck in terms of modal safety (2005: 156):

S knows that p (where p is a contingent proposition) only if S believes p in the actual world, p is true in the actual world, and in all nearby possible worlds in which S forms her belief about p in the same way as she does in the actual world, S believes p only when p is true.

As with his account of luck, Pritchard takes this to provide necessary conditions that are the most central to knowledge, but perhaps not fully sufficient. The same caveats to the strategy of arguing by counterexample apply here as did in the discussion of his account of luck.

Because this is simply a special application of his more general account of luck, one would assume that Lackey's formula for generating counterexamples would serve to undermine this account of knowledge as well. Pritchard is claiming with this principle that knowledge implies a non-veritically lucky true belief, and that veritically lucky true belief implies a lack of knowledge. But, if Lackey is right, we should be able to construct cases wherein an agent comes to hold a belief in a counterfactually robust way, and the belief is true in a counterfactually robust way, but where the connection between the way the agent came to hold the belief and the truth of that belief are not connected in the right sort of way, and so their having come to hold a true belief is lucky, despite meeting the modal conditions of the safety principle.

Indeed, these cases are constructible. What they amount to are counterfactually robust Gettier cases. Consider:

Suppose the North Korean government undertook a publicity campaign to inform its citizens that all Americans are racists who particularly despise Asians. This is done very convincingly, so that most North Koreans come reasonably to believe that all Americans are racists who particularly despise Asians. Meanwhile, the CIA decides to send a covert operative to investigate the possibility of a covert nuclear facility in a rural Korean village. As it happens, there is only one agent who even comes close to having the training and skills necessary for this mission. He is a singularly competent individual,

the likes of whom comes along once in a lifetime. He also happens to be a racist who particularly despises Asians. This prejudice is a deep and longstanding feature of his personality, produced by a series of life-shaping and counterfactually robust incidents in his past. When the agent arrives at the village, a local villager spots him and immediately recognizes him as an American. The villager subsequently comes to believe that the person he sees is a racist who particularly despises Asians.

It looks as though the villager's belief is true, and that there are no nearby possible worlds where he comes to this belief in the same way and the belief is false. All nearby possible worlds where the villager comes to this belief in the same way he does in the actual world are worlds where the very same operative with the very same prejudices is the subject of that belief. This assumes that both the operative's skill set and his racism are deep properties that hold of him in all nearby possible worlds, and that no other American operative would have been successful in arriving at the village at all. Yet it is still a matter of luck that his belief is true. So Lackey's formula holds for constructing counterexamples to the safety account of epistemic luck as well. Consequently Pritchard's definition of knowledge is faulty because the villager's belief meets the requirements of his definition though he clearly does not know. More damning still, he doesn't know *precisely because* his being right is a matter of luck, which is the consideration that Pritchard's account is supposed to be sensitive to.

At this point, Pritchard might bring out his notion of reflective luck to finesse the objection rather than meet it head on. He could agree, for example, that the Korean villager's being right about the CIA agent was a matter of luck. But, he might add, the luck involved is reflective luck, not veritic luck. A belief is reflectively lucky, remember, when, given only what the agent is able to know by reflection alone, it is a matter of luck that her belief is true. In Pritchard's 'safety' terms, this means that the agent's belief fails to match the truth across a wide range of the nearest possible worlds where these worlds are ordered in a non-standard way '*solely in terms of what the agent is able to know by reflection alone in the actual world*' (Pritchard 2005: 175; emphasis in original). This move would allow Pritchard to accept the intuition that the villager's belief was lucky in some epistemically significant sense, without having to concede that it is not thereby an instance of knowledge.

Let us look more carefully at what is being proposed here. The villager believes that all Americans are racist against Asians, and so the possible worlds by which we judge reflective luck must be ordered accordingly. The closest possible worlds will be those in which all Americans really are racist against Asians, and in which an American ends up being seen by the villager (presumably, the villager does not know that the American in the original example is a CIA agent, so this cannot be used to order the possible worlds). Given this ordering, it looks as though the closest possible worlds will be heavily dominated by ones in which the villager's

belief tracks the truth. In which case, of course, it is not reflectively lucky. This leaves Pritchard unable to explain away our intuition that the villager is lucky to be right, while still claiming that he knows. It would appear that an appeal to reflective luck won't help here.

But what about my rather weak argument against the sufficiency of Pritchard's conditions on luck? Do these carry over to similarly undermine his account of knowledge? The astute student of Gettierology will already have anticipated a somewhat counterintuitive commitment of my critique at this point. For what would an epistemic analogue of the Smarty case look like? Here's a proposal, chosen randomly from among the various alternatives.

Imagine that I am traveling through a picturesque countryside and I pull off into a driveway to a farmhouse because I like to look at barns. Indeed, as I approach the house I can see a barn quite clearly off in the adjacent meadow. Having accomplished my goal, and having formed the belief that 'there's a barn,' I drive away. As you might have surmised, I am correct in my belief. Yet, the neighbors for miles around have secretly erected fake barns to fool just such casual barnspotters as myself, and had I chosen any other driveway I would have been fooled by one of their fake barns.

So here is a case in which, I submit, it is not obvious that it is a full-blown matter of luck that I came to have a true belief. Contrast this with a case in which I am, say, riding an express train that is going so fast that I cannot distinguish one building from another as they speed past my window. I am thinking nostalgically of my upbringing on a farm, and remember quite fondly the barn where I used to milk the cows. Glancing briefly outside, and being primed by my reminiscences, I believe of the next building that I see go flashing past, that it is a barn. And, of course, it happens to be a barn. Let us construct both this case and the previous one so that the truth of my belief is equally counterfactually robust. That is, the ratio of fake barns to real ones is roughly the same as the ratio of barns to other buildings that are flashing by my train window. So in both cases the nearby possible worlds are equally populated with instances of my believing 'there's a barn' where that belief is false. I submit that, intuitively, it is not nearly so obviously a matter of luck that I am correct in the original case as it is in the modified case.

But that would imply that it is not obvious that my belief in the classic barn façade case is *not* an instance of knowledge. But it is an item of contemporary epistemological dogma that discounting one's true belief in the barn façade case as an instance of knowledge is a *desideratum* of a theory of knowledge, not a point of criticism. True enough. I cannot offer much of a defense of this judgment on my part, other than to point out that I am not quite alone in my judgment. Both Bill Lycan and Stephen Hetherington have recently argued that the barn façade case is not a real Gettier case because it represents a genuine case of knowledge. I also have anecdotal evidence from a number of epistemologists who say that

they have always found the barn façade case to be the least persuasive of the standard menagerie of Gettier cases. For now, I must simply leave this as I did my initial criticism of the sufficiency of Pritchard's account of luck. Clearly, this criticism is not going to convince anyone on its own, but it is a significant part of the case to be made against safety accounts of luck and knowledge nonetheless.

4. 'OUT OF CONTROL' LUCK

Let me turn now to my own conception of luck that I take to be superior to the safety account. As I have indicated, the notion of luck I defend places agent control (or, rather, the lack thereof) at the forefront of the analysis. One has control over some happening to the extent that the happening is properly considered something the agent has *done*. As I have also already indicated, this imposes two separable requirements. First, the event has to be the product of the agent's powers, abilities, or skills. Second, the event has to be, at least in some attenuated sense, something the agent *meant to do*. This second requirement does not demand an actual conscious intention on the part of the agent, but it does mean that a goal or desire or intention must be guiding the exercise of one's powers, abilities or skills that brings about the event in question.

I think that even this rough description of the control account of luck is enough to see that it gives the right answer in ordinary, run-of-the-mill examples of lucky occurrences. Winning the lottery or making a lucky pool shot are prime examples of events that are not brought about by an agent's powers, abilities, or skills. They are simply the result of the operation of chance, which dictates that, given time, pretty much anything can happen, and occasionally will. Other examples, like finding buried treasure while digging in your flower garden, are doubly lucky. Not only is it not the result of your powers, abilities, or skills that you found buried treasure, your desire or intention or goal of finding buried treasure, if you even had one, played no role in your choosing to dig right there.

Furthermore, I think that this account of luck handles easily the cases discussed previously with regard to the safety account of luck. The cases I have in mind are the buried treasure case, guessing the right answer on jeopardy, and the two cases involving Smarty the valedictorian. Consider first the buried treasure case. Recall that Vincent was looking for an appropriate place to plant roses in memory of his mother. He had no intention of looking for buried treasure, so his finding it is clearly lucky on my account. Moreover, his finding buried treasure was in no way due to his abilities, etc. Guessing the right answer on jeopardy is a different matter. In that case, presumably, you desire to win the game, and that desire is very much governing your deployment of your powers, abilities, and skills in coming up with an answer. But, by hypothesis, those powers, abilities, and skills

were not sufficient for you to know the answer. Therefore, your giving what turns out to be the correct answer is not due to any such powers, abilities, or skills.

And what of my two examples intended to show the insufficiency of the safety account of luck? In the first example, Smarty the valedictorian scored well on her college board exams because her nemesis inadvertently failed to change her questions. What does the control theory of luck say about this case? On the one hand, her score very much reflects her powers, abilities, and skills. She knew the answers to the questions given because of her native intelligence and her study. On the other hand, she would have failed miserably had the hacker not overlooked her exam and had changed her questions as he did for the others. Smarty had no control over this aspect of the situation, which seems to be crucial to her success. Notice that there is no such ambiguity in the second Smarty case, in which the hacker did change her questions, and Smarty managed to guess correctly on most of them. In that case, her high score on the test does not in any way reflect her powers, abilities or skills. This generates some ambiguity concerning whether Smarty does or does not have control over her performance on the exam. This ambiguity, I maintain, explains the difference in our intuitive responses to these two cases.

5. LUCK AND EXPLOITATION

Let me now turn to an objection to the control account of luck recently offered by Andrew Latus (2003). The objection is both very simple and very powerful. If luck is simply a matter of some event or state of affairs being out of one's control, then all manner of mundane natural occurrences, like the rising of the sun each morning, should be counted as lucky. For that matter, many unnatural occurrences, like the continued functioning of the electricity in my house every day, are not in my control, and thus should be counted as lucky on my view. This seems counterintuitive, to say the least.

This objection is to the sufficiency of the control theory of luck. Something's being out of S's control does not suffice for its being a matter of luck, or so it appears. Latus says,

The problem for the control account is that we can conceive of cases in which the occurrence of some event is both entirely out of a person's control and of some value, yet that event is not properly described as lucky. Consider, for example, that, while it is entirely out of your control whether the sun will rise tomorrow, it does not seem correct that you are lucky the sun will rise tomorrow. (2003: 467)

In a previous paper (Riggs 2007), I acknowledged this point, and conceded that the control account of luck was not sufficient—that mundane occurrences outside our control were not matters of luck for us, and so some kind of condition requiring that the occurrence be remarkable in some way was still necessary. In

this section, I will argue that I was too quick to concede this point. With certain modifications, the control account can be made sufficient, and can account adequately for these events that are mundane but out of our control.

However, the path is circuitous, and I will beg the reader's indulgence in following me around the bend with my assurances that we will eventually return with the resources necessary to reply to Latus's objection. The first thing we need to do is to consider the following scenario.

The scenario involves two gentleman adventurers in the golden era of such adventures. Let us call them Indiana Jones and New Jersey Smith. New Jersey Smith plans an expedition into the wilds of Africa where certain tribes of Africans with exotic customs were known to live. Smith is constrained by his schedule and finances to make this trip during a particular month of a particular year. He proposes the trip to his fellow adventurer Jones, including the specific times that he means to travel. Jones agrees to tag along. As it happens, the particular tribe that lives in the area that Smith and Jones visit has a custom of sacrificing people from outside the tribe on the equinoxes of the year. The autumnal equinox happens to fall during the time that Smith and Jones are in the area, so they are captured and held until that day so that they can be sacrificed. When the day of the autumnal equinox dawns, the tribe readies their captives for sacrifice at midday. As the tribesmen approach to kill them, Smith says to Jones, 'Only a miracle could save us now!' At that precise moment, there is a total eclipse of the sun. The members of this tribe always take such exotic natural occurrences to signal the anger of their gods at them for whatever they happen to be doing at the moment. Consequently, they set their captives free. Smith says to Jones, 'that solar eclipse was an amazing stroke of luck!'

Smith certainly seems to be right—it was very lucky for him that the solar eclipse happened when it did. So it is possible for such nomically certain events to be lucky for someone. So far we have no problem for the control theory. But there is more to the story. You see, Jones has a surprising and enlightening reply to Smith's breathless declaration of their good fortune.

Smith says to Jones, 'that solar eclipse was an amazing stroke of luck!' Jones replies, 'Don't be absurd! There was nothing lucky about it. I knew all along that these people would likely try to sacrifice us on the equinox if we were captured, but I also knew that there was a total eclipse of the sun due on that very day, and that this tribe would react to that event by letting us go. Did you really think I would be stupid enough to fall into such a situation without having a plan to extricate myself?'

According to Jones, the eclipse was *not* a matter of luck for *him*. But how can this be? He was no more in control of the eclipse than was Smith. And the eclipse was equally significant to him as it was to Smith. By any criteria that we have

appealed to so far, these two should be on a par with respect to the luckiness of the eclipse. Yet both seem right: Smith was lucky that the eclipse happened, and Jones was not.

Figuring out what to make of the luckiness or not of the eclipse is a subtle business, and it might help to pin down some of the clearer implications of the scenario first. It seems clear, and accords with the control theory of luck, that Smith is *lucky to be alive*, while Jones is not. Jones's *being alive* is a consequence of his deploying his powers, abilities, and skills in planning against his possible capture by the tribesfolk. Smith's being alive, on the other hand, has nothing to do with his powers, abilities, or skills. So far, so good. But what of the eclipse?

It is tempting to say that, since each adventurer's survival depends crucially on the eclipse, that each one's survival is lucky if the eclipse is lucky for him. This, at least, gets the facts right. Both the eclipse and his subsequent survival are lucky for Smith, and neither is lucky for Jones. On this view, Smith's survival is lucky for Smith because it is a consequence of the occurrence of the eclipse, which was lucky for Smith. The luck entered into the picture at the occurrence of the eclipse, and was simply carried along, as it were, down through the causal consequences of that lucky event.

The argument for this view would go as follows: It is intuitively clear in the scenario described that Smith's survival is a matter of luck for Smith and Jones's survival is not a matter of luck for Jones. But what distinguishes the two that can account for this difference? The main salient difference seems to be that Jones knew about the eclipse and what effects it would have on the tribesfolk, and Smith did not know this.⁵ But this suggests that events that are out of our control can nonetheless fail to be a matter of luck for us if we know about them in advance. This seems to deal a substantial blow to the control theory of luck. Technically, one could add an additional necessary condition to the definition of luck so that an event is lucky for S iff the event is out of S's control and S doesn't know about it, but this definitely has an ad hoc feel to it. It abandons the idea that the central feature of luck is the lack of control, and makes a lack of knowledge of the event or an inability to anticipate the event a coequal partner in its definition.

I think we can avoid this conclusion if we look at a slight variant of the scenario described above. Suppose that Smith knew, just as well as did Jones, that an eclipse was due that day, but he did not know that such an eclipse would save his life if he were captured. In other words, he did not realize (or at the least did not consider in his planning of the trip) that the tribesfolk would react the way they did to the eclipse, nor did he realize (or at the least did not consider in his planning of the trip) that it would save his life were they to be captured. Would we still say that Smith was lucky to be alive at the end? Absolutely. Would we still say that Smith was lucky that an eclipse occurred that day? I'm not sure intuition

⁵ Nothing hangs on Jones's awareness of these things actually counting as *knowledge*, as far as I can tell. All that matters is that he reasonably believes them.

speaks as clearly here as in the previous cases, but I think it is reasonable to say that Smith was lucky that the eclipse occurred, even though its occurrence was not a surprise to him. For now I will assume this is right, but I will return to this judgment shortly.

So, if I am right that Smith is still lucky that the eclipse occurred, even though he knew it would, then we have misdiagnosed above what made the eclipse lucky for Smith but not for Jones, since in the modified scenario they both knew about the eclipse, but it was lucky for Smith and not for Jones. What, then, is left to account for the fact that Jones's survival was not a matter of luck but Smith's is? I think we can find the answer in my careful description of the last scenario. I said that Smith knew about the upcoming eclipse, but he 'did not realize (or at the least did not consider in his planning of the trip) that the tribesfolk would react the way they did to the eclipse, nor did he realize (or at the least did not consider in his planning of the trip) that it would save his life were they to be captured.' What seems to distinguish Jones from Smith, and makes Smith lucky to be alive but not Jones, is not that Jones *knew* about the eclipse and whatnot, but that he *exploited* those facts to his own advantage. That is to say, he took them into account and planned a course of action that assumed that those things would occur. And the outcome that resulted, his survival, was a consequence of his having taken account of and exploited those facts. Thus, he was in control of his own destiny, even though he was not in control of every event that played an important role in that destiny.

Once this is pointed out, it seems obvious. After all, to the extent that we are ever in control of our actions, it is only against the backdrop of our environment. We can successfully plan and carry out actions only if the world cooperates by behaving more or less as we expect it to. And when the world does act as we expect it to, it is not a matter of luck for us that it does so. And when we successfully conclude a plan of action that depends upon the world acting in a way that we reasonably expect it to, then the successful conclusion of that plan is not a matter of luck for us either.

So it looks as though we can preserve the natural view that for each of Smith and Jones, his survival's being a matter of luck for him depends on whether the eclipse was a matter of luck for him. The eclipse was a matter of luck for Smith because it was both out of his control and unexploited by him. The eclipse was not a matter of luck for Jones because, though it was out of his control, he nonetheless exploited its occurrence to procure his survival.

This conclusion indicates that a modification of my definition of luck is required. Up until now, I have been assuming roughly the following:

E is lucky for S iff E is (too far) out of S's control, and

S controls E iff S brought E about (where this implies both that E was the result of the application of S's powers, abilities, or skills, and E was not inadvertent with respect to S).

Putting these together, we get:

E is lucky for S iff it is not the case that S brought E about (where this implies that either E was not the result of the application of S's powers, abilities, or skills, or E was inadvertent with respect to S).

But the preceding discussion shows that an event can be both beyond someone's abilities to produce or even affect, and yet still fail to be lucky. That is Jones's situation with respect to the eclipse. What must be recognized here is that S's lacking control over something is not *sufficient* for it to be lucky for S. Thus the biconditional linking the two is incorrect. Instead, we must define luck this way:

E is lucky for S iff

- (a) E is (too far) out of S's control, and
- (b) S did not successfully exploit E (prior to E's occurrence) for some purpose.

There is obviously a lot of work that needs to be done in getting clearer on what counts as 'exploiting' something in this context. For now, I will have to rely on an intuitive understanding gained from the foregoing examples.

It is important to realize that this does not affect the definition of credit, which is still in terms of control, not luck in general. It would be a mistake to start giving Jones *credit* for the eclipse, just because he wasn't lucky that it happened.

We have now come around the last curve in the road, and the pieces are in place to provide a response to Latus's objection that mundane events that are out of our control nonetheless seem not to be lucky. In fact, there are two aspects of the kind of example he provides that are problematic. On the one hand, the event in question is nomically necessary, and so well beyond my influence in any way; and, on the other hand, its occurrence is of no particular significance to me. Let me address these two aspects separately.

I hope to have shown that it is possible to exploit nomically necessary events, even when one has no actual control over them. This point generalizes to other events that are not nomically necessary, but are such that I cannot affect them. And when one successfully exploits such an event, it is not a matter of luck for you that the event occurs. Thus, by modifying the control account, I have introduced the theoretical resources necessary to get these cases right.

The problem remains, though, that there are all kinds of events that are out of my control and that I do *not* exploit, yet seem not to be a matter of luck because they do not impinge on my life in any way—the movements of the stars in a distant galaxy, for example. The response I'd like to make here is to bite the bullet and insist that the intuitions give way to the theory in this case. There is a sense of 'luck' that is entwined with our concept of 'fortune,' as in one's fate, or what befalls one in life. It is the pressure of this sense of luck that makes us

uneasy calling something that is completely irrelevant to our fortunes a matter of luck. Yet this sense of luck is completely separable from the sense of luck I am trying to get clear on, which is the sense of luck that is conceptually connected to credit and responsibility, rather than fortune. Yet I fear this may strike some as too presumptuous, especially for a view that is still in the early stages of its development and defense. The sell-out alternative, then, is to acknowledge the need for something like Pritchard's principle (L2), which added the additional requirement that a lucky event had to be one that was significant to the agent concerned. So now we have the following:

E is lucky for S iff

- (a) E is (too far) out of S's control, and
- (b) S did not successfully exploit E for some purpose, and
- (c) E is significant to S (or would be significant, were S to be availed of the relevant facts).

6. CONCLUSION

Put simply: the control theory of luck is better than the safety theory offered by Pritchard. I have tried to convince you that the safety account is not simply wrong, but fundamentally misguided. That is, I think that it is simply looking in the wrong places for luck. Ironically, my diagnosis here mirrors Pritchard's own diagnosis of control theories like my own. He says that control theories of luck seem plausible because they mimic in a wide range of cases the results that one gets from the safety theory. I agree, but conclude that this is why the safety theory *appears* plausible. It is the control theory that both gets the cases right and preserves the intuitive conceptual connections among luck, credit, responsibility, and other notions in that family. Consequently, it is the better account of luck to appeal to in an anti-luck theory of knowledge.

I think it is also fairly easy to see why such an anti-luck condition would make knowledge valuable to us in a way that merely having a true belief is not. On my view, roughly, one knows that *p* when one comes to have a true belief that *p* in a way that is sufficiently under one's control. This makes such a theory a version of the so-called 'credit theory' of knowledge. A crucial feature that distinguishes knowing from other cognitive states, on such a view, is the extent to which your having the other features of knowing (holding a true belief) is due to your own powers and abilities. A growing coterie of epistemologists have argued that this 'credit-worthiness' feature of knowledge is what makes it distinctively valuable.⁶ The general idea is that we value bringing about a positive state of

⁶ See e.g. Greco (2003); Sosa (2008); and Riggs (2007).

affairs (hitting a bulls-eye, for example) by way of our skill and effort more than we value that same state of affairs coming about through luck. The account of luck that I have articulated and defended puts meat on the bones of these kinds of contentions. Providing a theory of luck in terms of just the kind of control necessary to be credit-worthy in coming to hold a true belief brings the centrality of luck-avoidance and the importance of agency in our concept of knowledge harmoniously together.

REFERENCES

- Greco, J. (2003). 'Knowledge as Credit for True Belief,' in Michael DePaul and Linda Zagzebski (eds.), *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Oxford University Press.
- Lackey, J. (2006). 'Pritchard's *Epistemic Luck*,' *Philosophical Quarterly*, 56: 284–9.
- Latus, A. (2003). 'Constitutive Luck,' *Metaphilosophy*, 34: 460–75.
- Nagel, T. (1993). 'Moral Luck,' in Daniel Statman (ed.), *Moral Luck*. New York: SUNY Press.
- Pritchard, D. (2005). *Epistemic Luck*. Oxford: Oxford University Press.
- Riggs, W. (1998). 'What Are the "Chances" of Being Justified?,' *The Monist*, 81: 452–72.
- (2007). 'Why Epistemologists Are So Down on their Luck,' *Synthese*, 158 (3): 329–44.
- Sosa, E. (2008). *A Virtue Epistemology*. Oxford: Oxford University Press.

This page intentionally left blank

PART II

TRUTH AND EPISTEMIC
APPRAISAL

This page intentionally left blank

The Values of Truth and the Truth of Values

Michael P. Lynch

It is not that hard to be skeptical about value. You can't see values. You can't touch them. And many think we can't even define them. Partly as a result, our disagreements about value have an interminable and intractable feel. Philosophically speaking, skepticism about value seems an easy sell.

At least this is the case when we are talking about the type of value that most philosophers worry about—moral value. But we have other values besides moral values. And one of the most basic of these other values is the value of truth.

In this paper, I want to address two questions. What does it mean to say that truth is a value? And how seriously—from what we might call the meta-normative point of view—should we take this value? In a certain sense to be explained, I think we have no choice but to take it very seriously, simply because we lack any standpoint from which we can make skepticism about the value—or, as we shall see, values—of truth intelligible.

1. TRUTH AS A GOAL OF INQUIRY

The claim that 'truth is a value' can mean quite different things. Two in particular need sorting out.

One thing we might mean is the value of true beliefs. It is true beliefs we have in mind when we say that truth is a *goal of inquiry*.¹ Here's how I understand this. First, by 'inquiry' I mean the range of epistemic practices we engage in when asking and answering questions, whether banal ('where did I put my other sock') or sublime ('Can something come from nothing?'). Second, by 'goal' I don't mean something that is necessarily desired. It is true that when we say that something is a goal, of inquiry or anything else, we sometimes take this to imply that we desire it.² But truth often seems the faintest of human

¹ Note that I say 'a' not 'the'. I remain neutral on whether true belief is the only aim of inquiry. It may be its most fundamental aim; or simply one among several fundamental aims.

² Thus Sosa (2001) worries about whether we can desire the truth and nothing but the truth, as does Piller (forthcoming). I too have suggested elsewhere (Lynch 2004a and 2004b) that we can,

passions, undesired or actively avoided. In any event, these are matters of human psychology and best left to the experts. So let's instead say that truth is a goal of inquiry in the sense of being a proper end of our epistemic practices, where 'proper end' means something that is *worth* pursuing, whether we in fact do pursue it.

A natural way of characterizing the end in question is James's demand that we shouldn't just 'believe the truth!' we should 'shun error!' It is not just good to believe the truth, then, it is good to not believe what is not true; that is, it is good to believe only the truth. Thus one might suggest that

(TE) It is *prima facie* good that one believes all and only what is true.

I use the word 'good' here to qualify a general state of affairs: the state of affairs of believing all and only true propositions. Note that (TE) does not say that it is good for one's actual beliefs to be true. One's *actual* beliefs might be absurd. The point is that it is good to believe *whatever* turns out to be true and only what is true.

The intuitive thought behind (TE) runs on all fours with the thought that it is *prima facie* good to be omniscient. And that seems plausible; it is good to be God, as it were. Understood in this way, however, the value of truth is too much of an ideal. After all, humans aren't gods, and no human can believe everything that is true. Accordingly, it seems to make sense to relativize the truth-goal to a restricted set of propositions:

(TG) It is *prima facie* good that, relative to the propositions one might consider, one believe all and only those that are true.³

Unpacking the embedded modality here will be tricky, but the point should be clear: (TG) doesn't say that it is good for the propositions that I actually consider to be true.⁴ Rather, the point is that it is good, relative to the set of

arguing that this fact is good *evidence* for thinking we believe that having all and only true beliefs is good. But the key issue is not the psychological one; what matters is whether true beliefs are good or valuable, not whether they are desired. This betrays my conviction that what we desire is not equivalent to what is valuable, or even what we value.

³ Here I am influenced by Ernest Sosa (2001) and Marian David, whose work on this matter has helped (and provoked) me in numerous ways; see David (2001); see also my reply to David (2005) in Lynch, 2005a.

⁴ Alan Millar has suggested to me that one might capture the spirit of (TG) instead as:

(TG*) It is a *prima facie* good that, when one has a belief as to whether or not p, it be the belief that p iff it is true that p, and the belief that not-p iff it is true that not-p.

This has the advantage of not containing the restriction to beliefs that one might consider. But, unlike (TG) itself, it would seem to commit us, via classical logic, to: for all p, either p is true or not-p is true. Whether this is a problem (given that TS might be thought to commit one to the same), I leave to one side here. In any event, for purposes of what follows, the choice between (TG) and (TG*) is immaterial.

propositions I am able to consider, that I believe all and only those that are true.⁵

Principle (TG) claims that the state of affairs of believing what is true and only what is true on any matter that might come to hand is always good; but it is *always prima facie good*. Something is good in this way when there is always something to say for it; when it is good considered by itself but not necessarily good all things considered. Almost everything that is good is *prima facie good*. Keeping a promise, for example, is always good, other things being equal; but it is best to break a date to save the ubiquitous drowning child. Likewise, while it is always good that one believe only the truth, it is not always good, all things considered. This reflects the fact that while truth is a value, it is not our only value; and sometimes our values, whether they are cognitive or moral, conflict. Thus it might be good, all things considered, to believe something false when, for example, it is justified by the evidence. And not only are there good falsehoods, there are also bad truths. There are all sorts of trivial truths that are not worth believing, given my limited intellect and time. Nonetheless, were these limits not in place—were it to be the case that believing the truth was cost-free, so to speak—then it would be good to believe all and only what is true. And that is just to say that believing what is true is a *prima facie good*.⁶

It is helpful, in discussing these matters, to distinguish between the ultimate end or value that *governs* the practice and the more immediate aims that are justified in light of this ultimate value. The former is the light, however practically unreachable, by which the practice steers, so to speak. The latter are the direct goals practitioners typically aim to achieve. In saying that true belief is a proper end of inquiry, we take it to be an aim of inquiry in the first sense. An individual inquirer rarely has (TG) as a conscious aim in her everyday epistemological life. And even when she does, she cannot achieve that end—in the sense of (TG)—directly. One does not simply will oneself to believe the truth. Rather, we pursue truth indirectly, by pursuing those beliefs backed by reasons and supported by the evidence. Yet these more practical and immediate goals only make sense in the light of the value of truth. If it was not good to believe what is true, then the pursuit of justification would be unimportant. We pursue the truth indirectly by directly pursuing—aiming at—justification.

What sort of goodness is involved in (TG)? The question of how to sort goods is a thorny one, but we need not enter that thicket here. Instead, let us note

⁵ (TG) and (TE) are obviously different: one is an absolute ideal; the other is relativized to propositions I am able to consider. But intuitively, (TG) is justified by (TE). For unless it were good to believe all and only what is true, it would be difficult to see why it would be good to believe all and only what is true on any matter that comes to hand. That is, if the unrestricted ideal was not good, it is hard to see how restricting it could be.

⁶ For more remarks on the structural relationships between the value of truth and other values, see Lynch 2004a: ch. 4.

that it is at least initially plausible to think that true belief is not a moral good, because what is morally good is generally either a subject or an object of direct responsibility. But believing what is true and only what is true is not something we are directly responsible for. What we are directly responsible for is how we go about pursuing true beliefs (or not) in our everyday epistemic life. Thus in the sense described by (TG), true beliefs seem better described as an epistemic or, if you prefer, a cognitive good.

2. TRUTH AS A NORM OF BELIEF

The second idea we might be talking about when discussing the value of truth is the value of believing what is true. This is presumably what James was thinking of when he noted that truth is the good in the way of belief (1975: 42). He might have better said that truth is the right in the way of belief; for the idea here is that true beliefs are right or correct. That is,

(TN) It is correct to believe $\langle p \rangle$ if and only if $\langle p \rangle$ is true.

Where (TG) ascribes value to a general state of affairs, (TN) ascribes value—what I'm here calling 'correctness'—to *believing* true propositions.⁷ There are of course similarities between the two principles as well. Here too the value in question seems more cognitive than moral. While I can be responsible for how I go about forming beliefs, I am not, strictly speaking, responsible for the belief itself.⁸

So unpacked, (TN) appears to be not only true, but a truism.⁹ But it is not vacuous or trivial. For on the intended reading 'it is correct to believe $\langle p \rangle$ ' and ' $\langle p \rangle$ is true' are not merely two ways to say the same thing. For what is true is the propositional content of the belief, while what is correct is the believing of that content. Thus *the two sides* of (TN) state different facts; while (TN) *as whole* claims those facts are co-extensive.

(TN) might be thought to tell us that truth is the aim of belief. But this is at best a metaphor. As Wedgewood, notes, beliefs aren't little archers aimed at

⁷ Again, David 2001 and 2005 and Lynch 2005a are instructive.

⁸ I use the word 'correct' here to help us distinguish (TN) from (TG). But don't be misled by the thought that what is correct is always an action. Beliefs aren't actions, at least in the typical sense of that term. I can't, for example, simply will myself to believe that George Bush wasn't President in the direct way I can will myself to raise my arm. Thus in saying that it is correct to for you to believe that p when it is true, I am not to be understood as saying that your true beliefs are actions done well.

⁹ We might wish to add that believings are themselves only *prima facie* correct. At least this is what we should say unless we have an airtight argument that there are no other norms, cognitive or otherwise, that operate over belief. Justification and rationality, for example, are normative, and they operate over belief. Moreover, they can conflict with the norm of truth—what is justified isn't always true. Nor is believing what is false always irrational.

truth (2002: 267). Moreover, in saying that beliefs aim at the truth, we aren't saying that in deciding what to believe, I must somehow expend effort in trying to believe what is true. This means that in the literal sense, truth is not an aim or goal of belief. It is not something beliefs strive for.¹⁰ *True belief is the aim of inquiry but beliefs don't aim at the truth.* Nonetheless, as a number of recent writers have urged, (TN) does tell us something about belief (Boghossian 2003; Velleman 2000; Shah 2003; Wedgewood 2002). Namely, it tells us that belief's basic norm or standard of correctness is truth.

Three considerations suggest this is, moreover, a constitutive or essential fact about belief. First, the fact that truth is a norm of correctness for believing is part of what distinguishes believing from other cognitive attitudes. Imagining, assuming, and hoping, for example, are each governed by norms—assumptions can be justified or not, imaginings can be sharp or vague, hopes can be rational or irrational. But neither imagining that *p*, assuming that *p*, nor hoping that *p* are properly evaluated in terms of truth. Believing is.

Second, believing that *p* is not only properly evaluated in terms of the truth of $\langle p \rangle$, it is indirectly *responsive* to its being true. In the typical conscious, deliberative case, it is so by being directly responsive to evidence for $\langle p \rangle$ (Shah 2003). And this suggests, third, that truth is not just a norm of belief, it is a basic norm. For we take it to be correct to believe what is based on evidence *because* beliefs based on evidence are likely to be *true*, and thus the value of truth in this sense is more basic than the value of believing what is based on evidence.¹¹

In sum, that truth is the norm of belief serves to distinguish belief from other cognitive attitudes. And this norm seems more basic than the demand to believe what is justified or based on evidence. Thus it is plausible that (TN) is a necessary, constitutive fact about belief.¹²

Recently Shah (2003) and Shah and Velleman (2005) have gone farther and claimed that (TN) is a straightforward *conceptual* or analytic truth. As Velleman earlier put it:

The concept of belief just is the concept of an attitude for which there is such a thing as correctness or incorrectness, consisting in truth or falsity. For a propositional attitude to

¹⁰ Marian David (2005) and Nishi Shah (2003) both make this point.

¹¹ This point also illustrates the difficulty in maintaining that epistemic rationality or justification is an *equally basic* norm of believing as truth. For to maintain that view, one would have to deny that what makes it correct to believe what is justified or epistemically rational is that justified beliefs are likely to be true. In other words, one would have to deny the seemingly necessary link between the value of justification and justification's truth-conduciveness. For more on this score, see Lynch 2004a. Note that *this* issue, which concerns which norms are basic on *acts* of belief, is distinct from the issue of whether there are other goals of *inquiry* besides having true beliefs. For this latter issue, discussed further below, see DePaul (2001) and David (2001).

¹² Some might say more cautiously: it is a constitutive fact about full-fledged beliefs: perhaps animals might have proto-belief states without being subject to any norms, no matter how trivial. I leave this difficulty to one side here.

be a belief just is, in part, for it to be capable of going right or wrong by being true or false. (2000: 16)

So, as Shah suggests, one must accept the ‘authority of truth over one’s cognition when one views it as a belief’—simply because this is a conceptual consequence of taking your cognitive activity to be a belief at all (2003: 474). This is certainly reasonable, but for purposes of our discussion we needn’t go so far. We need only say that given (TN), those who have beliefs, and have the capacity to recognize that they do, are implicitly committed to accepting that truth constitutes a normative standard for belief. Whether or not this is, as Shah believes, because (TN) is a conceptual truth, or whether it is simply a fact about belief, we can leave to one side.

Although they are distinct claims, (TN) and (TG) are interdependent. One connection between them consists in the structure of their justification (Lynch 2005a; cf. David 2005; McGrath 2005). But they are also connected in terms of how commitment to one can drag in its wake commitment to the other.

For example, we’ve noted that (TN) implies that truth constitutes a normative standard for belief. A consequence of truth being a normative standard of belief is that the having of that property plays a regulative role for any practice that aims at producing belief. Since inquiry is just such a practice, truth plays a regulative role for inquiry. A property P plays a regulative role in a practice when, just by virtue of participating in that practice, one is normatively committed to regulating one’s moves in the practice by one’s judgments about what has or lacks that property (Wedgewood 2002: 268). Thus the property of being a winning chess move is regulative of chess: in playing chess I am committed to regulating my moves by my judgments of what is or isn’t a winning move. Likewise, in figuring out what to believe—that is, when engaging in inquiry—I am committed to regulating my doxastic practices by my judgments about what is or isn’t true. Indeed, I am regulated by the truth in inquiry in the most direct possible way: the recognition that p is true is a decisive reason to believe it (Shah 2003).

It is a quick step between being committed to doing what is correct and being committed to the goodness of that which is correct. If my action is morally correct just when what I do is right, then clearly, if I engage in moral deliberation, that is, if I am *trying to figure out what to do, morally speaking*, my activity is governed by the principle that it is good to do what is right. Engagement in a goal-directed practice commits me to the value of the goal I so pursue. Likewise, if my believing is cognitively correct just when what I believe is true, then if I am engaged in inquiry, that is, *I am trying to figure out what to believe*, then I am normatively committed to my doxastic practices being governed by (TG). Call this the *trivial connection principle*: If I am committed to (TN), and I engage in inquiry, I am committed to (TG).

3. THE META-ETHICS OF BELIEF

So there are two values of truth: the value of having true beliefs, or (TG), and the value of a belief's being true (TN). There is much more to say about the character of both values, but I now want to turn to our principal question: What sort of meta-normative stance should we take towards these values? Or perhaps we should say: what sort of meta-normative stance *can* we take towards them?¹³

Let's begin with (TN). Since (TN) is, we've claimed, a description of a norm of belief, asking about the ontological status of this norm amounts to asking for a meta-ethics of belief. (TN) says that beliefs are correct when true; for ease of exposition, let's call the correctness (TN) speaks of *alethic correctness*. Briefly put, the following options, at least, seem available towards alethic correctness:

Non-naturalism: It is a *sui generis*, non-natural objective fact that beliefs are correct if and only if true.

Naturalism: It is a natural, objective fact that beliefs are correct if and only if true.

Error Theory: It is (literally) false that beliefs are correct if and only if true.

Expressivism: It is a non-factual matter whether beliefs are correct if and only if true, since to describe a belief as correct is not to state a fact about it but to express a sentiment or attitude.

One version of naturalism about alethic correctness holds that true propositions are correct to believe only as a means towards something else (see e.g. Papineau 1999). I've argued against this claim elsewhere (Lynch 2004a); here my interest is in skepticism, not naturalism. Thus I won't be worrying here about either of the first two positions.¹⁴

¹³ I have characterized both values in a strong biconditional form. The arguments that follow, however, are largely independent of my own characterization of these values.

¹⁴ Here's one way of putting the worry about taking alethic correctness to be a purely instrumental value. To believe that it is amounts to holding that:

- (i) It is correct to believe <p> if and only if believing <p> is a means to X.

Now if we assume that an advocate of (SN) will accept (TN), then given

- (TS) <p> is true if and only if p,

We can derive from their position, via transitivity of the biconditional:

- (ii) <p> is true if and only if believing <p> is a means to x.

Now presumably, any such position will read (i) and (ii) as follows:

Of the latter two, expressivism seems the far more interesting position. In part this is because error theory about alethic correctness is presumably already ruled out if (TN) is a constitutive fact about belief. It can't be false to say that a belief is correct when true if it is a necessary truth about belief that beliefs are correct when true. At least so it seems to me.¹⁵ Thus in what follows I'll take expressivism about alethic correctness as my primary example of skepticism about that value; but I also take the arguments raised below to apply to any form of skepticism about alethic correctness.

As indicated, expressivism about (TN) amounts to saying that the normative import of ascribing truth to a belief is an expressive, rather than descriptive affair. The correctness of a true belief is not a property that the belief has *per se*, but an expression or projection of our own desires and sentiments. Consequently, it is not a new and mysterious feature of the world that the naturalist can't explain. It is no feature of the world at all, but a feature of us. Thus, for example, a toy version of the view would take (TN) to mean:

(TN_e) Hooray for believing $\langle p \rangle$ if and only if $\langle p \rangle$ is true.

Since the left hand-side ('Hooray for believing $\langle p \rangle$ ') lacks truth-value, the whole of (TN_e) presumably does as well. And this of course brings up the usual problems faced by expressivists, for at a minimum it demands an explanation for the two associated conditionals, that is:

(1) Hooray for believing $\langle p \rangle$ if $\langle p \rangle$ is true.

And (even more oddly)

(2) $\langle p \rangle$ is true if hooray for believing $\langle p \rangle$.

It is not clear what these statements mean, and thus they seem ill-suited to act in an analysis of (TN).

Yet this familiar type of worry is not the main problem facing the alethic expressivist. The main problem is that expressivism about (TN) is self-undermining.

(i*) It is correct to believe $\langle p \rangle$ in virtue of believing $\langle p \rangle$ being a means to X.

(ii*) Believing $\langle p \rangle$ is a means to x in virtue of $\langle p \rangle$ being true.

That is, the naturalist will claim that the alethic correctness of a belief is explained by its having some instrumental value, and its having that value is in turn explained by the proposition believed being true. Nonetheless, if (TS) and (TN) are accepted, (ii) remains true. And that means that the naturalist must make a plausible case that there is *prima facie* instrumental value of a specific sort that attaches to believing any particular true proposition. (Note that this is distinct from arguing that there is instrumental value to having all and only true beliefs.) Moreover, if their view is not to slip back into non-naturalism, they must argue that this is the *only* type of value that can attach to believing what is true. And that seems difficult to do.

¹⁵ It is possible that some form of fictionalism about alethic correctness might avoid this problem, but I won't explore the issue here.

In order to state expressivism, one must express a belief about expressivism, and in doing that, one is committed to (TN). That is, to assert that,

- (3) (TN) is neither true nor false

requires that you *believe* that (TN) is neither true nor false. But if (TN) is a constitutive fact about belief, then to believe that (TN) is neither true nor false in turn commits one to believing the following instance of (TN):

- (4) It is correct to believe <(TN) is neither true nor false> if and only if <(TN) is neither true nor false> is true.

Hence believing that (TN) is neither true nor false commits one to believing that it is true.

So the outlook seems grim for the expressivist about alethic value. But perhaps all is not lost. So far we've only looked at a toy version of the view. A more sophisticated version will attempt to dodge such complications by invoking a now-familiar two-stance approach. One stance is the stance we take when we *employ* evaluative language. Thus, from what Mark Timmons usefully calls the *morally engaged standpoint* (1998: 150–1), the moral expressivist can affirm all that the realist can affirm. She can do so, the claim goes, by adopting a minimalist or deflationary theory of truth (see Timmons 1998; Blackburn 1998).

Deflationism about truth comes in many forms. But the standard versions of the view tend to endorse the following three points.¹⁶ First, our grasp of the concept of truth is constituted by our grasp of (the instances of) the T-schema:

- (TS) <p> is true if and only if p.

Second, the concept of truth we so grasp is only an expressive device: its sole function is that it allows us to generalize over propositions as in 'Everything Socrates said was true.' Consequently, third, it is a mistake to think that truth has a 'nature'; we needn't appeal to truth to explain anything of philosophical importance.

If minimalism is true, the expressivist contends, then since there is nothing more to saying that <p> is true than, roughly, saying that p, our sentimental commitments just carry over the T-schema. So whatever attitude I express when affirming that slavery is wrong, I can climb 'Ramsey's ladder,' as Blackburn puts it, and express it by saying it is true that it is wrong, or that it is objectively true that it is wrong, or that it is really incredibly objectively true that it is wrong, and so on (see Blackburn 1998: 78–9). At least, all this is so from the morally engaged standpoint.

¹⁶ Representative examples of deflationists include Horwich (1998a); Field (2001); Williams (2001).

The expressivist, however, also employs another stance in regarding morality—what Timmons calls the *morally disengaged standpoint*. This is the stance from which the expressivist wishes to ‘give a story about how ethical thought functions’ (Blackburn 1998: 49). From this vantage point, the expressivist insists that evaluative thought and language look very different than the realist believes. Theorists describe the difference differently, saying variously that evaluative claims don’t aim to represent, or describe, they don’t express beliefs; they don’t correspond to the mind-independent world, or simply ‘they are neither true nor false’ (Timmons 1998: 151). But however it is cashed out, this need for a second stance is clear: it allows the expressivist to speak with the vulgar while still maintaining what is distinctive about her position—namely a form of meta-ethical skepticism or irrealism about value, one which affirms that, for example, ‘ethical properties of things are constructed precisely in order to reflect our concerns’ (1998: 80).

So according to the contemporary expressivist, from the engaged standpoint, slavery is wrong and so is anyone who believes otherwise. But from the disengaged standpoint, from which we are not using but explaining evaluative thought and language, differing moral ends are on a par, factually speaking. As Timmons puts it, there is no moral FACT—no mind-independent, objective fact—of the matter about which moral outlook is correct (Timmons 1998: 152).

Applied to the present case, we presumably arrive at something like the following. The expressivist will insist that, from the *alethically engaged perspective*, she can say whatever the realist can say about the truth-norm. It is correct to believe a proposition when it is true, and if a proposition is true, it is correct to believe. (TN) is therefore not, from this perspective, equivalent in meaning to (TNe). Rather (TN) is just true. But from the *alethically disengaged perspective*, ascriptions of alethic correctness, and hence (TN) itself, are neither true nor false but expressions of our desires and sentiments.

But here the expressivist faces an obvious problem. If the view is to be coherent, it must be possible for us to achieve the alethically disengaged standpoint. In order to even state her view, the expressivist about alethic correctness must be able to abstract from her own sentiments toward true beliefs in order to rise above them. And if we are going to be able to assess the view, let alone agree with it, we must be able to do the same. In the moral case, this is perhaps not too difficult—we don’t have to try too hard to imagine having been brought up with different values. But it is not clear, to say the least, whether I can abstract from my sentiments towards believing.

But hold on, one might think—Isn’t the expressivist simply offering a different sort of *explanation* of our doxastic practice? The expressivist, one might say, following Simon Blackburn, is merely doing a bit of Humean naturalist philosophy, helping us to see what *grounds* our positive evaluations of beliefs and then reminding us that this is all the explanation we need for

those evaluations—no appeal to ‘normative facts’ is necessary. Surely that is a reasonable and coherent enterprise?¹⁷

An explanatory stance of this sort in the case of moral evaluation does seem coherent. But I remain unconvinced in the present case, and the issue again hangs on whether there is conceptual room for the expressivist to make her needed two-stance distinction in the case of alethic value.

Consider again the moral expressivist. In offering her explanation of the grounds of our moral sentiments, she typically insists, with the realist, on the supervenience of the moral on the non-moral. Our moral evaluations are *determined* by natural facts; and as such, they don’t just come out of thin air; they are roped, via our use of our specific moral concepts, to the non-moral facts on the ground. Asked to say why it is that we have the sentiments we do towards, for example, pleasure, the expressivist will point us to the natural facts about pleasure, its causes, and our human reactions to it. These facts explain why we have the sentiments we do, and why we express them in this way rather than not.

The alethic expressivist will wish to say the same in the present case. That is, she will presumably ‘read’ (TN) ‘Socratically’—as telling us that we positively evaluate beliefs in virtue of their being true. In other words, the ‘natural’ fact that a given proposition is true is what *makes it the case* that it is correct to believe—or at least makes it the case that we express a positive sentiment towards believing that proposition. Once again, our sentiments do not come out of thin air. We cheer for doing what is pleasurable because of what pleasure is; we cheer for believing a true proposition rather than a false but flattering one because of what truth is. And this means that, just as in the moral case, we need to say something about what truth is if we wish to explain why our moral sentiments are the way they are.

But now the problem appears: for just as our moral expressivist employs the facts about pleasure to *explain* our positive evaluations of it, so the alethic expressivist invokes truth to *explain* our positive evaluations of beliefs. This is no problem just so long as the expressivist has a theory of truth that allows for this. But the contemporary expressivist we’ve been imagining is a deflationist; and this is just the sort of explanation the deflationist is presumably barred from using.

The point here *isn’t* that a deflationist can’t agree that TN is true. She’ll insist that she can, just as she’ll insist that she can accept any generalization involving truth that we might happen to mention. In this case, she’ll first point out that she can obviously accept every instance of the schema:

It is correct to believe <p> if and only if p.

And then claim that ‘true’ is employed merely as a device for generalizing over such instances, resulting in the ‘equivalent’ (TN). I am not contesting

¹⁷ Broadly speaking, this is how I understand Blackburn’s approach in his (1998).

this (familiar) move here, although I have elsewhere.¹⁸ Rather, my point is that *expressivism* about (TN) is unstable. On the one hand, the expressivist about (TN) needs a deflationary theory of truth if she is going to make sense of biconditionals like (TN). On the other, she must appeal to the nature of truth if she is going to explain why our sentiments are as they are. And it is this that is not allowed on deflationism, for the deflationist believes that truth has no nature, and has no explanatory role to play in philosophical explanation.

The point here is analogous to the familiar point—widely granted by deflationists—that deflationism isn't compatible with a truth-conditional theory of the nature of meaning.¹⁹ The deflationist *can* agree that, when explaining the meaning of any particular sentence, it is perfectly correct (if trivial) to cite its truth condition (e.g. to say that 'snow is white' means that snow is white). What she *can't* agree to is that we must appeal to truth to explain what meaning consists in. For to do that would be to concede that there are facts about the nature of truth—facts about its intimate relation to meaning—that go beyond the concept's expressive function. Similarly, the deflationist might be able to grant TN, but she won't grant the very point the expressivist needs: namely that it is correct to believe a proposition *in virtue of* its having the property of truth. For to say this commits one to holding that there are facts about truth and its explanatory power that go beyond what can be gleaned from TS.

This is perhaps unsurprising. After all, it seems reasonable to think that if (TN) tells us something about belief, then it also tells us something about truth—namely that truth just is, in part, a basic norm of correctness for belief. Truth and belief are clearly interrelated. And so it seems that if (TN) is a constitutive fact about belief, then it is also a constitutive fact about truth. Here Dummett's old analogy of truth and winning is on the mark: the fact that the aim of a game is to win is not just a fact about games; it is also a fact about winning (Dummett 1959). Similarly, the fact that the 'aim' of belief is truth is not just a fact about belief; it is a fact about truth. Of course, the nature and explanation of this fact, like its sister fact about belief, is a matter of dispute—but the conclusion, however understood, seems unavoidable if we read (TN) in the natural, Socratic, way we have here.

I conclude, therefore, that attempting to evade our initial arguments against expressivism about alethic correctness (that it can't explain the conditionals (1) and (2) and that it is self-undermining) by appeal to a two-stance approach towards (TN) fails. The expressivist about (TN) faces a dwindling number of options at this point. One is to deny that (TN) is a constitutive fact about belief. But that requires a non-standard account of belief and its norms. Another is to claim that statements of the form 'x believes that y' are themselves neither true nor false—that is, to embrace an expressivism about belief. This requires a very

¹⁸ See Lynch 2004*b*.

¹⁹ See Field 2001; Horwich 1998*b* for prime examples of deflationists who make this point.

non-standard theory of mind. An initially more plausible approach might be to reject the premise, appealed to just above, that the expressivist needs to read (TN) as

Socratic: It is correct to believe $\langle p \rangle$ in virtue of $\langle p \rangle$ being true.

How else to read (TN)? Well, one possibility is to read it as

Euthyphronic: $\langle p \rangle$ is true in virtue of it being correct to believe $\langle p \rangle$.

Another would be to take it that truth and alethic correctness are simply identical. Earlier we noted that this is implausible if we take correctness to attach to the act of believing and truth to the proposition believed. But we might try avoiding this by construing 'correct to believe $\langle p \rangle$ ' as equivalent to ' $\langle p \rangle$ is correct to believe.' If so, then we might claim:

Identical: $\langle p \rangle$ is true = $\langle p \rangle$ is correct to believe.

Both of these interpretations have their problems; but for present purposes, we need only note that neither will do for the expressivist. If either Euthyphronic or Identical is endorsed together with expressivism about alethic correctness, the latter position collapses. Here's why. Let us grant, for the moment, that the expressivist is entitled to deflationism after all. And thus let us grant that she is entitled to her distinction between the alethically engaged and disengaged perspectives. Thus, on her view, from the alethically disengaged perspective

$\langle p \rangle$ is correct to believe

does not state a fact—or a FACT, even if from the engaged perspective, we are entitled to say that it is true, and thus expresses a fact in the deflationary sense. But now consider: if either Euthyphronic or Identical are true, then it is difficult to see how

$\langle p \rangle$ is true

could fail to be non-FACTUAL as well. Take Euthyphronic, which says that the truth of a proposition is determined by its being correct. If it is non-FACTUAL whether x , and x determines y , then it can't be a FACTUAL matter whether y . Thus if being correct is not a FACTUAL matter, and correctness determines truth, then being true is not a FACTUAL matter either. Even more obviously, if correctness is just identical to truth, as Identical contends, then if correctness is non-FACTUAL, so is truth. Things get worse from here. For given the T-schema, where

$\langle p \rangle$ is true

fails to be FACTUAL, so must p . Which is to say that nothing at all is a FACT. And that seems bad, not only because it is absurd—but because it once again deconstructs the distinction between FACTS and facts that the expressivist requires.

4. EXPRESSIVISM ABOUT TRUTH AS A GOAL OF INQUIRY

So far then, we have found that skepticism about (TN) is difficult to sustain. What about (TG)? Here too, I think that skepticism is not promising. Since that is the subject of a companion piece to this paper, I will only briefly summarize my view on the matter here.²⁰

Let's again take as our target form of skepticism an expressivist view of the value in question. That value is again the value of true beliefs as a goal of inquiry. Call this epistemic expressivism. It amounts to the view that there is no FACT of the matter about whether (TG) is true.

Again, the epistemic expressivist will wish to distinguish two stances: the epistemically engaged and disengaged perspectives. The first perspective is the one she has when she engages in inquiry; the second is the perspective from which she offers her meta-normative explanation of the goals of inquiry. From this latter perspective, to say that truth is a proper end of inquiry is not to say something true but to express a sentiment. This is to agree with Hartry Field when he says,

There are no constraints on what one's epistemological goals ought to be: nothing makes it wrong for a person not to care about achieving truth and avoiding falsehood but adopting beliefs that will make him feel good about his cultural origins. (Field 2001: 385)

From the epistemically disengaged perspective, there is no question of whether one goal of inquiry is better than another.

In my view, the epistemic expressivist, like the alethic expressivist we just discussed, faces a problem right here. For there are serious reasons to think that, whatever we might say in the moral case, the epistemically disengaged standpoint is illusory.

In the contrasting ethical case, the moral expressivist asks us to consider people who have very different moral ends than our own. For it is by considering this possibility that we are able to perform the necessary 'disengagement' from our own moral ends and view them with a critical eye. Similarly, if we are to make sense of the epistemically disengaged standpoint, we need to consider the possibility of not having true belief as an epistemic goal. It isn't sufficient to simply imagine someone who has *more* than true belief among his aims (such as someone who wants justified *and* true beliefs). Rather, one would have to say that someone could engage in inquiry without having true belief among his aims at all. But here we have the problem: for an activity whose aims don't include true belief isn't bad inquiry; it is not inquiry at all.²¹

²⁰ See Lynch (forthcoming).

²¹ And the point can be put in terms of (TN) as well: the expressivist about the truth norm would have to be able to make sense of believers whose beliefs were not correct when true. But

Suppose we encounter someone, for example, who is committed to the goal of 'accepting' all and only that which flatters his cultural origin. What would make us think that in attempting to 'accept' propositions that have this property, he is actually engaging in inquiry? Our reflections on (TN) suggests that, unless his practices are regulated by the property of truth, we can't even understand him as engaged in figuring out what to *believe*, as opposed to engaging in an elaborate game of wishful thinking. And if he isn't engaged in figuring out what to believe, he isn't engaged in inquiry. On the other if he is forming beliefs when accepting his preferred type of proposition, he is implicitly committed to (TN); and what we've called the trivial connection principle tells us that if one is committed to (TN), and one engages in inquiry, one is committed to (TG). In sum, it seems that unlike the moral case, we can't readily conceive of different *epistemic* goals that don't include among them the goal of believing what's true. And that is just to say that we can't meaningfully abstract from our own epistemic goals, which in turn means that the idea of an epistemically disengaged standpoint is empty.

Of course, even if this argument is successful, it doesn't solve what might be a deeper question. For suppose we grant that people who don't care at all about the truth, and are motivated by entirely different goals, aren't engaged in inquiry, but rather schminquiry, and they form schmeliefs rather than beliefs. So what? This just pushes the question back. Why is it better to have beliefs rather than schmeliefs to engage in inquiry rather than schminquiry? Why is it good to be epistemically engaged at all?

Can we answer this? If the question is understood as a demand for justification for *inquiry as a whole*, it seems simply impossible to answer. For it asks us to provide an epistemic *reason*, an *argument*, and a *justification* for the practice of giving reasons, arguments, and justifications. And that obviously can't be done, for those activities are constitutive of inquiry, whose aim is the formation of true beliefs. I cannot give a non-circular justification of my belief that it is valuable to engage in inquiry; for in answering the question I am already committed to the value of the very practice in question. Consequently, I can't say what makes inquiry epistemically better than schminquiry. But the best explanation of this fact is not, surely, that there is no fact of the matter. Rather, it is because one can assess something epistemically only from the epistemically engaged standpoint. To ask for a justification from the epistemically disengaged standpoint is to ask for nothing.

Of course, this is not to say that we must remain mute in the face of a more realistic and hence more formidable skepticism, one which challenges us to tell us why we should care so much about truth, or claims that inquiry and true belief are only good as means to power, or some such. In the face of this more limited

that can't be made sense of, for it is a fundamental fact about belief that beliefs are correct when true.

skepticism about inquiry and truth, we can point to the connections between true belief and inquiry and our other values. We can, in short, demonstrate the constitutive role inquiry and its proper aim play in our lives. To do so would be to show how the values of truth are essentially part of, rather than mere means to, other things that matter, such as integrity, authenticity, and democratic political institutions. I have attempted some remarks to this end elsewhere, and will not repeat them here (2004a). My present point is simply that we are not without resources for giving a less global, if more realistic, explanation for why we should care about truth and inquiry.

Skeptics may respond to these more limited justifications by suggesting that they only give moral, or at any rate non-epistemic, reasons for the value of true belief and inquiry. Does this mean that true belief is not a truly epistemic value after all? Of course not—any more than a successful answer to the question ‘why be moral’ would show that there were no moral values. We cannot give reductive answers to such spade-turning questions; we can only point to the connections between our values.

Thus it appears that expressivism about (TG), like expressivism about (TN), is difficult to make sense of simply because we don’t seem to be able to reach a sufficiently impartial standpoint from which to understand the position.

5. CONCLUSION

Our reflections have led us to something of a curious conclusion. They suggest that, unlike some of our other values, we cannot sufficiently abstract away from either value of truth in order to be sufficiently skeptical about them. It is as if we cannot be objective enough about our commitment to the values of truth.

One reason this is curious is that it falls short of a straightforward defense of realism about the values of truth. By a ‘straightforward’ defense I mean an argument which proves that (TN) and (TG) are objectively true. Our reflections don’t show this. They show that we can’t help but think they are objectively true. Yet the fact that we can’t help but think that either (TN) or (TG) is true threatens to be as much a fact about us as it does about the world.

Thankfully the conclusion has bite either way. It shows that we cannot, even if we wish to, take a skeptical attitude towards all of our values; towards some, certainly, but not towards all, or at least, not towards all at once. In particular, we cannot take a skeptical attitude towards the values of truth. And that means that we can’t help but think that there are some objective values. If this is not realism about value in the strictest sense, perhaps it is realism enough.²²

²² Earlier versions of this paper were read at the Stirling workshop on epistemic value and at the SSPP meeting in Charleston, NC. I thank the audiences on these occasions and Alan Millar, Duncan Pritchard, Adrian Haddock, Chase Wrenn, Carrie Jenkins, Patrick Greenough, Crispin Wright, Tom Bontly, and Paul Bloomfield for helpful discussion and comments.

REFERENCES

- Alston, W. P. (1993). *The Reliability of Sense Perception*. Ithaca, NY: Cornell University Press.
- Ayer, A. J. (1952). *Language, Truth and Logic*. (1st Dover edn.) Mineola, NY: Dover Publications.
- Blackburn, S. (1993). *Essays in Quasi-Realism*. Oxford: Oxford University Press.
- (1998). *Ruling Passions*. Oxford: Clarendon Press.
- Bogohossian, P. (2003). 'The Normativity of Content,' *Philosophical Issues*, 13: 31–45.
- (1990). 'The Status of Content,' *Philosophical Review*, 99: 157–84.
- David, M. (2001). 'Truth as the Epistemic Goal,' in M. Steup (ed.), *Knowledge, Truth and Duty*. New York: Oxford University Press.
- (2005). 'On "Truth is Good",' *Philosophical Books*, 46: 292–302.
- DePaul, M. (2001). 'Value Monism in Epistemology,' in M. Steup (ed.), *Knowledge, Truth and Duty*. New York: Oxford University Press.
- Dreier, J. (1999). 'Transforming Expressivism,' *Nous*, 33: 558–72.
- Dummett, M. (1959). 'Truth,' *Proceedings of the Aristotelian Society*, 59: 141–62.
- Field, H. (2001). *Truth and the Absence of Fact*. Oxford: Oxford University Press.
- Frankfurt, H. (1958). 'Peirce's Account of Inquiry,' *Journal of Philosophy*, 55: 588–92.
- Gibbard, A. (1990). *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press.
- (2003). 'Thoughts and Norms,' *Philosophical Issues*, 13: 83–97.
- Horwich, P. (1998a). *Truth*, 2nd edn. Oxford: Oxford University Press.
- (1998b). *Meaning*. Oxford: Oxford University Press.
- (2002). 'Norms of Truth and Meaning,' in R. Schantz (ed.), *What is Truth?* Berlin: Walter de Gruyter.
- James, W. (1975). *Pragmatism and the Meaning of Truth*. Cambridge MA: Harvard University Press.
- Kvanvig, J. (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge, MA: Cambridge University Press.
- Lynch, M. (ed.) (2001a). *The Nature of Truth*. Cambridge, MA: MIT Press.
- (2001b). 'A Functional Theory of Truth,' in Lynch (2001a), 723–50.
- (2004a). *True to Life*. Cambridge, MA: MIT Press.
- (2004b). 'Minimalism and the Value of Truth,' *Philosophical Quarterly*, 54: 497–517.
- (2004c). 'Truth and Multiple Realizability,' *Australasian Journal of Philosophy*, 82: 384–408.
- (2005a). 'Replies to Critics,' *Philosophical Books*, 46 (2005): 331–42.
- (2005b). 'Functionalism and our Folk Theory of Truth: Reply to Cory Wright,' *Synthese*, 145: 29–43.
- (2006). 'ReWrighting Pluralism,' *The Monist*, 89: 63–85.
- (forthcoming). 'Truth, Value and Epistemic Expressivism,' *Philosophy and Phenomenological Research*.
- McGrath, M. (2005). 'Lynch on the Value of Truth,' *Philosophical Books*, 46: 302–10.
- Papineau, D. (1999). 'Normativity and Judgement,' *Proceedings of the Aristotelian Society*, 72: 17–44.

- Peirce, C. S. (1932–3). *Collected Papers*, ii–iv. Cambridge, MA: Harvard University Press.
- Pillar, C. (forthcoming) ‘Desiring the Truth and Nothing but the Truth,’ *Nous*.
- Shah, N. (2003). ‘How Truth Governs Belief,’ *Philosophical Review*, 112: 447–83.
- and Velleman, J. D. (2005). ‘Doxastic Deliberation,’ *Philosophical Review*, 114(4): 497–534.
- Sosa, E. (1997). ‘Reflective Knowledge in the Best Circles,’ *Journal of Philosophy*, 94: 410–30.
- (2001). ‘For the Love of Truth?,’ in A. Fairweather and L. Zagzebski (eds.), *Virtue Epistemology: Essays on Epistemic Virtue and Responsibility*. New York: Oxford University Press.
- Stevenson, C. (1937). ‘The Emotive Meaning of Ethical Terms,’ *Mind*, 46.
- Timmons, M. (1998). *Morality without Foundations*. Oxford: Oxford University Press.
- Velleman, D. (2000). *The Possibility of Practical Reason*. Oxford: Oxford University Press.
- Wedgewood, R. (2002). ‘The Aim of Belief,’ *Philosophical Perspectives*, 16: 268–97.
- Williams, M. (2001). ‘On Some Critics of Deflationism,’ in R. Schantz (ed.), *What is Truth?* Berlin: Walter de Gruyter.

Epistemic Normativity

Stephen R. Grimm

How should we make sense of our epistemic evaluations? To judge a belief to be *justified* or *rational*, for example, is obviously to think something positive about it, and similarly to judge a belief to be *unjustified* or *irrational* is to think something negative. But what is the source or basis of these judgments?

Among contemporary epistemologists, perhaps the most prominent way to make sense of our epistemic evaluations is in teleological terms.¹ On this way of looking at things, a belief earns positive marks, from an epistemic point of view, just to the extent that it seems to promote or in some way bring about the things with intrinsic epistemic value. And similarly, a belief earns negative marks just to the extent that it seems to *fail* to promote or bring about the things with intrinsic epistemic value. I will say more about the motivation for this view in Section 1, but one of my basic goals in this paper will be to show that the teleological view—at least, as it is popularly understood—is mistaken.² In short, the problem for the view is that our practice of epistemic evaluation is broader and more wide-ranging than the view can capture. After considering a recent proposal by Ernest Sosa that seems to improve on the teleological account, I then suggest that Sosa's proposal too faces significant difficulties. I close by recommending a way of thinking about the nature of our epistemic evaluations that seems to avoid the problems canvassed earlier.

1. THE TELEOLOGICAL ACCOUNT

Although the teleological account is widely popular, in the first part of this paper I will focus on the way in which three philosophers in particular—Alvin

¹ In addition to the figures to be discussed below, see e.g. BonJour (1985: 7–8), Foley (1987: ch. 1), and Lehrer (1990: 112). For a more extensive list, see David (2001: 152).

² The qualification 'as it is popularly understood' will prove to be important later. There I will suggest that if one thinks about the goals that are being promoted (respected, etc.) in a different way, then it could be more plausible to think that we appraise our beliefs in terms of how well they promote these other goals and concerns.

Goldman, William Alston, and Michael Lynch—develop the view. Goldman, Alston, and Lynch are worth considering as a group, because although they eventually differ on the question of the value of true belief, they all begin at least by stressing the following two points. First, that as human beings we often value possessing certain epistemic goods *for their own sake*, and not merely for the sake of whatever further goals—especially, further practical goals—we might happen to have. And second, that the reason *why* we value these goods for their own sake is because of, or due to, our natural curiosity.

In the following passages Goldman, Alston, and Lynch not only endorse both claims, but help to show the way in which they seem to be naturally related:

[Goldman:] Our interest in information has two sources: *curiosity* and practical concerns. The dinosaur extinction fascinates us, although knowing its cause would have no material impact on our lives. We also seek knowledge for practical reasons, as when we solicit a physician's diagnosis or compare prices at automobile dealerships. (Goldman 1999: 3; emphasis added)³

[Alston:] [Although having true beliefs furthers our practical goals] the attainment of knowledge and understanding are *also of intrinsic value*. 'All men by nature desire to know,' said Aristotle, and this dictum has been reaffirmed by many of his successors. Members of our species seem to have a built-in drive to get to *the truth about things that pique their curiosity* and to understand how and why things are as they are and happen as they do. So it is as close to truistic as we can get in philosophy to take truth as a good-making characteristic, and falsity as a bad-making characteristic, of beliefs and other outputs of cognition. (Alston 2005: 31; emphasis added)⁴

[Lynch:] We care about the truth for more than just the benefits it brings us. . . . There are times in our lives when we simply want to know for no other reason than the knowing itself. *Curiosity* is not always motivated by practical concerns. Consider extremely abstract mathematical conjectures. With regard to at least some such conjectures, knowing their truth would get us no closer to anything else we want. (Lynch 2004: 15–16; emphasis added)

What I am calling the 'first' point therefore seems to be the more fundamental one: namely, that certain epistemic goods seem to possess a kind of intrinsic

³ This passage highlights the role of curiosity, but it is not as clear as it might be that Goldman associates this with a true-belief-for-its-own-sake claim. It is thus helpful to read this passage in conjunction with a passage from his earlier *Epistemology and Cognition* (1986). There he writes: 'Even if the desire for truth-acquisition is ultimately traceable to biological fitness (curiosity about one's environment can promote survival), it still appears in the organism as an 'autonomous' desire. People do not desire true belief merely as a means to survival, or the achievement of practical ends. Truth acquisition is often desired for its own sake, not for ulterior ends. It would hardly be surprising, then, that intellectual norms should incorporate true belief as an autonomous value, quite apart from its possible contribution to biological or practical ends' (Goldman 1986: 98).

⁴ Notice that, although Alston begins this passage by suggesting (along with Aristotle) that it is knowledge and understanding that is desired for its own sake, by the end of the passage (and in keeping with the rest of the argument in the book) he claims that it is truth that is the 'good-making characteristic'—in other words, the intrinsically valuable thing.

value. That is, they seem to be goods worth acquiring for their own sake, and not merely for the sake of whatever further practical goods they might help to produce.⁵ What I am calling the 'second' point, concerning the role of *curiosity*, in turn seems to be offered as a kind of explanation or defense of the first. After all, it might be thought, although little needs to be said on behalf of the idea that certain practical goods (such as pleasure, perhaps) are worth realizing for their own sake, it might be less obvious that epistemic goods have the same kind of status. Indeed, it might be thought that in comparison with other goods we obviously value for their own sake, the notion of a purely *epistemic* good seems like little more than a fiction.⁶ The appeal to curiosity, it seems clear, is meant to cut off just these concerns. Just as there are a range of practical goods we naturally desire, so too, the above passages suggest, there are also purely *epistemic* goods—goods the wanting of which can be explained in terms of our natural curiosity—that we naturally want to possess.⁷

But what are these 'purely epistemic' goods, exactly? In the passages quoted above there is not as much consistency as one might expect. Goldman first speaks vaguely of acquiring information and then of gaining knowledge, Alston first of acquiring knowledge or understanding and then more vaguely of something like possessing the truth, and Lynch first of caring for the truth and then of knowing the truth.

Despite this initial diversity, as we will see in a moment the considered view of all three seems to be that *believing the truth* is the thing that possesses intrinsic epistemic value, at least for creatures like us. When we are uncertain about how things stand with respect to certain subjects (Why *did* the dinosaurs die so suddenly, anyway?) our curiosity is naturally piqued by those subjects. Finding out the truth with respect to such subjects—in other words, believing the truth with respect to such subjects—accordingly possesses an intrinsic worth or value all its own.

⁵ Perhaps, as Goldman suggests, the nature of the dinosaur extinction is like this—though one would have thought the practical relevance of this topic (sudden and catastrophic extinction!) was fairly clear.

⁶ According to Stephen Stich (1990: 131), for example, although we do value many things intrinsically—health, happiness, the welfare of our children, etc.—the truth (Stich's main candidate for a putatively epistemic good) is not one of them. Similarly, although Hilary Kornblith (2003: ch. 5) is critical of many aspects of Stich's view, he seems to agree that there are no epistemic goods that are worth pursuing for their own sake.

⁷ In addition to the authors cited above, Ram Neta is another who makes this connection explicit. As he writes: 'Knowledge and other positive epistemic statuses are worthy of pursuit by inquisitive creatures not (or not just) because they are instrumentally valuable. They may, of course, be instrumentally valuable—we need not disagree with Kornblith on that point. But that's not the only thing that makes them worthy of pursuit for inquisitive creatures. What makes them worthy of pursuit for inquisitive creatures like ourselves is that, like health, friendship, and love, their attainment is partly constitutive of our well-being. Knowledge, and epistemic excellence more generally, is part of what constitutes the natural and valuable phenomenon of an inquisitive creature's well-being' (2007: 352).

Once we take believing the truth to be intrinsically valuable from an epistemic point of view, at any rate, for many philosophers the following teleological account of epistemic appraisal has come to seem very natural and compelling:

The teleological account of epistemic appraisal: A belief earns positive marks (counts as justified, rational, virtuous, etc.), from an epistemic point of view, just in case it does well with respect to the things with intrinsic epistemic value (i.e. helps to promote them or bring them about). Likewise, a belief earns negative marks just in case it does poorly with respect to the things with intrinsic epistemic value.⁸

As the following passages suggest, Goldman, Alston, and Lynch all endorse this view in very similar terms, and make it clear (or, at least, clearer) that by their lights true belief is the thing with intrinsic epistemic value.⁹

[Goldman:] I shall attempt to make a case for the unity of epistemic virtues in which the cardinal value, or underlying motif, is something like true, or accurate, belief. . . . The principal relation that epistemic virtues bear to the core epistemic value will be a teleological or consequentialist one. A process, trait, or action is an epistemic virtue to the extent that it tends to produce, generate, or promote (roughly) true belief. (2002: 52)

[Alston:] We evaluate something epistemically (I will be mostly concerned with the evaluation of beliefs) when we judge it to be more or less good from the epistemic point of view, that is, for the attainment of epistemic purposes. . . . The evaluative aspect of epistemology involves an attempt to identify ways in which the conduct and the products of our cognitive activities can be better or worse vis-à-vis the goals of cognition. And what are those goals? Along with many other epistemologists I suggest that the primary function of cognition in human life is to acquire true beliefs rather than false beliefs about matters that are of interest to us. (2005: 28)¹⁰

⁸ This account is doubtless incomplete as it stands (though complete enough for our purposes), because presumably *withholdings* too can be appraised epistemically. For more on this see DePaul (2004).

⁹ I use the notion of 'intrinsic value' in what I take to be the standard way here, to mean a value that is worth pursuing and realizing for its own sake. When Goldman speaks of truth as a 'cardinal value' we pursue for its own sake, I therefore assume by this he means what we standardly mean by an 'intrinsic value.' (It is possible that by the talk of 'cardinal' instead of 'intrinsic' value, however, Goldman has something more like the Sosa view, which we will consider shortly, in mind.) Similarly, Lynch (elsewhere) prefers to speak of the for-its-own-sake value that believing the truth possesses as a 'constitutive value' (see e.g. Lynch 2004: 127)—'constitutive' in the sense that it is an essential constitutive part of a flourishing life, which is an end we all desire. As Lynch notes, the notion of a constitutive value is theoretically quite similar to the notion of an intrinsic value: 'Being constitutively good, like being an intrinsic good, makes something worth caring about for its own sake, as opposed to caring about it for what it leads to' (2004: 128).

¹⁰ Notice that, unlike Goldman (and Lynch, looking ahead), Alston claims not simply that true belief is the intrinsically valuable thing, but 'true beliefs rather than false beliefs about matters that are of interest to us.' This is not a trivial difference, as I will argue at length in Section 3, but for the moment we can put the distinction to one side.

[Lynch:] Once again, the key point is that the value of believing what is justified is parasitic on the value of believing what is true. Having justified beliefs is good because justified beliefs are likely to be true. (2004: 50)

Although more would need to be said here about many aspects of the view,¹¹ I hope that by now the basic idea is clear enough: again, that the reason why we think of an individual belief as good or bad is because of some sort of ‘doing well’ relationship that the belief bears towards the things with intrinsic epistemic value, where the things with intrinsic epistemic value are taken to be true beliefs.

2. TWO VIEWS OF VALUE

We just saw that Goldman, Alston, and Lynch all motivate their views by pointing out that, when we are uncertain about how things stand with respect to certain subjects, our curiosity is naturally piqued by those subjects. Suppose for the moment we grant that some subjects *do* in fact naturally elicit our curiosity: perhaps something like the dinosaur extinction falls into this category. Finding out the truth with respect to these subjects will then seem to be intrinsically worthwhile, from a purely epistemic point of view.

The question we now need to ask, however, and the one that reveals an instability in the teleological view, is whether it is really plausible to think that just *any* subject falls into this category. Suppose I am uncertain about how many motes of dust there are on my desk now, for example, or about the now-defunct phone number of some random person in Bangladesh. Is my curiosity really naturally elicited by *these* subjects? Are these really the kinds of subjects that Aristotle had in mind when (as Alston notes in his earlier passage) he claimed at the outset of the *Metaphysics* that ‘All men by nature desire to know’?

Significantly, this is where opinions begin to divide. According to Lynch,¹² for example, finding out the truth with respect to just *any* subject—even apparently trivial subjects like the number of motes of dust on my desk—possesses genuine, intrinsic epistemic value: it possesses a value worth pursuing for its own sake, from a purely epistemic point of view. Of course, Lynch is quick to acknowledge that the value to be found in trivial subjects of this sort is usually trumped by our other concerns—the value is therefore only *prima facie*, by his lights. But on his view, and had we world enough and time, finding out the truth with respect to any of these topics would indeed be intrinsically worthwhile, from a purely epistemic point of view.¹³

¹¹ For example, what sort of reliability matters? ‘Actual world’ likelihood (where ‘actual’ is a name, rather than an indexical)? ‘Normal world’ likelihood? Something else? This is none too clear, as Goldman’s various stances over the years suggest.

¹² Among others: Kvanvig (2003: 41) and Horwich (2006: 347) also defend this view.

¹³ As Lynch, responding to the sort of natural objection we will next consider, writes: ‘Come on, what about really trivial truths? Surely there are all sorts of true beliefs I could have that are not

Let's call this the *unrestricted* view of the value of true belief, according to which believing the truth with respect to just *any* subject possesses a kind of intrinsic epistemic value.¹⁴ The reason why opinions begin to divide here is that many philosophers have found the unrestricted view of the value of true belief incredible, including (most notably) fellow advocates of the teleological account such as Goldman and Alston.

Goldman offers his own counterexamples that tell against the unrestricted view (1999: 88; 2002: 61). What is the 323rd entry in the Wichita, Kansas, phone directory? Who was placed sixth in the women's breast stroke in the 1976 Summer Olympics? What was the full name of Domenico Scarlatti's maternal grandmother? According to Goldman, since finding out the answer to questions of this sort seems wholly lacking in value—even, it seems, from a 'purely' epistemic point of view—'We can no longer suggest that higher degrees of truth possession are all that count in matters of inquiry' (2002: 61). Instead, on his view, we need to shift to what we might now call a *restricted* or *qualified* view of intrinsic epistemic value, where what matters is not possessing the truth on any topic but rather only on 'topics of interest' (2002: 61; cf. Goldman 1999: 89).

In his (2005) book Alston too argues that an unrestricted view of the value of true belief cannot be maintained.¹⁵ Since the true beliefs that we could gain from activities like memorizing phone books apparently lack intrinsic value, Alston suggests, along with Goldman he concludes that we need to restrict the realm of those things with intrinsic epistemic value to truths concerning 'matters that are of interest or importance to us' (2005: 32).¹⁶

Despite their initial agreement, the advocates of the teleological account we have been considering so far therefore part ways conspicuously when it comes to identifying the thing (or things) with intrinsic epistemic value. Although all begin with the claim that true beliefs are the things with intrinsic epistemic value, in the face of certain obvious objections—especially, what we might think of as the 'trivial truths' objection—Goldman and Alston immediately back off their claim and relativize the intrinsically valuable things to true beliefs on, roughly, 'matters of interest or importance to us.'

even *prima facie* good? Without a doubt, there are all sorts of true beliefs that are not worth having, all things considered. But the fact that I should not bother with those sorts of beliefs doesn't mean that it isn't still *prima facie* good to believe even the most trivial truth' (2004: 55).

¹⁴ Lynch puts his point more formally as follows: 'It is *prima facie* good, for all p (to believe that p if and only if it is true that p)' (2005: 331).

¹⁵ It is worth noting that his 2005 book represents something of a change in his thinking about the nature of epistemic value. In his earlier 'Concepts of Epistemic Justification,' for example, he there characterized the epistemic goal as that of 'maximizing truth and minimizing falsity in a large body of beliefs,' although he explicitly called that a 'rough characterization' (1989a: 83–4).

¹⁶ And, indeed, many epistemologists writing on the topic of our 'epistemic goal' quite naturally relativize the goal to something like 'topics of interest and importance.' See e.g. Haack (1993: 199) and David (2005: 299). The temptation to make this move is obviously very powerful.

But how dramatic is this difference, and what does it have to teach us about the viability of the teleological view? If Goldman is to be believed, the qualification represents only a 'slight' revision to the teleological view. As he writes:

But can't we incorporate the element of interest by a slight revision in our theory? Let us just say that the core epistemic value is a high degree of truth possession *on topics of interest*. Admittedly, this makes the core underlying value a somewhat 'compound' or 'complex' state of affairs. But, arguably, this is enough to preserve the idea of thematic unity, and thereby preserve Unitarianism. (2002: 61)¹⁷

Alston too seems to think that the revision is quite slight; at any rate, he seems to even lack Goldman's misgiving that such a qualification immediately 'makes the core underlying value a somewhat "compound" or "complex" state of affairs.'¹⁸

What I want to argue in the following section, however, is that this difference concerning the nature of the intrinsic epistemic value is dramatic indeed, and that it exposes a fundamental problem at the heart of the teleological view.

3. A DILEMMA

To see why, suppose we take it, along with Goldman and Alston, that not all true beliefs are intrinsically valuable but only true beliefs with respect to subjects of interest or importance to us. This then leads us to the crucial question: How should we make sense of our epistemic appraisals with respect to those beliefs (or, better, those topics) that apparently *lack* intrinsic epistemic value—that is, that are not interesting or important, from a purely epistemic point of view?

If we take the teleological account at its word, such a belief would deserve a positive or negative appraisal only to the extent that it did well with respect to the things with intrinsic epistemic value. But by hypothesis a true belief on such a topic would *lack* any such value.¹⁹ And from this it would seem to follow that a positive or negative appraisal of the belief would simply be out of place.

But now the problem should be clear, for positive and negative appraisals of such beliefs clearly do *not* seem out of place. Suppose that on a lazy whim you decide to scan your desktop for motes of dust. After a bit of distracted counting, you then conclude that the desktop is harboring eighteen motes. Given the sloppiness of your method, however, we can suppose that this answer really amounted to little more than a guess: you might very easily have concluded, for example, that there were rather more motes or rather fewer.

¹⁷ 'Unitarianism' is Goldman's term for the view that the only intrinsic epistemic value—in his words, the 'cardinal' epistemic value—is true belief.

¹⁸ See e.g. Alston's discussion in 2005: 30–3.

¹⁹ Some philosophers even insist (on a slightly different note) that a *false* belief on this topic holds no intrinsic disvalue. See e.g. Kelly (2003: 624–5).

What now should we say about your belief? Even if it turns out to be true, is it justified? Well-formed? Rationally held? I take it that on all counts the answer is No. Given your lack of responsiveness to the truth about the motes, your belief would presumably earn low marks with respect to virtually any type of epistemic appraisal on offer.

The problem for the restricted teleological account²⁰ offered by Goldman and Alston, however, is to explain why this should be so. Recall that according to the teleological account a belief inherits its epistemic status from its 'doing well' relationship to the things with intrinsic epistemic value. But if there is nothing with intrinsic epistemic value to do well with respect to, then it is hard to see where this inherited value or status might come from. From nothing, nothing comes, it would seem.

This then leads us, by way of summary, to the following dilemma for the teleological account of epistemic appraisal. For suppose that, with Goldman and Alston, it is *not* the case that believing the truth with respect to just any subject is intrinsically worthwhile, from an epistemic point of view. If so, then the teleological account seems unable to explain why our beliefs with respect to subjects that lack this kind of value—for short, 'trivial' beliefs—are appropriate candidates for epistemic appraisal. Suppose instead that along with Lynch, we accept an unrestricted view of intrinsic epistemic value. Combined with the teleological account, we can then make sense of the fact that 'trivial' as well as 'non-trivial' beliefs are appropriate candidates for epistemic appraisal, for such beliefs would be derivatively either good or bad to the extent that they promote or respect the things with intrinsic value. But, again, the problem with this account is that an unrestricted view of intrinsic epistemic value is deeply implausible. It hardly seems to be the case that finding out how things stand with respect to just any subject is intrinsically worthwhile, even from a purely epistemic point of view.²¹

In the following section I will consider a recent proposal by Ernest Sosa that seems to allow us to keep the spirit of the teleological account while avoiding the sorts of problems that arise when we try to take a stand on a particular account of intrinsic epistemic value. Before moving on, however, I want to consider one way that Goldman and Alston might try to blunt their particular horn of the dilemma just described.

Recall that Goldman and Alston both qualified their initial views by restricting the scope of intrinsic epistemic value to (roughly) 'matters of interest or importance.' But now it might be thought that the sort of problem cases imagined above—as when we negatively appraise my sloppy counting of motes

²⁰ That is, the problem arises for an account that combines the teleological view of appraisal with a restricted view of intrinsic epistemic value.

²¹ For more on the selectiveness of our sense of curiosity, see Harman (1999: 100) and Grimm (2008).

of dust—are not *really* problems because they are not fairly described. As I described the mote-counting case, for example, I gave the impression that I could care less about how things stood with respect to this subject; instead, I was just looking to pass the time. But it is implausible to suppose that anything like a genuine *belief* could issue from such a process, for the very process of forming a belief seems to require that I care about how things stand with respect to the subject in question.²² Thus when a topic *interests* me enough to trigger a belief—in other words, when I care about it enough—it might be thought that this implies the presence of something *worth* caring about, hence the presence of a value that could be used to ground further epistemic appraisals.

The basic problem with this response, however, is that it loses track of the fact that not just any sort of good was supposed to ground the teleological account, but rather a good that was distinctively ‘epistemic’; in other words, a good that we took to be intrinsically worth realizing from a ‘purely’ epistemic point of view, or (apparently) simply insofar as we were *curious* beings. Assuming the sort of interest described in the desktop case was serious enough to issue in a belief, however, it was nonetheless still at bottom a practical interest—it was an interest that stemmed from my desire to put off my work for a little while longer, or to give my mind a little rest, and so on. It was, presumably, not an interest that derived from a purely (or even partly) curiosity-driven inclination of mine.

The teleological view, at least as developed by philosophers such as Goldman and Alston, therefore has significant problems trying to make sense of how our concepts of epistemic appraisal apply to apparently ‘trivial’ topics. But perhaps the teleological view is better understood in a different way—one that retains the appealing structural features of the view while bypassing the difficulties that surround the question of epistemic value. As we will see in the following section, Sosa’s recent work offers just such an alternative.

4. SOSA’S VIEW

Sosa begins his account by noting that as human beings we are ‘zestfully judgmental’ across a wide range of areas, including art, literature, science, politics, sports, food, wine, and even coffee (2007: 70). In Sosa’s terms, each of these areas of evaluation therefore represents a kind of *domain*—more exactly, each area represents a *critical* domain. Why ‘critical’? Because, Sosa suggests, once we identify the values that are *fundamental* within each domain we can then appraise or assess (hence criticize) the *derivative* value of other items in the

²² See Nishi Shah (2003), and Shah and David Velleman (2005). It is also true that a typical person will have what we might think of as ‘standing concerns’: thus if I hear a loud noise nearby, or see a flash in the distance, I will (as it were) automatically form a belief about these subjects. Thanks to Robert Audi for helping to clarify this point.

domain in terms of how well they promote, bring about, or perhaps in some way duly respect the domain's fundamental values. The fundamental values within a given domain therefore serve as the goal around which the critical domain is structured.

Consider, for example, the domain of assassinship. Because for an assassin killing one's target is the goal around which the practice of assassinship is structured (or so it seems), we can therefore evaluate various elements of the assassin's conduct in terms of how effectively the conduct realizes this goal. Or consider the critical domain associated with the card game whist. Since the goal of whist is to take the majority of tricks, particular moves in whist can therefore be evaluated in terms of how well they realize *this* fundamental goal.

Moreover, what focusing on these more unusual sorts of domains helps to bring out, Sosa suggests,²³ is that our ability to evaluate particular items within a domain does not turn on our judgments about the worth of the fundamental values that structure the domain. Thus with respect to some domains, such as whist playing, we might think that the fundamental values involved are too trivial to possess any intrinsic worth. And with respect to others, such as assassinship, we might even think that the fundamental values that structure the domain possess positive *dis*value. Nonetheless, as Sosa notes, this hardly seems to get in the way of our ability to appraise particular elements within the domain:

Paradoxically, one can be an adept critic within such a domain even while discerning no domain-transcendent value in it. Thus, someone knowledgeable about guns and their use for hunting, for military ends, etc., may undergo a conversion that makes the use of guns abhorrent. The good shot is thus drained of any real value that he can discern. Nevertheless, his critical judgment within that domain may outstrip anyone else's, whether gun lover or not. Critical domains can be viewed as thus *insulated*, in ways suggested by our example. (2007: 73–4)

The basic insight Sosa wants to build on, then, is that we can evaluate items within a domain in terms of how effectively they promote or bring about the fundamental values of the domain, while all the while remaining agnostic about whether the domain's fundamental values are valuable or worth pursuing intrinsically.

So, how does this basic insight help to shed light on our concepts of epistemic appraisal? According to Sosa, epistemic appraisals too take place within an insulated critical domain, a domain in which the fundamental value is true belief. Unlike Goldman, Alston, and Lynch, however, Sosa argues that in order to make sense of our concepts of epistemic appraisal there is no need to take a stand on whether true belief is something that possesses intrinsic value or is worth pursuing for its own sake. In Sosa's words: 'Truth may or may not be intrinsically

²³ I should add that the examples just mentioned are meant to *illustrate* Sosa's view; they are not Sosa's own.

valuable absolutely, who knows? Our worry requires only that we consider truth the epistemically fundamental value, the ultimate explainer of other distinctively epistemic values' (2007: 72).

5. A CLOSER LOOK

What should we make of Sosa's proposal? On the positive side, the view naturally accommodates our ability to appraise beliefs on *any* topic, even apparently 'trivial' topics. Since the fundamental epistemic value for Sosa is simply true belief, it follows that 'trivial' beliefs can be appraised and evaluated just as readily as more 'important' beliefs. In this way the account captures the full scope of our epistemic appraisals; unlike the Goldman–Alston view, it doesn't leave the trivial out. The view can also make good sense of our appraisals concerning how *effectively* a particular believer reaches the truth goal. Although the terms he uses are a bit Sosa-specific, there clearly seems to be a sense in which we can evaluate a belief as 'adroit' or as 'maladroit' (i.e. as deriving, or not, from a reliable competence on the part of the believer to realize the truth), or as 'apt' or 'inapt' (i.e. as realizing, or not realizing, the truth *because of* such a competence).²⁴ Sosa's view therefore interestingly unites our appraisal of beliefs to our appraisal of performances more generally. Thus, just as we can judge an archer's shot to be adroit or maladroit (relative to the goal of striking the bullseye) or we can judge a tennis player's serve as apt or inapt (relative to the goal of hitting the ball in the appropriate box), so too we can judge the truth-oriented merits of someone's believing: as a performance that manifests various degrees of skill and efficiency relative to the truth goal.

Despite these virtues, what I want to suggest now is that by remaining agnostic about the domain-transcendent value of true belief, Sosa seems to introduce a new problem—seems to, indeed, lose sight of one of the most important aspects of our epistemic appraisals. For notice: when we judge a belief to be unjustified or irrational, we seem to be doing more than just evaluating (in this case, in a negative way) the skill or virtuosity of the believer's performance. In addition, we seem to be in some sense *criticizing*, perhaps even *reproaching*, them for believing in this way.²⁵ To judge someone's belief to be unjustified or irrational is thus to judge that the person's attitude towards the content of the belief *should* be

²⁴ Sosa has been working with this distinction for some time; see e.g. Sosa 1991. For his most recent version, see his 2007: ch. 2.

²⁵ As Nicholas Wolterstorff notes: 'We say to each other such things as, "You should have known better than to think that Borges was an English writer," "You should be more trusting of what our State Department says," and "You should never have believed him when he told you that the auditors had approved that way of keeping books." Not only do we *regret* the knowledge and ignorance of other human beings, their beliefs, disbeliefs, and non-beliefs; we reproach them, blame them, chastise them, using the deontological concepts of ought and ought not, should and should not. Of course we also praise them for believing and not believing, knowing and not knowing, as they do' (2005: 326).

reconsidered, in some apparently binding sense of 'should.'²⁶ As Hilary Kornblith puts the point:

If you tell me that a belief of mine is unjustified, this gives me reason to give up that belief. The epistemic claim is something about which I should care, and an account of the source of epistemic norms must explain why it is that I should care about such things. (2002: 145)

What's more, even what we referred to earlier as the Sosa-specific appraisals such as 'apt' and 'inapt' seem to carry with them this normative force. Thus, and to extend Kornblith's point, if I accept that a certain belief of mine is 'inapt' I seem now to have a reason to do something about my attitude toward the content of the belief: perhaps to change my attitude altogether, or perhaps to try to get into a better epistemic position with respect to the subject at issue, and so on. Simply sticking with the original attitude in the face of the 'inapt' judgment does not seem to be acceptable. If at all possible, it seems that I *should* try to do something about my position.²⁷

Can Sosa make sense of the way in which this binding sense of 'should' attaches to our epistemic evaluations (especially, it seems, our negative evaluations)? In one sense it might be thought that he can, for there does seem to be a natural place for a 'should' even within Sosa's insulated domains of critical appraisal. Thus, we might say that, given that such-and-such is the goal, one *should* proceed in this way—and not in that way—in order to realize the goal. So, for example, given the goal of acquiring a true belief with respect to a given subject, one *should* base one's belief on good evidence, rather than hazard a random guess, because basing one's belief on good evidence is a more effective way of realizing the goal at issue.

²⁶ Notice that in suggesting that beliefs are subject to criticism in this way we do not have to accept that belief is subject to our direct voluntary control. I think (along with Alston 1989*b*, Plantinga 1993, and virtually everyone else) it is obvious that it is not. Instead it seems that all we need suppose is that belief is under *enough* control to make critical judgments appropriate. At any rate, these types of judgments are central enough to our epistemic appraisals that any theory of epistemic normativity should seek to accommodate them (see e.g. Audi 2001, as well as the previous footnote). Notice as well that while I have argued that a judgment that a belief is (say) truth-unjustified carries with it the judgment that the subject of the belief should try to improve her cognitive position with respect to the belief, I do not mean to say that it is always psychologically possible to give up the belief. Sometimes, as a result of brainwashing, perhaps, or possible psychological trauma, it might not be. But it does not follow that the 'unjustified' judgment does not have this binding sense of 'should' attached to it. Consider a comparison with the judgment *cruel*. I take it that when we judge a particular action to be cruel, this brings with it the idea that the agent should not act in this way, even that the agent has a binding reason not to act in this way. But it seems equally clear that we might apply this judgment to the actions of a particular agent even if, for some peculiar psychological reasons, the agent felt compelled to act this way, perhaps to the point where he could not have acted otherwise. Indeed, in a loose, analogical way, we sometimes even apply the judgment 'cruel' to the behavior of animals, even though it seems unlikely that they have the sort of voluntary control over their actions that would make judgments of blame and censure strictly appropriate.

²⁷ Alan Millar (2004: 92–9) and Terence Cuneo (2007: 67–70) likewise tie the notion of normativity to the notion of *having a reason*.

The problem of course is that this sense of 'should' is quite weak; in roughly Kantian terms, it is the 'should' of calculation rather than the stronger, binding 'should' associated with duty or obligation. In this weaker sense of 'should,' after all, the assassin *should* use a high-powered rifle, rather than a flimsy slingshot, in order best to realize his goal of killing the target. The sense of 'should' associated with our judgment that a particular belief is unjustified, however, seems deontologically more substantial than that.

To make better sense of the binding sense of 'should' that seems to attach to our epistemic appraisals, it is worth recalling that we can appraise or evaluate beliefs relative to several different goals.²⁸ Suppose, for example, I am wondering whether God exists. It might be the case that if I were to believe that God did *not* exist then I would experience tremendous psychological distress: I would find it very hard to go on in a world that suddenly seemed devoid of meaning. We can therefore appraise how well my belief about God does not just with respect to the goal of realizing the truth but also with respect to this other goal—roughly, what we might think of as the goal of 'psychological comfort.'

Imagine now that after soberly weighing the evidence I decide that God does not exist, thus (as expected) bringing with it significant psychological distress. Relative to the goal of psychological comfort ('comfort,' for short), we can therefore appraise my belief in a variety of different ways: we can say that it was 'comfort unjustified,' 'comfort irrational,' 'comfort inapt,' 'comfort maladroit,' and so on. Suppose we settle on one of these judgments: that the belief was 'comfort unjustified'—that is, unjustified from the point of view of psychological comfort. If I accept this judgment, does it now follow that I have *reason* to give up my belief, or that I *should* give up my belief? I take it that in some weak, calculative sense of a 'reason' or of 'should' this might be right. Thus, relative to the goal of psychological comfort, I have a reason to give up my belief; alternatively: relative to this end, I should give it up. But it seems clear that neither the 'reason' nor the 'should' at issue here is binding in the way considered above—in neither case does it seem that we can be justly blamed or criticized for failing to orient our belief towards the goal of psychological comfort, for example.

Once we consider things from the truth perspective, however, we can see the normative force of our evaluations has a dramatically different character. Suppose, for example, that instead of soberly weighing up the evidence I formed my belief about God by hazarding a random guess. Relative to the goal of believing the truth, naturally, this belief will earn a variety of negative appraisals: thus we might say (extending our artificial evaluations for the moment) that the belief was 'truth unjustified,' 'truth irrational,' 'truth inapt,' 'truth maladroit,' and so on. Notice now, however, that if I were to agree that my belief was

²⁸ I do not want to suggest that this point is unfamiliar to Sosa (indeed, he begins many of his papers with the very distinction)—just that its force needs to be appreciated properly here.

truth-unjustified, for example, I would now have more than a calculative reason to give up my belief, a reason that would be potentially dispensable, if for some reason I no longer cared about the truth. Instead, I would seem to have a *binding* reason. In accepting this judgment, I would agree that I should not be holding this belief, in some non-optional sense of 'should.'

It seems clear enough, then, that even though we can evaluate beliefs relative to countless different 'fundamental values' and hence countless different domains, the end of realizing the truth enjoys a special sort of status when it comes to the evaluation of belief. What's more, the fact that realizing the truth enjoys this special status seems to account for the particular normative force that our epistemic appraisals possess. Given Sosa's agnosticism about the domain-transcendent value of true belief, however, it is not clear that he can make sense of the fact that the truth perspective is in some sense the privileged perspective—as we might say, the binding perspective—when it comes to the evaluation of belief.

Before moving on, I should note that, towards the end of his most recent discussion of epistemic normativity, Sosa offers a distinction that might seem to accommodate the sort of bindingness we've just been emphasizing. Thus Sosa suggests that we need to distinguish between two sorts of epistemic normativity: on the one hand, the sort of normativity that is *constitutive* of knowledge, and on the other hand, the sort of normativity that is relevant to the study of 'intellectual ethics' (2007: 89–91). On this way of looking at things, the normativity associated with 'intellectual ethics' has to do with appreciating and pursuing the sorts of intellectual topics that are in some sense 'finer' (2007: 89) and hence more *worthy* of our attention. But according to Sosa this sort of normativity—which would seem to bring with it a kind of binding character—has little if anything to do with the sort of normativity that is *constitutive* of knowledge. Thus while we might blame or criticize someone for spending their life counting blades of grass (say), these judgments are separable from our normative judgments about whether a particular belief counts as apt or inapt, adroit or maladroit, and so on.

Although Sosa seems right that there is an important distinction to be made here, the question we still need to ask is whether it helps to explain the distinctive normative force of our epistemic appraisals. And it seems to me that it does not. As I noted at the outset, in trying to offer an account of epistemic normativity we are presumably looking for an explanation of why our epistemic appraisals have the particular force that they evidently have. I have argued, moreover, that our epistemic appraisals have a particular binding or reason-giving force that other sorts of appraisals lack (say, appraisals concerning comfort-aptness or inaptness). But Sosa's distinction seems to offer no explanation of this last fact. Even if we grant him (as I think we should) that a judgment that S should not be counting blades of grass seems categorical rather than merely hypothetical, we still have no explanation for why a judgment that S's grass belief was maladroit (say) goes hand in hand with a judgment that S should try to improve his cognitive position

with respect to the grass, if possible—where the ‘should’ here again seems to be a categorical one, rather than a merely hypothetical one.²⁹ It seems that we need to look further, then, in search of the source of the special sort of normativity that attaches to our epistemic appraisals.

6. OUR POSITION

Let’s take stock. Recall that Section 3 closed with a dilemma for the teleological account of epistemic appraisal. If we suppose that only truths of interest or importance are intrinsically valuable (from a ‘purely’ epistemic point of view—whatever exactly that comes to), then it looks like we lose our ability to explain how our epistemic appraisals apply to unimportant, ‘trivial’ beliefs. If we suppose instead that *any* truth is intrinsically valuable, then it looks like we’ve reached an absurdity; only someone really desperate, apparently, would think *that*. Our discussion of Sosa’s view, in turn, suggested that attempts to make sense of our epistemic appraisals should not lose sight of the distinctive normative force of these appraisals. Thus to judge a belief to be justified (for example) is not simply to judge that it is skillfully oriented to the truth but rather that it *should* be so oriented, in some binding sense of ‘should’—just as to deem a belief to be unjustified is to judge that it should *not* be so oriented, in some binding sense of ‘should not.’ To suppose that it is only, or even primarily, the skillfulness of the belief that we are appraising when we make positive appraisals of this kind seems to lose sight of the fact that a belief can be skillfully aimed at almost any goal. There thus seems to be something special about the truth goal that Sosa’s truth agnosticism apparently leaves out.

Overall, this leaves us with two main points in need of reconciliation. First, the fact that our epistemic appraisals not only can be, but patently are, applied to the full range of our beliefs. Second, that our epistemic appraisals seem to have a distinctive normative (binding) force, suggesting that the truth goal is not simply one goal among others when it comes to the evaluation of belief, but rather a goal with a special status.

What I want to propose in this section is that the best way to make sense of these two claims is to accept a *modified* version of the thesis that any true belief has a special value or worth—a version that attempts to explain the unrestricted value of true belief in a slightly different way, or (better) from a slightly different perspective. More exactly, I want to argue that the best way to make sense of the two claims is by shifting away from the standard first-person question about the value of true belief—wherein we ask (as Goldman, Alston, and Lynch

²⁹ It is worth noting that Sosa’s position, as far as I can see, does not exclude the possibility of identifying a further source for the bindingness that characterizes our epistemic appraisals. The problem is only that the view as it stands leaves this further sort of normativity unexplained.

asked) about the value of true belief in terms of our own intellectual goals or well-being—and by moving instead towards a more communal or social view of the value of truth.³⁰

To appreciate how the switch from a first-person perspective to a social perspective might help, consider first the following passage from Thomas Kelly. Here Kelly makes a point that will by now be familiar: namely, that it hardly seems to be the case that believing the truth with respect to just any subject is worthwhile or valuable, even from a ‘purely epistemic’ point of view. More radically, Kelly also insists that even believing something *false* about many subjects does not bring with it any obvious disvalue. As he writes:

In addition to those many truths such that my believing them would contribute to the achievement of some goal that I have, there are also (countless) truths such that my believing them would not contribute to any goal that I actually have. Whether Bertrand Russell was right- or left-handed, whether Hubert Humphrey was an only child—these are matters of complete indifference to me. That is, I have no preference for having true beliefs to having no beliefs about these subjects; nor, for that matter, do I have any preference for having true beliefs to false beliefs. There is simply no goal—cognitive or otherwise—which I actually have, which would be better achieved in virtue of my believing true propositions about such subjects, or which would be worse achieved in virtue of my believing false propositions about them. (2003: 624–5)

Let’s grant for the moment that Kelly is right about this: that believing the truth with respect to such topics would not contribute to any personal goals he might happen to have. What’s it to him, then, if he forms a belief about whether (say) Humphrey was an only child by means of a random guess? And yet, as we saw before, such a random guess would not only earn low marks from an epistemic point of view (count as unjustified, irrational, etc.), but also earn our criticism and perhaps even reproach. Why so?

Considered from a social point of view, it would seem that the answer to this question can be found by noting that even though we might not care less about some belief (or better, some topic), it is nonetheless the case that other people might care about the topic a great deal. For example, while finding out the truth with respect to whether Humphrey was an only child may not hold any value for us, or may not elicit our curiosity in any way, presumably for Humphrey’s biographer (say) getting this right will be quite important—if not in itself, or for its own sake, then at least for the sake of producing an accurate account of Humphrey’s life. And given that someone in the biographer’s shoes might depend on us as potential sources of information about this topic, it seems

³⁰ It might sound odd to accuse Goldman, at least, of being insensitive to the social value of true belief, for perhaps more than any current epistemologist he has emphasized the importance of the social dimension of knowing! To my mind, however, his teleological framework of epistemic appraisal does not sufficiently reflect this fact. Thus, for example, his basic framework is essentially indistinguishable from Alston’s, who does *not* stress the social in the same way.

that we have an obligation not to be cavalier when we form beliefs about the question—in other words, an obligation to try to position ourselves well with respect to this question, even though, given our own cognitive goals, we might very well be thoroughly uninterested.³¹

In short, what turning away from one's personal goals and concerns and towards our broader role in the information economy (as it were) helps to remind us is that the concerns of others—especially the practical concerns of others—are remarkably plastic and unpredictable. As such, it reminds us that, even though a topic may hold no interest or value from our first-person point of view, it may well hold interest or value for someone else. And as a potential source of information for others, we have an obligation to treat *any* topic or any question with due respect.³²

Suppose this approach works with respect to the Humphrey question. But what about the *really* trivial topics, like the 323rd entry in the Wichita, Kansas, phone directory? Or the number of motes of dust on my desk at this very moment? Although it is harder to imagine how someone might have an interest in these questions, once again we need only bear in mind how odd and varied people's practical concerns can be. For the Wichita phonebook fact-checker, for example, it might well be important—not epistemically important, but practically, presumably—to know the name of the 323rd person in the directory. Less realistically, but still possibly, we can imagine that someone with a particular antipathy towards dust might well wonder whether his new 'anti-dust' strategies have really succeeded in cutting down the number of motes of dust on his desk, as he fondly hopes.

It therefore seems that the basic idea we need to make sense both (a) of the apparently unlimited range of our epistemic appraisals as well as (b) of their normative force is that, given our nature as information-dependent and information-sharing creatures, we have an obligation not just to be sources of information for others but to be *good* sources of information. This obligation stems, moreover, not from the fact that believing the truth with respect to just any subject is intrinsically valuable, but rather because any subject might *come* to have value—if only value of a practical sort—in light of the varied and unpredictable concerns others might have.³³

Plausibly, then, the best way to make sense of the value of true belief is to think of it along the lines of a *common good*. Consider, for example, the value

³¹ There is a difficult question here about whether we have positive obligations to (say) seek out information concerning subjects about which we are currently ignorant. It seems to me that the answer is sometimes yes, sometimes no, though I will not take a stand on that question here.

³² My colleague Allan Hazlett reminds me that this sort of view, right down to the concern with apparently 'trivial' truths, can be found in W. C. K. Clifford's classic essay 'The Ethics of Belief.' For these reasons I am tempted to call the view of normativity developed here Cliffordian, at least in the sense that it suggests that epistemic normativity is ultimately grounded in moral normativity.

³³ Cf. Craig (1990), Greco (forthcoming), Weinberg (2006), Kusch (2009).

we associate with other classic examples of common goods such as *clean water*.³⁴ Given the central place of clean water in all of our lives, there is strong temptation to regard the value of clean water as intrinsic, as inherently worth possessing or acquiring, or at least as inherently worthy of our respect. From a first-person point of view, however, this does not always seem to be the case. Suppose, for example, I have my own guaranteed lifetime supply of water (perhaps a valet carries it around in jugs behind me, wherever I happen to go), more than enough to satisfy whatever thirst I might happen to have. Would any particular parcel of clean water still seem intrinsically valuable in this case—worth acquiring or possessing? It doesn't seem so, at least from my first-person point of view. And yet it looks like the nature of an intrinsic good that its goodness is a necessary feature of the thing, not the kind of property that can come and go.

That said, to appreciate the sense in which any particular parcel of water would nonetheless be worthy of our respect, suppose my imaginary valet and I are crossing a bridge over a wide expanse of clean water. What should my attitude towards this water be? By hypothesis, I am not interested in taking a drink from it; I have my jugs, after all. But while drinking this water is no goal of mine—while it is something I can find no personal value in—it seems clear enough that this stretch of water possesses a value that is worthy of my respect. Were I to dump a barrellful of sludge in the water, for example, this would clearly be something for which I would deserve blame or censure. But why's that, exactly?

Some might appeal at this point to the intrinsic value of preserving natural systems, which I would here be damaging.³⁵ But another, more straightforward answer should again strike us as plausible: namely, that other people might well need this water to satisfy their needs. Indeed, even if I have some reason to think that no one would really be harmed by the loss of this particular parcel of water, that would not seem to justify the dumping. For, given the unpredictable nature of the needs of others, and given how contamination of this sort can spread in unpredictable ways, others very well might turn out to depend on this water. And since clean water plays such an indispensable role in human well-being, we plausibly have an obligation not to pollute in this way, but rather to treat the water with due respect.

The comparison between the value of true belief and clean water is not perfect, but I think it nonetheless focuses our attention in the right way. It is not perfect, because while we could potentially flourish without clean water of any kind (perhaps we could flourish on Twin Earth, for example, with *twater*, not water), it hardly seems to be the case that we could flourish without truth of any kind.

³⁴ Goods such as *clean air* also come to mind, but focusing on water should be enough to make the point. Kusch (Ch. 3 this volume) also interestingly compares the value of true belief to the value of clean water.

³⁵ e.g. Aldo Leopold (1966: 240–2).

The comparison focuses our attention in the right way, however, because it helps us to see the way in which both belong to the category of common goods, goods that are, at least contingently, crucial to human well-being. Even though particular instances of these goods might thus not intrinsically contribute to my well-being, they should nonetheless be duly respected because of the central role that they might play in the lives of others, and perhaps even (who's to say?) in our own.

7. CONCLUSION

We can now offer a few tentative conclusions.

The first conclusion, and one that we can draw from our discussion of Goldman and Alston in particular, is that any attempt to relativize the teleological account of epistemic appraisal to 'questions of interest and importance' (or the like) faces multiple problems. For one thing, there is the problem of offering an account of what it is that makes some questions important and others unimportant, from a purely epistemic (as opposed to practical) point of view. For another thing there is the problem of how to make sense of our epistemic appraisals with respect to those questions that lie on the 'unimportant' side of the ledger (assuming such a side exists). Although in this paper I focused on the second sort of criticism, to my mind the significance of the first problem has yet to be fully appreciated, and will almost certainly prove to be the more important (and difficult) issue going forward. For example: supposing that it is true, *why* exactly is it true that counting motes of dust counts as trivial and lacking in value, from a 'purely epistemic' point of view, while finding out (say) whether one has a hand does not? Or if the hand question too counts as trivial then when, exactly, *does* one come to a question that counts as epistemically important or significant? It is hard to know how even to begin to answer these questions. And yet one often hears appeals to the epistemically 'significant' or 'important' as if these notions were well understood, or could be put to good theoretical use.

The second conclusion is that, if we follow Sosa in adopting a kind of agnosticism about the value of true belief, then we seem at a loss to explain the distinctively normative—that is, binding or action-guiding—character of our epistemic appraisals. Thus, given certain arbitrarily specified ends, we can evaluate how well someone does with respect to those ends—just in the way that Sosa describes. But it doesn't follow that a judgment that someone has failed to do well with respect to those ends is binding or action guiding in the way that a judgment that someone's belief has done poorly with respect to the truth is binding or action guiding. In order to get at the distinctively normative (as opposed to merely evaluative) character of our epistemic appraisals, we need to dig deeper.

We were then left with the following question: how *can* we make sense of the distinctively normative force of our epistemic appraisals? And according

to the proposal I sketched in the final part of the paper the best way to do this is to appreciate true belief's status as a common good (rather than, for example, as an intrinsic good, one the having of which always adds value to the life of the possessor). The way I developed this idea, moreover, was by emphasizing the fact that, as information-dependent and information-sharing creatures, we naturally—and, it seems, rightfully—depend on others as sources of information.

I will close by suggesting that if this final proposal is on the right track, moreover, then it is a mistake to suppose that the sort of normativity that characterizes our epistemic appraisals is basic or irreducible. Instead, epistemic normativity would seem to be explicable in terms of a deeper, and more obviously moral, sort of normativity: namely, the sort of normativity that derives from our obligation to help others carry out their projects and concerns (broadly understood). Although in one way this suggestion does not make the normative force of our epistemic appraisals any less mysterious—for what is the source of our non-hypothetical obligation to assist others, after all?—it does suggest that it is a mistake to try to offer an account of epistemic normativity in isolation, and without attending to what, if anything, we owe to one another from a moral point of view.³⁶

REFERENCES

- Alston, W. (1989a). 'Concepts of Epistemic Justification,' in his *Epistemic Justification*. Ithaca, NY: Cornell University Press.
- (1989b). 'The Deontological Concept of Epistemic Justification,' in his *Epistemic Justification*. Ithaca, NY: Cornell University Press.
- (2005). *Beyond Justification: Dimensions of Epistemic Evaluation*. Ithaca, NY: Cornell University Press.
- Audi, R. (2001). 'Doxastic Voluntarism and the Ethics of Belief,' in M. Steup (ed.), *Knowledge, Truth, and Duty*. New York: Oxford University Press.
- BonJour, L. (1985). *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.
- Clifford, W. C. K. (1877 [1999]). *The Ethics of Belief and Other Essays*. Prometheus Books, Amherst, NY: Harvard University Press.
- Craig, E. (1990). *Knowledge and the State of Nature*. Oxford: Clarendon Press.
- Cuneo, T. (2007). *The Normative Web: An Argument for Moral Realism*. New York: Oxford University Press.
- David, M. (2001). 'Truth as an Epistemic Goal,' in M. Steup (ed.), *Knowledge, Truth, and Duty*. New York: Oxford University Press.

³⁶ Thanks to Robert Audi, Heather Battaly, Jason Baehr, Michael DePaul, John Greco, Adrian Haddock, Allan Hazlett, Martin Kusch, Alan Millar, Duncan Pritchard, Wayne Riggs, Ernest Sosa, and Fritz Warfield for helpful comments on earlier versions of this paper.

- (2005). 'Truth as the Primary Epistemic Goal,' in E. Sosa and M. Steup (eds.), *Contemporary Debates in Epistemology*. Oxford: Blackwell Publishing.
- DePaul, M. (2004). 'Truth Consequentialism, Withholding, and Proportioning Belief to the Evidence,' *Philosophical Issues*, 14: 91–112.
- Foley, R. (1987). *The Theory of Epistemic Rationality*. Cambridge, MA: Harvard University Press.
- Goldman, A. (1986). *Epistemology and Cognition*. Cambridge, MA: Harvard University Press.
- (1999). *Knowledge in a Social World*. New York: Oxford University Press.
- (2002). 'The Unity of the Epistemic Virtues,' in his *Pathways to Knowledge*. New York: Oxford University Press.
- Greco, J. (forthcoming). *Achieving Knowledge*. New York: Cambridge University Press. *Epistemic Evaluation: A Virtue-Theoretic Approach*.
- Grimm, S. (2008). 'Epistemic Goals and Epistemic Values,' *Philosophy and Phenomenological Research*, 77.
- Haack, S. (1993). *Evidence and Inquiry*. Oxford: Blackwell Publishing.
- Harman, G. (1999). 'Pragmatism and Reasons for Belief,' in his *Reasoning, Meaning, and Mind*. Oxford: Clarendon Press.
- Horwich, P. (2006). 'The Value of Truth,' *Noûs*, 40: 347–60.
- Kelly, T. (2003). 'Epistemic Rationality as Instrumental Rationality: A Critique,' *Philosophy and Phenomenological Research*, 66: 612–40.
- Kornblith, H. (2002). *Knowledge and its Place in Nature*. New York: Oxford University Press.
- Kusch, M. (2009 this volume). 'Testimony and the Value of Knowledge'
- Kvanvig, J. (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge, MA: Cambridge University Press.
- Lehrer, K. (1990). *Theory of Knowledge*. Boulder, CO: Westview Press.
- Leopold, A. (1966). 'The Land Ethic,' in his *A Sand County Almanac*. New York: Oxford University Press.
- Lynch, M. (2004). *True to Life: Why Truth Matters*. Cambridge, MA: MIT Press.
- (2005). 'Replies to Critics,' *Philosophical Books*, 46: 331–42.
- Millar, A. (2004). *Understanding People: Normativity and Rationalizing Explanation*. New York: Oxford University Press.
- Neta, R. (2007). 'How to Naturalize Epistemology,' in D. Pritchard and V. Hendricks (eds.), *New Waves in Epistemology*. London: Palgrave Macmillan.
- Plantinga, A. (1993). *Warrant: The Current Debate*. New York: Oxford University Press.
- Quine, W. v. O. (1986). 'Reply to Morton White,' in *The Philosophy of W. V. Quine*, ed. L. Hahn and P. Schlipp. La Salle, IL: Open Court Publishing Company.
- Shah, N. (2003). 'How Truth Governs Belief,' *Philosophical Review*, 112(4): 447–82.
- and Velleman, J. D. (2005). 'Doxastic Deliberation,' *Philosophical Review*, 114(4): 497–534.
- Sosa, E. (1991). 'Intellectual Virtue in Perspective,' in his *Knowledge in Perspective: Selected Essays in Epistemology*. New York: Cambridge University Press.
- (2003). 'The Place of Truth in Epistemology,' in M. DePaul and L. Zagzebski (eds.), *Intellectual Virtue: Perspectives from Ethics and Epistemology*. New York: Oxford University Press.

- Sosa, E. (2007). *A Virtue Epistemology: Apt Belief and Reflective Knowledge*. New York: Oxford University Press.
- Stich, S. (1990). *The Fragmentation of Reason*. Cambridge, MA: MIT Press.
- Weinberg, J. (2006). 'What's Epistemology For? The Case for Neopragmatism in Normative Metaepistemology,' in S. Hetherington (ed.), *Epistemology Futures*. New York: Oxford University Press.
- Wolterstorff, Nicholas (2005). 'Obligation, Entitlement, and Rationality,' in M. Steup and E. Sosa (eds.), *Contemporary Debates in Epistemology*. Oxford: Blackwell Publishing.

12

Curiosity and the Value of Truth

Michael S. Brady

It is uncontroversial to suppose that the truth—or more correctly, true belief—is often instrumentally valuable. But many philosophers also want to maintain that true belief can be intrinsically valuable, or valuable for its own sake.¹ This raises the question of whether we can *show* that truth has this kind of value.² In a number of recent papers, an interesting approach to this question has emerged. This approach posits a link between the value of truth and the answering of our inquiries. On this line, the value of truth is conditional upon our inquiries: that is, truth is valuable *because* (and only because) it answers our questions.³ This does not imply that the value of truth is always instrumental, however. For we sometimes want to know an answer simply for the sake of knowing that answer—in these cases we have an ‘intellectual’ interest in the truth, which is grounded in our natural curiosity—and this suggests that the truth is valuable for its own sake when it answers such an interest. The idea is, therefore, that truth is valuable for its own sake because it answers intellectual inquiries grounded in our natural curiosity.

In this paper, I’ll argue that this approach fails to provide an adequate explanation of the intrinsic or final value of truth. After seeing how the

¹ As Christine Korsgaard (1983) has made clear, it is a mistake to simply identify something’s being valuable for its own sake (or finally valuable) with its being intrinsically valuable. This is because there are cases of final value where the value depends upon something external to the object, and as a result fails to count as intrinsic value—at least on a strict understanding of ‘intrinsic value’. On a strict understanding, intrinsic value ‘depends solely on the nature of the thing in question’ G. E. Moore (1922: 260). There are, however, looser ways to understand the notion of intrinsic value. As Thomas Hurka (2001: 6) puts it, ‘a looser view equates intrinsic goodness just with non-instrumental goodness, or with that portion of the overall goodness of the world that is located in or assignable to a particular state. It is the state’s own goodness, whatever its basis, rather than some other’s.’ In what follows, I’ll employ this looser understanding. Nothing, apart from ease of exposition, will rest upon this usage.

² It is, of course, difficult to provide arguments or justifications for claims about intrinsic value. Indeed, as Jonathan Kvanvig suggests, the fact that ‘we are hard-pressed to say anything informative at all’ in favour of the value of some object or state is evidence that we value that object or state intrinsically. See Kvanvig (forthcoming, 2009). See also Lynch (2004: 127).

³ A different way of making the same point is to claim that truth is the *goal* of inquiry, and that truth is valuable because it constitutes the achievement of our goal.

explanation is supposed to work, I argue that proponents of this view face the difficulty of closing a gap between our valuing the truth on some issue, and the truth on that issue being valuable. I then consider a number of ways in which this gap might be closed, and show why none of them is ultimately plausible. This indicates, moreover, that the approach gets things the wrong way around: it's not that truth is valuable because it satisfies our curiosity; rather, satisfying our curiosity is valuable because it promises to bring about the truth. The fact that we are naturally curious about some subject or question can be a reliable indicator that the truth on that subject is valuable; but it is not a condition of that value.

1

The idea that truth is valuable because it answers our questions is implicit in the work of a number of philosophers. For instance, William Alston, Marian David, Richard Foley, Alvin Goldman, Carl Hempel, Jonathan Kvanvig, Michael Lynch and Ernest Sosa all suggest this approach to epistemic value,⁴ insofar as they maintain that there is a connection between our inquiries or questions and our *regarding* the truth as intrinsically valuable.⁵ The idea that truth *is* intrinsically valuable, or that it is a proper end of our inquiries, seems to follow naturally from this connection.⁶ In this section I want to consider how this explanation of the value of truth is supposed to work.

The first stage of the explanation is to note the connections between inquiry and the value of truth. When we inquire about something, we ask and attempt to answer questions about that thing; in doing so we attempt to find something out about the object of our inquiry.⁷ This indicates a clear connection between inquiry and truth, assuming that a successful answer to our questions constitutes the truth on some issue. This notion of successful inquiry is typically expressed

⁴ See Alston (2005), David (2005), Goldman (1986; 1999), Foley (1987), Hempel (1965), Kvanvig (2003; forthcoming, 2009), Lynch (2004), Sosa (2003).

⁵ The idea that there is a distinctive approach to the value of truth to be found in the writings of these philosophers is suggested by Stephen Grimm's excellent paper 'Epistemic Goals and Epistemic Values' (2008). I have learnt a lot from this paper and share Grimm's scepticism about the approach. However, in this paper I develop arguments against the approach that differ from those raised by Grimm. In particular, I focus on the difficulty in moving from claims about our valuing the truth for its own sake to claims about the intrinsic or final value of truth.

⁶ Alvin Goldman (1986: 98) is explicit about this. He writes: '[t]ruth acquisition is often desired and enjoyed *for its own sake*, not for ulterior ends. It would hardly be surprising, then, that intellectual norms should incorporate true belief as an autonomous value, quite apart from its contribution to biological and practical ends.' But even if this is unsurprising, we'll see that we cannot simply move from the fact that people have the truth as an end to the claim that the truth is a *proper end*.

⁷ Christopher Hookway (1996: 7) writes that 'When we conduct an inquiry, or deliberate on some matter, we attempt to formulate questions and to answer them correctly.'

by saying that truth is the *goal* or *aim* of inquiry. A second connection here is between inquiry and a desire for or an interest in the truth: to inquire about *x* is to want the truth about *x*, or to be interested in the truth about *x*.⁸ So an interest in truth in these cases is an interest in arriving at correct answers to our questions. A third connection is between inquiry and the value of truth—or at least between inquiry and our *valuing* the truth. For the idea that truth is the aim of inquiry suggests that truth has a certain status: the notion of a goal or aim is the notion of something valuable or desirable, something that a person has reason to pursue. It is not surprising, then, that an inquirer will regard success in her inquiry—and hence truth—as something that is valuable. So if I inquire about *x*, this implies that I value the truth about *x*. We might, however, go further than this and suppose that our interest in truth is *exhausted by* our interest in answering our questions. We might suppose, that is, that it is *only* because we care about answering our questions that we care about the truth. This, at least, is suggested by Ernest Sosa, who writes that ‘our interest in the truth is an interest in certain questions or in certain sorts of questions’, and who claims that, in the absence of antecedent interests, one’s attaining the truth on some subject would lack value.⁹ If so, we have an interest in and value the truth because—and only because—we have an interest in and value answering our questions.

The second stage in the explanation is to distinguish different motivations that lie behind our interest in the truth. To do this, we might ask what generates an interest in answering our questions.¹⁰ Clearly, many of our inquiries are responses to our *practical* interests and concerns. When we inquire about something, most of the time we are interested in how it has a bearing on other things that we want or need. For instance, we inquire about cinema times because we have an interest in seeing a particular film; or we want the truth about what happened in the conservatory because we want to solve the murder. We are therefore interested in getting answers to our questions because answering our questions promises to contribute to these other goals. As a result, the truth is often regarded as *instrumentally* valuable, or valuable as a *means* to something else.

However, it is also clear that we are sometimes interested in certain questions or motivated to inquire about certain issues, but not because the truth about such things is a means to some further end that we have. Instead, we sometimes engage in what Jonathan Kvanvig calls ‘inquiry for its own sake’, or pursue what Stephen Grimm terms ‘a purely epistemic or intellectual interest in finding the

⁸ Ernest Sosa (2003: 158) writes: ‘We may want true beliefs, in this sense: that *if*, for whatever reason, we are interested in a certain question, we would prefer to believe a correct rather than an incorrect answer to that question.’

⁹ Sosa (2003: 156) writes, of a ‘trivial’ truth about the number of grains of sand in a random handful, that ‘absent any such antecedent interest, it is hard to see any sort of *value* in one’s having the truth’ on this issue.

¹⁰ Since, according to this line at least, we are interested in the truth because we are interested in answering questions, we cannot appeal to an interest in truth as that which motivates our inquiries. As Sosa implies, without an *antecedent* interest, we would have no interest in truth.

truth'.¹¹ Inquiry for its own sake aims at the truth, but not for any ulterior purpose or concern; we simply want to know the answer to a question *for the sake of knowing that answer*.¹² Now whereas the first kind of interest is generated by our practical concerns, an interest in truth for its own sake seems to result from or reflect our *natural curiosity*. Thus, Carl Hempel maintains that inquiry follows on from 'sheer intellectual curiosity, [from our] deep and persistent desire to know and to understand [ourselves] and [our] world' (1965: 333). And Alvin Goldman writes that 'Our interest in information has two sources: curiosity and practical concerns. The dinosaur extinction fascinates us, although knowing its cause would have no material impact on our lives' (1999: 3). Grimm comments:

According to both Hempel and Goldman . . . it seems that the reason why we desire truth for its own sake, and quite apart from our practical goals, can be traced to the fact that we are naturally curious beings. Even when nothing of practical importance seems to ride on finding out how things stand with respect to a certain subject, given our natural curiosity we simply have a natural interest in finding out how they *do* stand. (Grimm 2008: 727).

It is our natural curiosity which sparks an interest in finding out the truth on certain issues for the sake of finding out. This suggests that truths which answer our natural curiosity are regarded by the subject as good in themselves, or valuable as ends. If this is correct, then sometimes we value getting the right answers to our inquiries because we value satisfying our practical interests, whilst at other times we value getting the right answers simply because we are naturally curious. Getting the truth in the former case is seen as instrumentally valuable, whilst we regard the truth in the latter cases as valuable for its own sake.

The third and final stage of this explanation is to connect these claims about valuing the truth with claims about the *value* of truth. That is, the third stage will be to show the bearing of our valuing the truth because we value answering our questions on the issue of whether the truth is *valuable*. It seems plausible to hold that truth is instrumentally valuable insofar as it contributes to bringing about a valuable practical end. So if we are right in thinking that some truth will contribute to our practical goals and concerns, then this truth is instrumentally valuable. The link between valuing some answer instrumentally and the instrumental value of this answer is relatively straightforward, therefore. But what of the value of truth where there is no practical end at stake, and so where there is nothing for the truth to contribute *to*? How might we connect an interest in the truth for its own sake and that truth's being valuable for its own sake, or with it's being a *proper end* to our intellectual inquiries? The idea here seems to be that truths are valuable for their own sakes simply insofar as they answer our intellectual inquiries and thereby *satisfy* our natural

¹¹ Kvanvig (forthcoming, 2009: p. 3 in manuscript); Grimm (2008: 726).

¹² See Lynch (2004: 16).

curiosity. As Grimm notes, in explaining the line taken by Goldman and Hempel:

according to this way of thinking, our curiosity about how things stand in the world is . . . importantly like the thirst we (characteristically, at least) feel when our body is dehydrated. When our body is dehydrated—when we experience thirst—*satisfying* our thirst is naturally thought to possess a kind of intrinsic value: it seems to be a good in its own right, quite apart from whatever further contributions it might make to our well-being. (Grimm 2008: 727).

On this view, then, there seems to be something good about satisfying our natural curiosity which is independent of any contribution that the truth makes to our practical goals or concerns. If we identify this independent value with the value that truth has for its own sake, then true beliefs are valuable for their own sakes when and because they satisfy our natural curiosity.

I think that this approach to the value of truth, although intriguing and deserving of attention, faces serious difficulties. Or so, at least, I'll argue in the following sections.

2

The proposal to be considered is that true beliefs are valuable for their own sakes because (and only because) they answer questions generated by our natural curiosity. Insofar as truths satisfy our curiosity, they have value as ends. But there seems to be a strong reason to be sceptical about this line on epistemic value. For it is a general truth in value theory that, although the fact that I do desire or care about something might incline us to think that that thing is worth desiring or caring about, it does not guarantee that it is.¹³ There is always the possibility that I desire or care about something that I ought not to desire or care about, that is, something that is not *worthy* of my concern. In other words, there is always the possibility that one of my ends or goals is not a *proper* end or goal. If so, we might think that the fact that I desire the truth on a particular subject for its own sake does not guarantee that the truth on that subject is worth desiring, or is valuable as an end.¹⁴ Moreover, there are a number of more particular reasons

¹³ The idea that desires based upon false beliefs or defective understanding are insufficiently normative to ground practical reasons or genuine value is both widely held and eminently plausible. The idea that desires can be criticized as 'unreasonable' when based upon false beliefs is familiar from Hume. It has been developed, with regard to practical reasons and values, by Bernard Williams, Peter Railton, Michael Smith, and many others. I assume, therefore, that there is a gap between what we as a matter of fact desire and what possesses final value.

¹⁴ This is the case even when people in general have some desire or pattern of concern. The fact that very many people are attracted by thoughts of fame and power does not guarantee that these things are worth pursuing; the fact that benevolent concerns are strongly influenced by the proximity of those in need does not show that it is right for one's concerns to have this pattern. So

why we might think that there is a gap between natural curiosity about the truth on some issue, and the truth on that issue having value as an end—as I'll now explain.

(1) It is possible for natural curiosity to reflect a *practical* rather than an *intellectual* concern for the truth. So although I regard the truth on some issue as worth getting for its own sake, my inquiry might in fact be stimulated and directed by an unacknowledged practical goal. It is, furthermore, possible that my interest in the truth is exhausted by my interest in this unacknowledged practical goal, such that I would have *no* interest in the truth on this issue if this practical interest were absent. This suggests that there is a gap between natural curiosity and intrinsic value in this instance, since there is something dubious about the idea that the truth has intrinsic value because it answers to a subject's *practical* concerns. The fact that we might be unaware of the real motive or reason for desiring the truth on some issue therefore casts some doubt upon the idea that the truth is valuable simply in virtue of satisfying our natural curiosity.

(2) It is possible for natural curiosity about some issue to reflect a false belief or a lack of understanding, in which case we might once again doubt that satisfying this curiosity has final value. For instance, suppose that I am a believer in crystal healing, and as a result I am naturally curious—I desire to know, simply for the sake of knowing—which crystals are thought to be most effective for healing which ailments. Since my desire for truth in this case rests upon a false belief in the efficacy of crystal healing, it is plausible to deny that the resulting truths—that crystal healers typically prefer to use quartz because of its shape and colour, that healers maintain these crystals by immersion in salt water with the aim of preventing 'environmental imbalance', and so on—are valuable for their own sakes. It is therefore tempting to deny that truths have final value because they answer our inquiries, in cases where these inquiries are generated by false beliefs or mistaken understanding.¹⁵

(3) It is possible for natural curiosity to give rise to desires or interests that the subject fails to endorse; as a result, it is possible for natural curiosity to generate inquiries that the subject does not value answering.¹⁶ For instance, perhaps my natural curiosity on some issue is *compulsive*: suppose that I feel a strong urge to

we ought to resist any simple move from the claim that X is generally valued for its own sake to the claim that X is valuable for its own sake.

¹⁵ We might also be tempted to deny that truths have final value insofar as they result from morally dubious desires or inclinations. Think of Leontius's morbid fascination in *The Republic*, or a salacious curiosity about the sex lives of one's neighbours. Nevertheless, there is room here to claim that, insofar as they satisfy natural curiosity, such truths have *some* intrinsic value, even though that value is outweighed by the immorality of engaging in such inquiries or satisfying such interests.

¹⁶ If so, we should deny the third purported connection between inquiry and truth stated above, namely that my inquiring about x implies that I value the truth about x.

count the number of steps I take each time I walk to work, or to determine the ratio of blue to red books in the university library, for no other reason than to know the truth about these things. Here I will experience an urge or desire to discover the truth, but will not regard the issues as worth caring about or the questions as worth answering. Again, there seems to be something objectionable about the idea that the truth has value because it answers the subject's inquiries or questions, in those instances where the subject does not *value* getting the truth on these questions. Desires that a subject would reflectively reject are, therefore, poor candidates if we are looking for factors capable of grounding final values.

In light of these considerations, we can identify a gap between the thought that people are naturally curious about discovering the truth on some issue, and the idea that the truth on that issue is valuable for its own sake. Truth does not seem to be intrinsically valuable *just* because it satisfies a subject's natural curiosity. So how might this gap be closed?

3

How can we move from the claim that we are naturally curious to discover the answers to particular questions, to the claim that answers to those questions are valuable in themselves? This problem is pressing, given that there might be something amiss with our curiosity or concern, and which therefore casts doubt upon the value of the truths which constitute the object of that curiosity or concern. A simple solution is to *idealize* the relevant concern for truth. Thus, we might claim that the truth on a certain issue is valuable, not if someone does care about or desire the truth on that issue, but only if the person *would* care about the truth under certain idealized conditions: if, for instance, the person would desire the truth on that issue were she fully rational. In this way we might rid ourselves of problems (1)–(3), on the grounds that a process of rational idealization will bring to light whether the subject's interest is instrumental or intellectual, will ensure that inquiries are not based upon false beliefs, and will rule out curiosity that results from irrational compulsions. We might therefore maintain that it is the satisfaction of natural and *rational, idealized* curiosity which has final value.

Such a move is not without its problems. The main difficulty is this: the usual method of idealizing a subject's actual cares and concerns so as to close the gap between her desires and her reasons or values involves appeal to what the subject would care about if she were better informed, and had a more coherent and consistent desiderative profile.¹⁷ But then we need to determine how

¹⁷ See e.g. Michael Smith (1994: ch. 5).

much information is required in order for someone's curiosity to be *sufficiently* informed, and hence sufficiently rational. If the person is provided with too little information, it is not clear that she will be rid of all of her false beliefs, or able to determine whether her interest is instrumental or intellectual. If so, we might worry about whether her curiosity is sufficiently idealized. If the person is provided with too much information, however, it is more than likely that her curiosity will disappear. So the proponent of idealization faces a problem in determining how it is possible to rule out inquiries based upon false beliefs, let us say, without ridding the subject of things to be curious about. A traditional move in value theory aimed at idealizing our desires and concerns, and thus at closing the gap between what we care about and what is valuable, fails to be plausible when applied to intellectual curiosity.

A better response, it seems to me, is to accept that (1)–(3) represent normative failings in our inquiries, and yet argue that true beliefs which result from those inquiries are nevertheless valuable for their own sakes. This is possible because true beliefs in these cases will *also* satisfy an interest in the truth which is general, open-ended, and unrestricted; as a result, the true beliefs will possess value because they satisfy *this* concern. So the fact that true beliefs might result from particular inquiries that are problematic does not count against the claim that they are valuable in themselves. In slightly different terms: a better response is to argue that all true beliefs have value in virtue of satisfying an interest in the truth as such, or in the truth *simpliciter*. We do not need to restrict valuable truths to those that constitute the right answers to *particular* questions or issues that interest us. This move enables us to reply to the problems listed at the end of § 2. By invoking a general concern, we render the possibility that our particular concern is instrumental moot. We might very well be interested in the truth for instrumental reasons, but insofar as we have a general concern for the truth as such—a concern to believe all and only true propositions—then the resulting true belief will have value insofar as it satisfies this general concern. And insofar as our natural curiosity is open-ended or unrestricted, the fact that our particular inquiries sometimes rest upon false beliefs, or are sometimes motivated by irrational compulsions, will also fail to be important. For again, we can say that truths have value in virtue of satisfying our open-ended and unrestricted natural curiosity, rather than in virtue of satisfying particular restricted interests. If we assume, then, that we are naturally curious about the truth *simpliciter*, then the worries about how to close the gap between our valuing the truth and the value of truth disappear.

This assumption might remove the worries about explaining the value of truth in terms of the satisfaction of particular but potentially suspect inquiries; but is it justified on any other grounds? In other words, do we have independent reasons for thinking that our natural curiosity is open-ended and unrestricted, or that we are naturally curious about the truth *simpliciter*? Certainly some philosophers think so: Jonathan Kvanvig and Michael Lynch each maintain that our natural

curiosity is directed towards the truth as such.¹⁸ Kvanvig, for instance, claims that it is in

the nature of interests to lack specificity: we do not have an individual interest in the truth of the claim that our mothers love us, that the president is not a crook, that Wyoming is north of Mexico, and so on. What we have is a general interest in the truth, and that interest attaches to particular truths in the manner of instantiation in predicate logic. The default position for any truth is that our general interest in the truth applies to it. (2003: 41)

On this picture, 'finding out the truth with respect to *any* subject would seem to be worthy of our interest.'¹⁹ Of course, it doesn't seem like this to us in our everyday lives. From our perspective, we are only curious about the truth on particular questions or issues; we regard the vast majority of truths with indifference. But for Kvanvig and Lynch this indifference is a function of our general interest in the truth being overridden by our *practical* concerns. Thus, our interest attaches to particular truths and not others due to the practical importance that the particular truths possess. This means that there are very many truths that we are not *in fact* motivated to discover; there are 'all sorts of true beliefs that are not worth having, all things considered.'²⁰ Nevertheless, if we could abstract from our practical concerns, then we would have some interest in even the most trivial truth. The view that our everyday interests are particular is thus compatible with the thought that it is '*prima facie* good to believe even the most trivial truth' (Lynch 2004: 55), on the grounds that even the most trivial truth will satisfy our natural curiosity.²¹

There are, however, good reasons to think that the gap between valuing and value cannot be closed in this way. Kvanvig, as noted, claims that it is in the nature of interests to lack specificity. But it seems to me that this is false. Insofar as our interests are generated by our natural curiosity, they will be specific interests, since curiosity, by its very nature, involves *selective* attention: when we are curious, we focus and attend to some things rather than others. It is this fact,

¹⁸ Lynch (2004: 128–36) provides an argument that we *ought* 'to care about the truth in general or *as such*' because this is 'essential to intellectual integrity'. I do not have time to consider Lynch's interesting argument in detail here. But even if Lynch is correct, my argument in § 5 tell against the claim that some truth is valuable because it results from a concern that we *ought* to have.

¹⁹ Grimm (2008: 728). Grimm, however, disagrees with this line on the value of truth. He asks, 'why think that, in virtue of our curiosity, it is attaining the truth *per se*, or finding out how things stand with respect to *any* subject, that has a standing value for us? (p. 730)' And: 'if we think that pursuing the truth is intrinsically valuable, then why are we unapologetically indifferent to so many truths?' (p. 726). A similar line is taken by Ernest Sosa (2003), Duncan Pritchard (2007), and others.

²⁰ Lynch (2004: 55). These will include, I assume, truths that satisfy morally dubious desires.

²¹ Kvanvig makes a similar point, claiming that 'it is certainly true that we view some truths as simply unimportant . . . that fact need not be taken to undermine the intrinsic value of truth, for it may be that our practical needs, goals, and interests interact with the intrinsic value of truth so that some truths are simply unimportant, all things considered, even though truth is still intrinsically valuable from a purely cognitive point of view or *from the point of view of inquiry for its own sake*' (2003: 6).

rather than any appeal to practical concerns or considerations, which explains why our curiosity generates particular interests in our actual circumstances. If so, it is implausible to suppose that we could have a general interest in the truth which is generated by our natural curiosity.

To see this, we can note that curiosity is an emotion.²² Now it is widely accepted that emotions constitute reactions to objects and events which are (potentially) significant to us. As Annette Baier writes:

We all accept the idea that emotions are reactions to matters of apparent importance to us: fear to danger, surprise to the unexpected, outrage to insult, disgust to what will make us sick, envy of the more favoured, gratitude for benefactors, hate for enemies, love for friends, and so on. And sometimes the felt emotion can precede knowledge of precisely what the danger, the insult, the nauseating substance, and so on is. Emotion then plays the role of alerting us to something important to us—a danger, or an insult. (Baier 2004: 200)

Emotions play this role by capturing and directing our attention onto important objects and events. As Aaron Ben Ze'ev puts it, 'like burglar alarms going off when an intruder appears, emotions signal that something needs attention' (2000: 13). Thus, when we are afraid our attention is rapidly and automatically fixed upon some potentially dangerous object or event; moreover, our attention typically remains fixed—it is both captured and consumed—so that it is difficult for us to shift our focus away or concentrate on other things. This indicates that fear necessarily involves *selective* attention: when we are afraid our attention is triggered by and drawn towards particular objects and events, from amongst 'the vast array of stimuli that constantly impinge on the senses' (Izard and Ackerman 2000: 257). But what is true of fear is true of other emotions, such as disgust, shame, anger, guilt, joy, surprise—and curiosity. Each of these emotions has the role of alerting us to something significant or important, and fulfils this role by directing our attention onto certain objects and events and not others. Curiosity, like other emotions, involves selective attention.²³ Given this picture of what emotions are, therefore, we can doubt that our natural curiosity is open-ended and unrestricted.²⁴ And if so, the claim that all truths have value because they

²² It is therefore a mistake to identify curiosity with a simple desire for or interest in truth, since emotions involve more than mere desire or interest.

²³ The need for a particular mechanism to fulfil the role of alerting us to important or significant objects and events stems from the fact that human beings are presented with vast amounts of information about the state of the world and the state of themselves, only some of which will be important to them. Given that human beings have limited mental resources, they thus face a problem of efficiently locating or identifying which information in their environment is important. Our emotional systems, at least in part, are thought to have evolved in order to solve this problem, and they do so in part by reflexively and automatically capturing and focusing the subject's attention. For details of this line, see e.g. Vuilleumier, Armony, and Dolan (2003).

²⁴ It might be argued that curiosity is more like an intellectual character trait than an emotion, and as such underlies a general desire or interest. For instance, someone who has the character trait of being open-minded is understood to be receptive to new ideas and ways of thinking on

satisfy our open-ended natural curiosity is false. The gap between our curiosity and the value of truth remains.²⁵

4

Interestingly, these reflections on natural curiosity as an emotional response suggest a different way in which the gap between inquiry and the value of truth might be closed. Consider the general and uncontroversial point made above, namely that emotions function to alert us to things of (potential) importance or significance. It seems clear, however, that different emotions function to detect different *kinds* of importance or significance. Thus, fear alerts us to danger, anger to insults, guilt to moral wrongdoing, disgust to contamination, and so on. Here different kinds of importance can be explained in terms of different kinds of evaluative properties, the possession of which by an object or event determines a 'standard of correctness' for the relevant emotion. So on this view, fear is correct if and only if its object is dangerous, anger is fitting if and only if its object is insulting, disgust is appropriate if and only if its object is contaminated, and so on. Now since natural curiosity is an emotion, we might wonder about the kind of evaluative property which determines its standard of correctness. That is, we might inquire about the kind of importance or significance that natural curiosity is meant to alert us to, given that it is not truth *simpliciter*. A tempting answer to this question is that curiosity functions to alert us to matters that are, in and of themselves, *interesting* or *fascinating*. If so, we might propose that true beliefs have final value if and only if they constitute answers to intrinsically interesting or fascinating questions.

any subject that she considers. By the same token, to describe someone as a curious person is to say that she wants to find out about any subject that is brought before her. But I don't think that the distinction between emotions and traits supports the idea that curiosity is open-ended and unrestrictive. First, we might, in describing someone as a jealous person, attribute to them a character trait; but this does not mean that jealousy is not a fully-fledged emotion. The fact, therefore, that we can describe someone as a curious person does not cast doubt upon curiosity's status as an emotion either. Second, the curious person *is* selective in what she attends to. Someone who literally wants to find out about any subject that is brought before her would seem to suffer from a psychological *disorder*; given the number of subjects that we encounter every day, such a person would be constantly switching her attention to each new subject as it occurs, and as a result would fail to attend properly to very many of them. The curious person, by contrast, focuses on some subjects *at the expense of* others.

²⁵ It is open to Kvanvig and Lynch to appeal to what we would want if we abstracted from the fact that we are limited, finite epistemic beings facing the problem of how to detect important or significant information. They might suggest that in such circumstances we would be curious about everything. But it seems to me that in such circumstances we would be creatures with no need for any of the central emotions, and hence creatures with no curiosity to fulfil. If we assume that the natural emotions are needed precisely *because* we are limited, finite epistemic beings, then we cannot appeal to the natural curiosity of beings that are unlimited or infinite. God, for instance, has no need for automatic and reflexive emotional responses to events of potential significance, any more than God has need for hunger and thirst.

Now talk of certain questions and subjects being intrinsically interesting might strike us as implausible, even on a loose understanding of 'intrinsic'. Such talk suggests that there are questions that call for curiosity, or subjects that we should take an interest in; by the same token, it suggests that there are other questions that we should not be curious about, or subjects that should not be a source of fascination. Against this, it might seem that there is little room for a gap between what interests me and what is interesting, or between what bores me and what is boring. It might seem, that is, that there is nothing more to the claim that some question is interesting than the fact that it is a question that interests me, or a question that I am curious about. Nevertheless, I think that there are good reasons to understand the claim that some question or subject is interesting as a normative claim, a claim about what *merits* interest and curiosity, rather than a descriptive claim about my present psychological state. Something similar can be said about claims that a topic or event is boring. In support, note that the fact that I am currently bored by *x* is clearly insufficient for me to judge that *x* is boring; perhaps I'm aware that I have taken a sedative which is responsible for my listless state, and that in other conditions I wouldn't be bored by it. This suggests, at the very least, that the judgement that *x* is boring is a judgement that I would be bored by *x* when conditions are *appropriate*, or that *x* has certain features that would, in the right conditions, give rise to boredom on my part. It suggests, in other words, a 'response-dependent' account of what it is for something to be boring, in which case there will be some normative distance between what we find boring (or interesting) and what we judge to be boring (or interesting). If so, however, there is no obvious problem in claiming that although someone finds *x* fascinating they ought not to do so: think, for instance, of a person who is high on drugs and is fascinated by the (mundane) pattern on the living room carpet. By the same token, there is no obvious problem in claiming that although a person on sedatives fails to find the news that they are directly related to Genghis Khan fascinating, they nevertheless should. All that these claims require for plausibility is the idea that there are certain (often hard to specify) features of objects and questions in virtue of which boredom and fascination are merited, and, relatedly, certain (perhaps easier to specify) conditions which count as appropriate for considering such objects and questions.²⁶

²⁶ Of course, claims about the features that merit curiosity and boredom will be disputable, and it might be difficult to identify any feature that all and only interesting things have in common other than the feature of eliciting curiosity. This, after all, is why a response-dependent account of such properties as 'interesting' and 'boring' has a good degree of plausibility. Nevertheless, there are certain more-or-less public standards which govern our practice of evaluating objects and events in terms of whether they are interesting or not, which again suggests that claims about these properties are normative claims. There are publicly accepted paradigms of boring activities, for instance: we talk of watching paint dry or grass grow. By the same token, we are unsurprised when people around the world become fascinated by the prospect of a full lunar eclipse, or are interested in the opening ceremony of the Olympic games. The fact that we have these practices and standards suggests that we are right to regard certain things as meriting interest and other things as meriting boredom.

If we accept the idea that certain questions can indeed be intrinsically interesting or fascinating, then the proposal that true beliefs have final value if and only if they constitute answers to such questions is attractive for a number of reasons.²⁷ The first is that it respects what we say and think about (the objects of) our curiosity. Thus, when asked to explain why we are curious about a subject which seems to have no practical pay-off, we claim simply that we find the subject interesting or fascinating in itself. We note that certain questions strike us as interesting in their own right, or that we are fascinated by certain subjects simply because of the nature of those subjects. So from our own perspective we are curious about and value the truth on issues which seem intrinsically interesting or fascinating. A second reason is that it supplies the right kind of normativity to our interest in the truth, and thus avoids making the value of truth a matter of what answers our *actual* particular inquiries or questions. We saw above that what we are naturally curious about might be insufficiently normative; and one reason for this is that we can be naturally curious about things that are not intrinsically interesting, in the same way that we can be naturally afraid of things that are not dangerous.

Finally, the idea that natural curiosity is an *emotional* response directed towards what is interesting or fascinating can accommodate the fact that the range of subjects and questions that people find intrinsically interesting is very wide. Thus, some people might be interested in questions about how and why the dinosaurs became extinct or about how Stalin came into power, whilst others are curious about the prospects for life on Mars or about the mental capacity of chimpanzees, whilst still others are naturally intrigued about which Scottish football team boasts the longest losing run or about whether anything eats wasps. But such variety in the objects of our natural curiosity is to be *expected*, given that curiosity is an emotional response. After all, it hardly needs saying that different people are afraid of different things: Adam might be afraid of house spiders while Beth thinks that they are cute; Clare might be terrified of rollercoasters while Dan finds them thrilling. It is equally obvious that there is a similarly wide variety in objects of anger, shame, guilt, disgust, and joy. Now clearly the explanation for this variety is that there are a whole range of contingent factors that can influence or trigger or modify a particular emotional response. Think, for instance, of how fear can depend upon or be influenced by a subject's beliefs (that the snake is poisonous, or harmless) or desires (for safety, or for thrills), or of how guilt might be influenced by religious upbringing, or disgust by early memories of and associations with food. Given the great variety in what people experience, believe, desire, remember, and feel, and the role that all of these things play in

²⁷ As before, we should understand 'intrinsic' here in a loose way, as meaning 'non-instrumental', to allow for the fact that what is interesting or fascinating to depend upon a certain *relation* between questions and responses. If so, then the fact that something is interesting will not depend entirely on features which are intrinsic to that thing.

the emotional life, it should come as no surprise that emotions have an extremely wide variety of objects. Since curiosity is also an emotion, and hence subject to influence and modification by a wide range of factors, we should therefore expect that there is a great variety in the kinds of questions and topics that people find intrinsically interesting, and which trigger natural curiosity. The emotional status of natural curiosity thus explains why natural curiosity is both selective *and* wide-ranging.²⁸

In light of this, the best option for those wishing to explain the value of truth in terms of the satisfaction of our natural curiosity is to appeal to the idea that certain questions trigger our natural curiosity because they are intrinsically interesting or fascinating, and to maintain that true beliefs have final value if and only if such beliefs constitute answers to intrinsically interesting questions.

5

Unfortunately, even if we accept the idea that there is a class of intrinsically interesting questions which trigger our natural curiosity, we should reject the claim that true beliefs have value if and only if they answer such questions. For this biconditional is false in both directions. To see this, consider the claim that true beliefs have value only if they answer interesting questions. Now we can assume that if there is a class of intrinsically interesting questions, then there are also certain subjects or questions that are intrinsically *uninteresting*. Suppose, then, that I am naturally curious about the truth of an uninteresting question: I am naturally curious about something that I ought not to be curious about, an issue which ought not to trigger my natural curiosity. But it does not follow that the truth on this question *lacks* value. Consider, for instance, these truths:

- If the average man never trimmed his beard, it would grow to nearly 30 feet long in his lifetime.
- On average, right-handed people live nine years longer than their left-handed counterparts.
- The average American will eat 35,000 cookies in a lifetime.
- The average lead pencil will draw a line 35 miles long or write approximately 50,000 English words.
- The only country in the world that has a Bill of Rights for Cows is India.

²⁸ In a sense, then, claims that a subject merits interest or is intrinsically fascinating will display a *sensitivity* to such influences. Still, it makes sense to claim that someone with certain interests ought to find a particular subject fascinating even if she does not, or that someone with these experiences and feelings ought not to be fascinated by those mundane topics. So we can maintain that claims about what is interesting and boring are normative, whilst also accommodating the fact that what is interesting for one person in his circumstances (which include his experiences, feelings, etc.) might not be interesting for another person in different circumstances.

- Ancient Chinese artists would never paint pictures of women's feet.
- During the Alaskan Klondike gold rush (1897–8), potatoes were practically worth their weight in gold. Potatoes were so valued for their vitamin C content that miners traded gold for potatoes.
- Duelling is legal in Paraguay as long as both parties are registered blood donors.
- It is forbidden for aircraft to fly over the Taj Mahal.
- A Saudi Arabian woman can get a divorce if her husband doesn't give her coffee.
- Offered a new pen to write with, 97% of all people will write their own name.
- In 1913, the US tax on a \$4,000 annual income was one penny.²⁹

I take it that at least some people will find at least some of these truths interesting or fascinating in their own right. As a result, some people will be glad to know them: these truths have *value* for the relevant people. However, it seems clear that such truths can result from inquiries which are intrinsically uninteresting, boring, or mundane. For instance, I might discover the fact about the value of potatoes during the Alaskan Klondike gold rush because I'm naturally curious about the value of tuber crops during the late nineteenth century. But isn't this a prime example of an intrinsically uninteresting or boring subject? By the same token, it is difficult to see how the class of intrinsically interesting or fascinating questions could include the question of whether there are any links between the consumption of hot beverages and divorce rates in Saudi Arabia, or the issue of the average number of words a lead pencil will write. There might even be interesting truths which emerge from the standard example of a boring or uninteresting inquiry, namely an inquiry into the number of grains in a handful of sand. Suppose, for instance, that I start counting the number of grains in a handful of sand on Devil's Beach in Rio de Janeiro, and discover that the number is 666. This strongly suggests that there are some truths which are valuable—we find them interesting or fascinating or curious or surprising, and as a result are glad to have them—even though they result from intrinsically uninteresting inquiries. In short, we can say that *boring* subjects can contain fascinating truths. So we have reason to doubt the claim that truth is valuable only if it results from an inquiry into an intrinsically interesting or fascinating subject.

Indeed, we can make a stronger claim here, and maintain that there are true beliefs whose value depends upon the fact that they do *not* result from inquiry, whether this is an inquiry into an intrinsically interesting subject or not. This is a particular instance of a more general point—made persuasively by David Velleman—that there are things 'whose value depends on their having been unanticipated and unsought' (2000: 91). Velleman argues that

²⁹ Information collected from: www.corsinet.com/trivia, www.healthypotato.com, www.myworldsouthamerica.com, and www.trivia-library.com.

a person's well-being or happiness can include *windfalls*—‘things such as unsolicited affection or spontaneous merriment’—that would lose their value if they resulted from the person's own efforts (Velleman 2000: 91). Something similar might be said about valuable truths. That is, we might think that there are *epistemic* windfalls, truths whose value depends upon the fact that they were unsought, and so depends upon the fact that they were not the results of inquiry. For example, if unsolicited affection constitutes a positive value in our lives, we might think that unsolicited knowledge of affection does as well. Thus, I might learn that ‘she loves me’ because of her unsolicited declaration of love. Here my true belief has value that it would lack if it resulted from inquiry on my part. There seem to be a great number of surprising but welcome truths that fall into this category. So the efforts of inquiry are sometimes incompatible with the intrinsic value of true beliefs.

What of the other conditional, that a truth has final value if it results from inquiry into an intrinsically interesting or fascinating subject? Here too there are good reasons to be sceptical. For sometimes inquiries into fascinating subjects yield truths which are mundane, uninteresting, and unimportant. This should come as no surprise to people who have eagerly embarked on an interesting research project, full of questions which spark both their curiosity and the curiosity of the grant-awarding body, only to realize, towards the end of the inquiry, that any truths to be had here are trivial and uninteresting. If this were not the case, there would hardly be the widespread phenomenon of people being *disappointed* with the answers to their inquiries. Of course, some of the disappointment might be due to frustration of practical interests: an interesting answer on a hot research topic might make one's name, advance one's career, secure a job in a prestigious university, and so forth. But some of the disappointment might be due to the fact that no one, including oneself, is remotely interested *in* or fascinated *by* the truths one has discovered. Suppose that I'm intrigued by the question of whether the pyramids at Teotihuacan in Mexico and at Giza in Egypt were built by the same people, and engage in a research project in order to find this out. Whilst the question is an interesting one, it seems clear that the answer—‘no’—has little in the way of intrinsic value. In fact, we can see that ‘yes/no’ answers to interesting questions are often uninteresting. The question of whether anything eats wasps is certainly interesting, but the true answer ‘yes’ has little in the way of intrinsic value. Similarly, it's an interesting question whether being cold raises the risk of catching a cold. Once again, however, we gain little of intrinsic value when we learn that the true answer is ‘no’. What *does* have value is knowing which creatures eat wasps—apparently 133 species do, including birds, frogs, crabs, bats, and humans—and knowing *why* being cold does not raise the possibility of catching a cold.³⁰ These kinds of answers certainly are interesting and valuable. But this indicates that there

³⁰ For these questions and answers about wasps and colds, see O'Hare (2005: 85–8; 2006: 33).

can be different answers to the same question, each of which is true, but which differ in their value. It will be difficult to maintain, in light of this, that *all* true answers to interesting questions are themselves intrinsically valuable; sometimes true answers to interesting questions are not worth having.

6

What conclusions can we draw from this? The first is that some truths have value for us, but not because they answer intrinsically interesting questions. We are, that is, glad to know certain things once we become aware of them, even though such valuable truths do not, and in some cases could not, result from inquiry on our part. By the same token, we have seen that there can be boring answers to interesting questions. This indicates that the value of truths when possessed is only contingently related to our natural and legitimate curiosity. In particular, it is a kind of value which is only contingently related to the interest we have in answering intrinsically interesting questions. If so, we have good reason to reject the claim that true beliefs have value if and only if they answer intrinsically interesting questions. The gap between valuing the truth on some intrinsically interesting topic, and the value of the truth on that topic, remains.

This strongly suggests a second conclusion, about the nature of our epistemic aim or goal. At the beginning of § 1 we noted the common view that truth is the aim of inquiry, and the suggestion that truth is therefore something valuable or desirable. However, insofar as there can be truths which answer our inquiries and yet lack value, we might think that the aim of inquiry is not simply truth, nor simply the truth on questions that interest us, nor even the truth on questions that are intrinsically interesting, but rather *intrinsically interesting* truths. That is, if we are to equate the aim of inquiry with something that is genuinely valuable, then it should be identified with the goal of attaining interesting or fascinating or surprising true beliefs. So on this view, it is *pro tanto* good, for all *p*, to believe that *p* if and only if *p* is true *and* interesting.³¹

Finally, this take on our epistemic aim tells us something important about the relation of priority between the value of truth and natural curiosity. The idea that truth is the aim of inquiry was taken to imply, again at the beginning of § 1, that we value getting the truth *because* we value getting answers to questions that interest us. Since we are sometimes interested in the truth for its own sake,

³¹ This is a version of Grimm's 'Prima Facie Good Principle' (2008: 730). As with claims about subjects or questions, I take it that the judgement that a truth is intrinsically interesting or intrinsically boring is itself a normative claim. I thus assume that the fact that we are not, in our present circumstances, glad to possess some truth does not mean that the truth is uninteresting, and I also assume that the fact that we are currently interested in some truth does not mean that it is interesting. As with subjects and questions, I assume that an interesting truth is one which *merits* a reaction of interest when possessed.

due to our natural curiosity, the thought was that truth is valuable for its own sake because it satisfies our natural curiosity. But we can see now that this gets things the wrong way around. If our epistemic aim is to amass interesting truths, then it seems that we value getting answers to interesting questions *because* this is the best or most reliable way of getting interesting truths. The fact that there is only a contingent relation between answering interesting questions and attaining interesting truths does not mean that there is no connection. And it is surely plausible to suppose that we are going to hit upon interesting truths more often if we investigate interesting, rather than boring, subjects. We can deny, in other words, that truth is valuable because it satisfies our natural curiosity. Instead, we should maintain that satisfying our natural curiosity is valuable because this is the best or most reliable way of attaining valuable truths.

The thought that truth is valuable because it satisfies our natural curiosity thus misrepresents the role of curiosity in our epistemic lives. If I am right, then the real role of natural curiosity is to alert us to interesting or fascinating subjects. Since inquiring about these subjects is the best or most reliable way of hitting upon interesting or significant truths, then natural curiosity does indeed play a vital role in helping us to attain our epistemic goals. Natural curiosity is importantly, although contingently, related to the value of true belief.³²

REFERENCES

- Alston, W. (2005). *Beyond Justification*. Ithaca, NY: Cornell University Press.
- Baier, A. (2004). 'Feelings that Matter', in Robert Solomon (ed.) *Thinking about Feeling*. Oxford: Oxford University Press.
- Ben Ze'ev, A. (2000). *The Subtlety of Emotions*. Cambridge, MA: MIT Press.
- David, M. (2005). 'Truth as the Primary Epistemic Goal', in M. Steup and E. Sosa (eds.), *Contemporary Debates in Epistemology*. Oxford: Blackwell Publishing.
- Foley, R. (1987). *The Theory of Epistemic Rationality*. Cambridge, MA: Harvard University Press.
- Goldman, A. (1986). *Epistemology and Cognition*. Cambridge, MA: Harvard University Press.
- (1999). *Knowledge in a Social World*. Oxford: Oxford University Press.
- Grimm, Stephen (2008). 'Epistemic Goals and Epistemic Values', *Philosophy and Phenomenological Research*, 77: 725–44.
- Hempel, C. (1965). 'Aspects of Scientific Explanation', in *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*. New York: Free Press.

³² An earlier version of this paper was presented at a workshop on *Epistemic Value* at the University of Stirling in November 2006. I would like to thank the participants at the workshop, and my commentator, Peter Baumann, for their feedback. I would also like to thank the anonymous referees for this volume for their useful criticisms and advice. And I am particularly grateful to Fiona Macpherson and Duncan Pritchard for their helpful comments on, and conversations about, earlier drafts of this material.

- Hurka, T. (2001). *Virtue, Vice, and Value*. Oxford: Oxford University Press.
- Hookway, C. (1996). 'Questions of Context', *Proceedings of the Aristotelian Society*, 96: 1–16.
- Izard, C., and Ackerman, B. (2000). 'Motivational, Organizational, and Regulatory Functions of Discrete Emotions', in M. Lewis and J. Haviland-Jones (eds.), *Handbook of Emotions*, 2nd edn. New York: Guildford Press.
- Korsgaard, C. (1983). 'Two Distinctions in Goodness', *Philosophical Review*, 92: 169–95.
- Kvanvig, J. (2003). *The Value of Knowledge and the Value of Understanding*. Cambridge: Cambridge University Press
- (forthcoming, 2009). 'The Value of Knowledge and Truth', D. M. Borchert (ed.), *Encyclopedia of Philosophy*, 2nd edn. London: Macmillan Reference Books.
- Lynch, M. (2004). *True to Life*. Cambridge, MA: MIT Press.
- Moore, G. E. (1922). 'The Conception of Intrinsic Value', in his *Philosophical Studies*. London: Routledge & Kegan Paul.
- O'Hare, M. (2005). *Does Anything Eat Wasps?* London: Profile Books.
- (2006). *Why Don't Penguins' Feet Freeze?* London: Profile Books.
- Pritchard, D. (2007). 'Recent Work on Epistemic Value', *American Philosophical Quarterly*, 44: 85–110.
- Schaffer, J. (2005). 'Contrastive Knowledge', in T. Gendler and J. Hawthorne (eds.), *Oxford Studies in Epistemology*. Oxford: Oxford University Press.
- Smith, M. (1994). *The Moral Problem*. Oxford: Blackwell Publishing.
- Sosa, E. (2003). 'The Place of Truth in Epistemology', in M. DePaul and L. Zagzebski (eds.), *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Oxford University Press.
- Velleman, J. D. (2000). 'Is Motivation Internal to Value?', in his *The Possibility of Practical Reason*. Oxford: Oxford University Press.
- Vuilleumier, P., Armony, J., and Dolan, R. (2003). 'Reciprocal Links between Emotion and Attention', in R. Frackowiak et al. (eds.), *Human Brain Function*, 2nd edn. San Diego: Academic Press.

13

The Trivial Argument for Epistemic Value Pluralism, or, How I Learned to Stop Caring about Truth

Berit Brogaard

It is generally agreed that truth is an epistemic goal. Some think it is the highest epistemic goal. It is what we strive towards when we engage in inquiry. Let us follow Michael Lynch in taking ‘inquiry’ to mean ‘the range of epistemic practices we engage in when asking and answering questions, whether banal (“where did I put my other sock”) or sublime (“can something come from nothing?”)’ (p. 225 above). And let us call the view that truth is the highest epistemic goal ‘epistemic value monism.’ Epistemic value pluralism then is the view that truth is only one of many epistemic goals, and that its supremacy can be overridden.¹

Epistemic value monists are not oblivious to the fact that there are other things besides truth that we might be after when we engage in the practices associated with asking and answering questions. For instance, in asking ‘how do I look?’ we might be seeking a compliment, and in asking ‘did you ever feel this way when you were my age?’ we might be looking for support and encouragement. But epistemic value monists think that most of these other goals are not *epistemic* goals but goals of a different kind. Nor are epistemic value monists oblivious to the fact that there are other epistemic goals besides truth, for instance, avoiding the gambler’s fallacy, forming beliefs while you are sober, being justified in believing that it is raining, being open-minded, and so on.² But they think that these other goals are intermediary. They are goals only insofar as they facilitate our getting to the truth.

Which of the two theses, epistemic value pluralism or epistemic value monism, is correct is thought to be a difficult epistemological question: it can be decided only by reflection on the practices we engage in when asking and answering questions. In particular, it is not something that can be decided by looking at the semantics of a particular group of expressions.

¹ Riggs (manuscript). Riggs takes epistemic value pluralism to be consistent with truth being the overarching aim of inquiry. But this is not how I will use the term.

² See Douven (2005).

This, anyway, is how things are thought to be. But if a certain very plausible semantic hypothesis is correct, then this is not how things are. The plausible hypothesis is that there are genuinely relative expressions—expressions whose extension depends not only on a context of use but also on a context of assessment. If there are genuinely relative expressions, then epistemic value monism is false. I will call this *the trivial argument for epistemic value pluralism*. After formulating the argument, I will look at three possible ways to refute it. I will then argue that two of these are unsuccessful, and defend the third, which involves denying that there are any genuinely relative expressions.

1. THE TRIVIAL ARGUMENT FOR TRUTH VALUE PLURALISM

Epistemic value monists take truth to be our highest epistemic goal. What they mean by that is that it is *prima facie* good to believe what is true. So, if I go outside in order to find out whether it is raining, then the aim of my inquiry is to believe that it is raining if it is raining and believe that it is not raining if it is not. But, of course, I haven't satisfied the epistemic goal if I believe that it is raining and also believe that it is not raining. The epistemic goal is not merely to believe what is true, but rather to believe all and only what is true. Where a proposition *p* is simply true iff it has the truth-value *true* relative to the actual world, and its truth-value does not depend on any parameters besides the world parameter, we may formulate the epistemic goal as follows:

(EG) For any proposition *p*, the epistemic goal is for you to believe that *p* iff it is simply true that *p*.³

(EG) seems *prima facie* plausible. It is a normative thesis about what we should aim to believe, and if *p* is true, then it is *prima facie* plausible that we should aim to believe *p*. (Where 'should' (or 'ought') is the epistemic 'should' (or 'ought'), I assume that the epistemic goal is for you to believe that *p* iff you should (or ought) to believe that *p*.) Of course, one could hold that since it is practically impossible for us to believe all true propositions, we shouldn't aim to believe all true propositions, but just as many as we can. We return to this objection below. Even if we set aside this concern, however, (EG) is not entirely happy. For on the assumption that for any proposition *p*, we ought to believe either *p* or not-*p*, (EG) turns out to entail a particular 'Fregean' semantic thesis, namely:

(FT) For any proposition *p*, the truth-value of *p* is relative only to the world of evaluation.

³ Compare Lynch (this volume, Ch. 10).

Of course, it is open to deny the inference from (EG) to (FT) by rejecting the assumption that for any proposition p , we ought to believe either p or not- p . If p doesn't have a determinate truth-value, then arguably we shouldn't aim to believe p or not- p . But then (EG) does not entail (FT).

There is, however, a weaker connection between (EG) and (FT). If (FT) is false, then there are propositions which are not simply true or simply false. And if there are such propositions, then it is very plausible that we ought to believe some of them. But (EG) implies that if there is some proposition p such that p is not simply true, then you shouldn't believe p . So, if there is just one proposition p such that p isn't simply true and you ought to believe p , then (EG) is false. So, even if we reject the assumption that for any proposition p , we ought to believe either p or not- p , the plausibility of (EG) stands and falls with the plausibility of (FT).

(FT) gains *prima facie* plausibility from the fact that propositions are supposed to be whatever it is that we believe when we believe something. That is, propositions are supposed to be the objects of propositional attitudes such as beliefs. But the objects of belief seem to possess their truth-values eternally. Quite plausibly I do not simply believe that John is sitting. My belief seems more specific than that. For example, I might believe that John is sitting at 3 p.m. on March 8, 2006. But the proposition that John is sitting at 3 p.m. on March 8, 2006, possesses its truth-value eternally. If the proposition that John is sitting at 3 p.m. on March 8, 2006, is true at 3 p.m. on March 8, 2006, then it is also true at 4 p.m. on March 8, 2006. If entities which do not possess their truth-value eternally are not specific enough to be the objects of belief, and propositions are objects of belief, propositions are true or false only relative to a world of evaluation, which is to say that (FT) is true.

But, as David Kaplan and others have made vivid,⁴ there are also arguments against (FT). Propositions are thought to be the semantic values of sentences in contexts. For example, the semantic value of an utterance of the sentence 'John is currently a firefighter' at 3 p.m. on March 8, 2006, is supposed to be the proposition that John is a firefighter at 3 p.m. on March 8, 2006. But the view that propositions are the semantic values of sentences in context makes trouble for (FT). Consider, for instance, my utterance at 3 p.m. on March 8, 2006 of 'John is a firefighter' and your utterance at 3 p.m. on March 8, 2006, of 'it has been that John is a firefighter.' In standard semantics, 'it has been that' is treated as a sentential operator. It takes us from our current circumstance of evaluation to a circumstance in the past. Your utterance of the sentence 'it has been that John is a firefighter' is true just in case 'John is a firefighter' is true at a time before the time of utterance. So, if my utterance is false and yours true, then the semantic value of 'John is a firefighter' has different truth-values depending on whether the sentence occurs in a temporal context or not. But this strongly

⁴ Kaplan (1989). See also Lewis (1980) and King (2003: 195–246).

suggests that 'John is a firefighter' does not express a Fregean proposition. For sentences are supposed to have the same semantic value regardless of the linguistic environment in which they are uttered.

So, if propositions are both the objects of belief and the semantic values of sentences in context, then it would seem that they both do and do not possess their truth-values eternally.

David Lewis and others have proposed to resolve this tension by distinguishing between the compositional semantic value of a sentence in context and the proposition expressed or asserted by the sentence.⁵ 'It was the case that John is a firefighter' expresses a proposition that is the object of your belief that John was a firefighter, but it also has a distinct compositional semantic value that is composed of the semantic values of the parts.

The distinction between the propositional semantic values and the compositional semantic values of sentences allows us to preserve (FT). But there is a different way to preserve (FT). Jeffrey King has recently argued that temporal expressions (including the tenses) are not operators on sentences but object-language quantifiers over times.⁶ On this view, 'John is a firefighter' has a hidden indexical variable whose value is supplied by context. This variable is bound when 'John is a firefighter' is embedded in a temporal context, as in 'John was a firefighter.' The latter expresses, relative to context, the proposition that there is a time t such that t is earlier than the time of speech, and John is a firefighter at t . If the tenses are object-language quantifiers, then there is no need to recognize a special compositional semantic value for tense operators to operate on. Treating tense operators as object-language quantifiers thus makes it possible to maintain the view that propositions are both the objects of belief and the semantic values of sentences.

There is much to be said for both of these proposals. However, I think that there is independent reason for rejecting (FT).⁷

One reason is this: when I say 'John is a firefighter' and you say 'Nuh-uh, John is not a firefighter. He is a police officer,' it seems we are disagreeing. But this is not so if (FT) is correct. For when I say 'John is a firefighter' at t_1 , I assert that at t_1 John is a firefighter, and when you say 'John is not a firefighter' at t_2 , you assert that at t_2 John is not a firefighter. But at t_1 *John is a firefighter* and at t_2 *John is not a firefighter* do not contradict each other. Note, how unnatural-sounding the following exchange is.

⁵ Lewis (1980). See also Salmon (1986), Dummett (1991), and Stanley (1997a; 1997b).

⁶ King (2003) backs up his proposal with linguistic evidence. However, I think the evidence he presents does not actually support a quantificational treatment of the tenses. One example King gives is this: 'yesterday, I turned off the stove.' King thinks that the tense operator approach is bound to say that 'yesterday' is an operator. Treating 'yesterday' as an operator does indeed give us the wrong result. However, there is good reason not to treat 'yesterday' as an operator. 'Yesterday' is better treated as a component part of the composite tense operator 'it was the case yesterday that'. This treatment gives us the right result. For the details of this argument, see Brogaard (manuscript: ch. 2).

⁷ For a lengthier defense of temporalism, see Brogaard (manuscript: chs. 1 and 3).

(A and B are talking on the phone. B is standing outside the door of an office where a conversation is taking place between John and his superior).

A: . . . John is a firefighter.

(Behind closed doors the superior shouts: 'you are fired!')

B: I guess you are right. But John is not a firefighter.

If A asserted that John is a firefighter at t_1 , then we should expect B to reply exactly as he does. For it is still true at t_2 that John is a firefighter at t_1 . But it would make no sense for B to respond in this way. It would make much more sense for him to reply with 'no, you are wrong. I am standing outside the superior's office, and his superior just told him that he was fired.'

A critic may rejoin that the argument is too quick. Perhaps the semantic value of the hidden indexical variable in 'John is a firefighter' is not always the time of speech. 'John went to a local bar' need not be interpreted as 'John went to a bar that is local to the speaker' but may be interpreted as 'John went to a bar that is local to John,' 'John went to a bar that is local to the hearer,' 'John went to a bar that is local to John's grandmother,' and so on. The speaker can more or less freely fix the hidden indexical variable associated with 'local.' Perhaps the same goes for 'John is a firefighter.'

I think, however, that this version of eternalism is even less plausible. Consider the following dialogue between A and B:

AS I SAID

A: What does John do for a living?

B: He is a firefighter.

C: Nuh-uh, he was fired two years [weeks/days/minutes/seconds] ago. He is a lawyer now.

B: [*aggravated*] So, as I said, John is a firefighter.

B's last remark is exceedingly odd. If, however, the value of the alleged indexical variable can be freely chosen, then we should expect it to be felicitous.

A second reason to doubt that (FT) is correct is that it seems to yield an inadequate account of belief revision. If I learn that John was fired, intuitively I will not simply add one more belief to my belief system; I will dispose of my belief that John *is* a firefighter and add the belief that John *was* a firefighter but is no longer.

But this is not what (FT) gives us. Suppose I believe that John is a firefighter at 3 p.m. on April 5, 2004. If (FT) is true and you inform me that John is not a firefighter at 3 p.m. on August 8, 2004, then this would give me no reason to discard my old belief. After all, I have been given no reason to believe that my old belief is false. If eternalism is true, then all of our beliefs are temporally specified in this way. So if they are true at all, then they remain true when the world changes. Contrary to appearances, then, changes in the world will give us no reason to revise our original beliefs.

But changes in the world do give us reason to revise our beliefs. If I am told on August 8, 2006, that John has been fired, then there is something I cease to believe: I cease to believe that John is a firefighter. The following exchange illustrates this point:

RANKINGS

(A and B are talking in the hallway)

A: Our department is number two on the Leiter Report!

(One week later)

B: What you said last week isn't true anymore. I just saw the rankings.

A: Of course, it is. Haven't you read Frege? Maybe that's why we dropped.

If (FT) is true, then the proposition A believes on the earlier occasion cannot change its truth-value over time. So A's last remark should be perfectly fine.

A third reason to question (FT) is that it yields an unrealistic account of belief retention. If (FT) is true, then I do not believe that John is a firefighter, period. I believe that John is a firefighter at 15:00 on July 5, 2004, that John is a firefighter at 15:01 on July 5, 2004, and so on. So, to 'continue' to believe what I would express with 'John is a firefighter' for five years is to believe an indefinite number of propositions of the form 'John is a firefighter at t ,' one for each moment of time during the five years. But that can't be right. Surely, there is just a single belief that I continue to have for the five years, namely the belief that John is a firefighter.

Relatedly, when we say 'I still believe that p ,' the advocate of (FT) is required to say that we mean that we still believe the same time-indexed proposition. But if that were right, then it would be hard to explain cases of the following sort:

DECEIT

Wife: When I married John I thought he was a police officer. Thirty years later I still believed he was a police officer. Turns out that he was fired two years into our marriage.

or:

LOST LOVE

Friend: Yes, Barbara did love you ten years ago. So you were right back then. But you *still* believe that she loves you, don't you Peter?

or:

DEFENSE

Student: I think my dissertation is done.

Supervisor: You do? Well, I think you are wrong. Work on it for a few more weeks. Then read it again. If you *still* think that it's done, then we'll talk.

If (FT) is true, then what we believe when we believe something is a time-indexed proposition. But it is hardly the case that the wife in DECEIT means that she still believed the same time-indexed proposition after thirty years, viz. the proposition *my husband is a police officer at t* , where t is some time thirty years ago, that the friend in LOST LOVE means that Peter still believes the proposition *Barbara loves me at t* , where t is some time ten years ago, or that the supervisor in DEFENSE is asking S to return if S still believes the proposition *S's dissertation is done at t* , where t is the time of their exchange. To my mind, such cases raise one of the most pressing kinds of problems for advocates of (FT).

But if there is at least one proposition p such that p is true at one time but false at another, then (FT) is false. A more plausible candidate for being the object of my belief that John is a firefighter is what we might call a *tensed proposition*. Tensed propositions change their truth-values across time. In fact, as Arthur Prior has argued,⁸ there is overwhelming evidence for thinking that present-tensed propositions just are temporal propositions, that is, propositions that determine functions from worlds and times to truth-values. If he is right, then (FT) is false. The existence of tensed propositions, of course, does not rule out that there could not also be tenseless propositions, for instance, the proposition that John is a firefighter at 3 p.m. on July 5, 2006. But (FT) states that for any proposition p , the truth-value of p is relative only to the world of evaluation. And this is false, if some tensed propositions are worthy of belief.

Let us now return to the above formulation of the epistemic goal. Above we formulated the epistemic goal as follows:

- (EG) For any proposition p , you should believe that p iff it is simply true that p .

As we saw, (EG) is plausible only if (FT) is:

- (FT) For any proposition p , the truth-value of p is relative only to the world of evaluation.

So, if (FT) is false, so is (EG). To conclude on the grounds that (EG) is false that epistemic value monism should be rejected would be too quick, however. For there is a semantic thesis (KT) in the neighborhood that may be true, namely:⁹

- (KT) For any proposition p , and any time t , p has a truth-value relative to t .

⁸ Prior (1957: 9–10; 1967: 8–10, 14–15; 1968a: 17–23; 1968b: 101–19). Prior points out that there is a difference between the present tensed sentence 'John *is* a firefighter' and the indexical sentence 'John is currently a firefighter.' Unlike tense operators like 'it has been that,' 'it will be that' and 'it is the case that,' 'it is currently the case that,' and 'it is now the case that' take primary scope with respect to other operators.

⁹ 'KT' stands for 'Kaplan's thesis'. The world is kept fixed.

Keeping the world fixed, the epistemic goal can then be formulated as follows:

- (EG') For any proposition p and any time t , the epistemic goal is for you to believe that p at t if and only if p has the truth-value true relative (only) to t .

A problem arises, however, if there are genuinely relative expressions in the language, as a number of relativists have argued. John MacFarlane, for example, has argued that 'know' and 'might' are plausible candidates for being relative non-sentential expressions, whereas future contingents (e.g. 'there will be a sea battle tomorrow') are plausible candidates for being relative sentential expressions. Others have offered a relativistic semantics for predicates of personal taste (e.g. 'fun' and 'tastes great'), color expressions (e.g. 'is the same color as'), and non-standard uses of gradable adjectives (e.g. 'huge' and 'rich').¹⁰

A relative expression is one whose *extension* varies with parameters determined by the context of use *and* a context of assessment.¹¹ A context of assessment is a context in which a sentence is considered or evaluated for truth. Any context of use is a context of assessment, but not every context of assessment is a context of use. In a context in which I evaluate John's utterance of the sentence 'I am hungry,' for example, 'I am hungry' is evaluated for truth in my context but, according to the relativists, it is not used in that context.

Relativists claim that the existence of relative expressions in the language requires us to revise standard semantics. The content of relative expressions is context-invariant. A relative expression has the same semantic value at any context of use and any context of assessment. Its semantic value is (or determines) a function from circumstances of evaluation to extensions. But the relativists' circumstances of evaluation are non-standard. They contain a world and a time parameter determined by the context of use, and a judge parameter determined by the context of assessment. So, if I utter the sentence 'this chili tastes great' at some time t in the actual world @, and you evaluate my utterance at some time t^* in the actual world @, then our contexts together determine a $\langle @, t, \text{you} \rangle$ circumstance. The sentence 'this chili tastes great' is true relative to my context of use and your context of assessment iff the proposition expressed by 'this chili tastes great' is true at $\langle @, t, \text{you} \rangle$.

Standard semantics does not allow for circumstances that are determined partially by a context of use and partially by a context of assessment. The circumstances of standard semantics are fully determined by the context of use. So, if there are any relative expressions in the language, then standard semantics is in need of revision.

¹⁰ See e.g. Egan et al. (2005), Kölbel (2002; 2003; 2008), MacFarlane (2003; 2005a; 2005b), Richard (2004), and Weatherston (2008).

¹¹ The extension of an uttered sentence, relative to a circumstance, is its truth-value, the extension of a predicate, relative to a circumstance, is the set of things that possess the property expressed by the predicate.

The relativists' proposed revision of standard semantics does not affect the truth-value and content of just any old sentence. For instance, if I utter the sentence 'I am human' at a time t in the actual world @, my context of use and my context of assessment determine a $\langle @, t, \text{me} \rangle$ circumstance. If you then evaluate my utterance for truth at t^* , then our contexts together determine a $\langle @, t, \text{you} \rangle$ circumstance. If, however, none of the expressions in 'I am human' is a relative expression (which is very plausible), the truth-value of my utterance of the sentence 'I am human' will be the same regardless of who evaluates the utterance.

But if relativism is right, then there are some propositions which have no truth-value relative only to a world and a time. For instance, the proposition expressed by my utterance of the sentence 'this chili tastes great' has no truth-value only relative to my context of use. Rather, it is true or false only relative to a context of use and a context of assessment. So, if John considers my assertion and dislikes the chili, then the proposition is false, and if Mary considers my assertion and likes the chili, then the proposition is true.

But this makes trouble. If I enjoy this chili at time t , I will be inclined to believe at t that it tastes great. But if 'tastes great' is a genuinely relative expression, then 'this chili tastes great' does not have a truth-value only relative to a world and a time. By (EG'), I should not believe that this chili tastes great. But if 'tastes great' is a genuinely relative expression, then 'this chili does not taste great' does not have a truth-value only relative to a world and a time either. So, by (EG') I should not believe that this chili does not taste great. But this seems just plainly false. Vagueness to one side, I should believe either that this chili tastes great or that it does not taste great. But then what I should or shouldn't believe isn't fully determined by what is or isn't true relative to my temporal location in the actual world, which is to say, some other value or goal besides truth-at-a-time regulates inquiry as well. Hence, if epistemic value monism is committed to (EG'), then epistemic value monism is false.

2. FIRST REPLY: REVISING THE EPISTEMIC GOAL

Both epistemic value monists and epistemic value pluralists should be worried about this argument. For, where the conclusion of the argument concerns the aim of the practices that we engage in when asking and answering questions, the main premise is an empirical hypothesis about English (or some other language). But intuitively, the question of what the ultimate aim of the practices we engage in when asking and answering questions cannot be decided by looking at the semantics of a particular group of expressions. There is thus good reason to think that the trivial argument is flawed. The question is where it goes wrong.

It might perhaps be thought that the existence of relative expressions in English does not prove that epistemic value monism is wrong but only that we

need to revise epistemic value monism to take account of assessment-sensitivity. Arguably, there are independent problems with the above articulations of the epistemic goal. As Lynch has argued, we cannot believe everything that is true, and it is highly doubtful that our epistemic goal is to be God.¹² Lynch suggests that we revise the standard formulation of the epistemic goal to take account of our human deficiencies. The epistemic goal, he suggests, may be formulated as follows. For any p that we are able to consider, our epistemic goal is to believe p if and only if p is (simply) true. Lynch admits that it might be difficult to unpack the modality of 'is able to' but adds that whatever it means, it is not supposed to mean that we should believe only the true propositions that we happen to consider.

Lynch's reformulation of our epistemic goal seems to fit the relativist's purposes better than our original formulation, for it concerns only propositions that we are able to consider. It thus appears to be about believers in contexts of assessment rather than believers in contexts of use. It is natural, therefore, to suggest that we formulate the epistemic goal as follows:

- (EG'') For any time t and any proposition p that you are able to consider, the epistemic goal is for you to believe that p at t if and only if it is true that p relative to the standards in play in your context of assessment C_A at t .

So, if John is able to consider the proposition that this chili tastes great, then he should believe that this chili tastes great if and only if this chili tastes great relative to the standards in play in his context of assessment.

Unfortunately, (EG'') is mistaken if relativism is right. For relativists say that the extension of a sentence is a function of the world and time of the context of use, and the standards in play in the context of assessment. A proposition is not simply true relative to parameters determined by a context of assessment; it is true relative to parameters determined by a context of use and a context of assessment. The context of use determines a time and a world at which to evaluate the proposition and the context of assessment determines the standards with respect to which the proposition is to be evaluated. If John said last year before graduating, 'I am a student,' then the proposition expressed is true relative to the time determined by John's context of use but not true relative to the time determined by my current context of assessment. So the truth-value of the proposition that John is a student will depend on the context of use.

The following principle might be thought to do better as a formulation of our epistemic goal:

- (EG''') For any time t , any proposition p that you are able to consider and any context of use C_U , the epistemic goal is for you to believe that p

¹² See p. 226 above.

at t if and only if it is true that p relative to the time t^* determined by C_U and the standards in play at your context of assessment at t .

But this principle can't be right either. For the present-tensed proposition that John is a student may be true relative to the time determined by Peter's context of use but false relative to the time determined by Mary's context of use. By (EG''), we should and shouldn't believe that John is a student. So, (EG''') instructs us to have inconsistent beliefs.

Our best option, it seems, is to formulate the epistemic goal as follows:

(EG''') For any time t and any proposition p that you are able to consider, the epistemic goal is for you to believe that p at t if and only if it is true that p relative to the time determined by your context of use C_U at t and the standards in play at your context of assessment C_A at t .

But (EG''') is not entirely happy either. One problem with (EG''') is that your context may not be a context of use with respect to the proposition that you are considering. If, for example, John utters the sentence 'this chili tastes great' and you are considering the proposition expressed by John's utterance at t , then your context is not the context of use. Of course, we could treat any sequence of parameters of the right sort as a context, as David Kaplan has proposed (1989), and there is much to be said for Kaplan's notion of a context; but there is a further, and more serious, problem with (EG''').

(EG''') says that our highest epistemic goal is to believe what is true in the special case where the context of use C_U and context of assessment C_A are identical. But this seems like a refutation of the claim that there are relative truths. We have a notion of being true at a context of use C_U and a context of assessment C_A , but it is not one that has any value to us except when $C_U = C_A$.¹³ Truths at a context of use C_U and a context of assessment C_A are not truths that are worthy of belief; they are not truths we should believe. In fact, they are truths that we should *not* believe. To put the point another way: the only truths worthy of belief are those that might count as truths by non-relativist lights. So, if the highest epistemic goal is to believe what is true in the special case where the context of use C_U is identical to the context of assessment C_A , then our highest epistemic goal is to believe all and only non-relative truths; genuine relative truth is of no value to us. But if there are genuinely relative truths, then this is unacceptable. So, it seems doubtful that reformulating the epistemic goal will help to block the trivial argument.

¹³ MacFarlane raises a related objection to a related formulation of the norm of assertion. See MacFarlane (2005a: 331).

3. SECOND REPLY: REJECTING RELATIVISM

A second way to respond to the trivial argument is to insist that the fact that relativism cannot take truth to be our prime epistemic goal shows that the assumption that there are genuinely relative expressions is mistaken. The problem with this reply, however, is that the evidence for there being genuinely relative expressions is strong.

The most obvious candidates for being relative expressions are predicates of personal taste (e.g. 'fun' and 'tastes great'). As Peter Lasersohn points out, 'with predicates of personal taste, we actually operate from a position of epistemic *privilege*, rather than the opposite. If you ride the roller coaster, you are in a position to speak with authority as to whether it is fun or not; if you taste the chili, you can speak with authority as to whether it is tasty.'¹⁴ But there is equally good evidence for thinking that epistemic expressions (e.g. 'might,' 'know,' and 'justified'), color expressions (e.g. 'the same color as'), and so on are relative expressions. For expository simplicity, I shall here focus on predicates of personal taste. But my argument can easily be extended to take account of other expressions thought to be genuinely relative.

The first piece of evidence for treating predicates of personal tastes as relative expressions comes from the fact that there is widespread disagreement about what tastes great, tastes horrible, is fun, and so on. From the fact that I find this chili tasty I cannot infer that you will find it tasty as well. Semantic invariantists think expressions like 'taste great' have the same value and extension in any context of use and any context of assessment. So they cannot account for the insolubility of disputes about what is tasty without ascribing an implausible systematic semantic incompetence to ordinary speakers.

Contextualism seems more plausible. For example, it might be suggested that predicates of personal taste contain a hidden indexical variable which has the speaker as a default value. So, when I say 'this chili tastes great,' I am in effect asserting that this chili tastes great to me. And when you say 'this chili tastes great,' you are asserting that this chili tastes great to you. So construed, contextualism can explain why the truth-value of sentences containing predicates of personal taste seems to vary from speaker to speaker. But it cannot explain why disagreements about what tastes good, what is fun, and so on, often seem faultless. The following disagreement, for example, seems faultless:

(I) John: This chili tastes great.

Mary: I disagree. It does not taste great at all.

¹⁴ Lasersohn (2005: 655). See also Lasersohn (2008).

The following exchange, on the other hand, points toward semantic incompetence on Mary's part:

(II) John: I am hungry.

Mary: I disagree. I am not hungry at all.

If the contextualists are right, then (I) and (II) both indicate that Mary is semantically incompetent. But most of us fail to see the alleged semantic mistake made by Mary in (I). So, if contextualism is right, then most of us are ignorant of the semantics of predicates of personal taste.

A related problem for contextualism is that when our internal affective states change, we often retract our earlier judgments. Consider the following exchange:

(III) John: . . . chocolate ice cream tastes great.

Mary: But a few months ago you said that it didn't taste great.

John: I was mistaken. I didn't really appreciate the taste back then.

If John means different things by 'tastes great' on the two different occasions, as contextualism tells us, we should expect John to respond differently, for instance, with 'yes, but I was right back then as well. I just didn't mean the same thing by taste great.' But this is not how we behave at all.

Other expressions typically taken to be context-sensitive, such as 'tall' and 'flat,' seem to function differently.¹⁵ If we are talking about the height of men in general, I might say 'Michael Jordan is tall. He is 6'6"!' If we are talking about the height of NBA basketball players, on the other hand, I might say 'Michael Jordan is not tall. He is only 6'6".' If reminded of the apparent discrepancy, I will not say: 'I was mistaken. I didn't realize how small he really is.' Instead, I will say something like: 'I only committed myself to Jordan being tall for a man. The average height for a man is 5'10". But the average height for an NBA player is 6'6". So, Michael is tall for a man, but not for an NBA player.' This sort of difference between predicates of personal taste and other expressions generally regarded as context-sensitive count against treating the content of the latter as contextually variable. It seems that we do not, as a matter of fact, take the properties expressed by 'tastes great' and 'fun' to vary across contexts of use.

A third problem for contextualism turns on the behavior of predicates of personal taste in propositional attitude reports. Suppose that after her conversation with John Mary meets Peter who is throwing a party. Peter wants to know how John likes chocolate ice cream. Mary replies:

(1) He thinks it tastes great.

¹⁵ See e.g. MacFarlane (2005*b*).

(1) seems to be an adequate report of what John believes, provided, of course, that John complied with the norms of assertion on the earlier occasion. The report seems adequate even if Mary believes it doesn't taste great. But if she believes it doesn't taste great, then the contextualist is bound to say that John and Mary mean different things by 'tastes great.' Since Mary is the speaker of (1), the content of 'tastes' depends on her internal affective states. So (1) represents John as believing that chocolate ice cream tastes great, by Mary's lights. But evidently this is not what John believes.¹⁶

Both invariantism and contextualism apparently have difficulties accounting for the linguistic data involving predicates of personal taste. Relativism fares considerably better.

First, relativism explains why there is widespread and insoluble disagreement about what is tasty, what is fun, and so on. There is widespread disagreement about what is tasty, what is fun, and so on because the truth of sentences containing predicates of personal taste is relative to the standards in play in the context of assessment. So the proposition that this chili is tasty is true when assessed by John only if John finds this chili tasty, and it is true when assessed by Mary only if Mary finds this chili tasty.

Second, relativism explains why the disagreement appears faultless. For relativists about 'fun' and 'tastes great' hold that the semantic content of these expressions is constant across contexts. So, if John and Mary disagree about whether this chili tastes great, then there is a proposition the truth-value of which is the object of their disagreement, namely the proposition that this chili tastes great.

Third, relativism explains why we are strongly inclined to retract our earlier judgments when our internal affective states change. We are inclined to do this because sentence truth is relativized to a context of use and a context of assessment. 'Chocolate ice cream is disgusting' might be true when assessed from John's earlier perspective but false when assessed from his later perspective.

Fourth, relativists claim to be able to explain why we are quick to disquote. Since the content of 'tastes great' is contextually invariant, Mary's report 'John believes that chocolate ice cream tastes great' simply expresses the proposition that John stands in the *believe* relation to the proposition that chocolate ice cream tastes great.

Relativism about predicates of personal taste thus provides a powerful explanation of the linguistic data involving predicates of personal taste without having to resort to such dubious phenomena as semantic blindness or widespread semantic incompetence. Rejecting the thesis that there are relative expressions without offering an alternative that explains the linguistic evidence equally well would be ad hoc and desperate.

¹⁶ See e.g. Cappelen and Lepore (2003).

4. THIRD REPLY: OFFERING AN ALTERNATIVE

A third thing one might do when faced with the trivial argument is reject the assumption that there are relative expressions but offer an alternative theory that explains all the same linguistic data as relativism. This is the option I will defend.¹⁷ I propose the following theory as a replacement for relativism. Predicates of personal taste have the same content across context; the content determines a function from circumstances of evaluation to extensions. The default circumstance is fully determined by the context of use which, like Kaplan, I will take to be a sequence of a speaker, an addressee, a world, a time, and a location. But unlike Kaplan's default circumstances my default circumstances contain a judge parameter whose default value is the speaker. So my assertion of 'this chili tastes great' is true only if it is true at a circumstance containing me as the judge. To set it apart from standard contextualism I shall call this view *perspectivalism*.¹⁸ Perspectivalism, I will now argue, explains all the same data as relativism without relativizing sentence truth to a context of use and a context of assessment.

The prime reason for relativizing sentence truth to contexts of use and contexts of assessment is that there seem to be cases of sentences that are true relative to a <world, time, judge> circumstance, where the value of the world and time parameters are determined by the context of use and the value of the judge parameter is determined by the context of assessment. Consider, for instance, John's utterance at *t* of:

- (2) The sentence 'this chili will taste great tomorrow' was false when Mary uttered it on the earlier occasion.

(2) ascribes falsehood to a sentence-in-context, namely the sentence reported in (2). The truth-value of the reported sentence depends on the time and world at which Mary uttered it, and the standards in play in John's context. Thus Mary's utterance of the sentence 'this chili will taste great tomorrow' is true relative to John's context of assessment iff the chili in question tastes great by John's standards, the day after Mary uttered the sentence.

But, as I will now argue, we do not need to distinguish between a context of use and a context of assessment to get this result. We only need to treat 'tastes great' as a perspectival expression, that is, an expression whose extension is a function of a circumstance containing a judge parameter whose default value is the speaker.

¹⁷ This position is further developed in Brogaard (2008a; 2008b; 2008c).

¹⁸ MacFarlane calls this brand of contextualism 'nonindexical contextualism.' See MacFarlane (2007; 2008). As we will see, my account differs from MacFarlane's nonindexical contextualism with respect to how it deals with reported sentences.

Cases like (2) seem to pose a problem for all brands of contextualism because the embedded sentence is mentioned but does not appear to be used. This appearance is illusory, however.¹⁹ The quoted sentence in (2) is a direct speech report. But, as François Recanati has argued (2004: § 3.2), there is good reason to think that a reported sentence is used, not simply mentioned. One reason is that the quoted material in a direct speech report may be available for copying, as in:²⁰

- (3) ‘I’m going to find a way out,’ she said; and she did.

‘And she did’ is elliptical for ‘and she did find a way out.’ So the elided material ‘find a way out’ is available for copying. But this suggests that the quoted material is used. For if the sentence were merely mentioned, the quoted material would not be available for copying, witness “‘I’m going to find a way out’” is a sentence; and she did.’

Another reason to think a quoted sentence in a direct report is used not simply mentioned is that expressions in the matrix clause may depend anaphorically on expressions in the quoted sentence. Consider, for instance:

- (4) ‘Give me your money_i, or I’ll shoot,’ he said, but I didn’t give it_i to him.

The pronoun ‘it’ in the matrix clause is anaphoric on ‘your money.’ But this requires that ‘your money’ picks out an individual for ‘it’ to refer to.

Recanati suggests that direct speech reports create a shifted context with respect to which the expressions that occur in the scope of the report are assigned a semantic value. For example, if I assert ‘and then John said, “I am leaving now”’ at *t*, then the direct speech report creates a shifted context in which John is the value of the speaker parameter, and some past time *t** is the value of the time parameter. So the value of ‘I’ is John, and the value of ‘now’ is *t**. Direct speech reports can thus be seen to function as context-shifters, or monsters, as Kaplan called them.²¹ Kaplan thought there weren’t any context-shifting *operators* in English, but didn’t rule out direct speech report shifts.²²

Let us return to John’s utterance at *t* of:

- (2) The sentence ‘this chili will taste great tomorrow’ was false when Mary uttered it on the earlier occasion.

¹⁹ See Brogaard (2008*a*).

²⁰ The first argument is Recanati’s (2004: § 3.2, example 4); the second is my own.

²¹ Kaplan (1989: 510): ‘no operator can control the character of the indexical within its scope, because they will simply leap out of its scope to the front of the operator. I am not saying we could not construct a language with such operators, just that English is not one. And such operators could not be added to it.’

²² As he puts it, ‘there is a way to control an indexical, to keep it from taking primary scope, and even to refer it to another context. Use quotation marks. If we mention the indexical rather than use it, we can, of course, operate directly on it.’

If direct speech reports create shifted contexts, we can explain the intuitive truth-value of (2) without relativizing sentence truth to a context of use *and* a context of assessment. John's context is a context of use with respect to which the expressions in (2) are assigned a semantic value. The quoted sentence 'this chili will taste great tomorrow' is not assigned a semantic value with respect to John's context. Instead, it is assigned a semantic value with respect to a shifted context in which John is the speaker, and the time is some time t^* before t . So the semantic value of any indexical expressions in the quoted sentence will depend on parameters of the shifted context. The semantic value of 'tomorrow' is thus the day after t^* , not the day after t . But 'tastes great' is not an indexical; its content does not vary across contexts of use; so it has the same semantic value relative to John's and Mary's contexts. Its content is a function from $\langle \text{world, time, judge} \rangle$ triples to extensions.

Besides fixing the semantic values of the expressions in (2), John's context determines a $\langle w, t, \text{John} \rangle$ circumstance with respect to which (2) is evaluated. The quoted sentence in (2) is not evaluated with respect to this circumstance. But the reason is not that it is quoted. For direct speech reports do not by themselves shift the parameters of the circumstance of evaluation. To shift the parameters of the circumstance we need a circumstance shifter. For example, the proposition expressed by 'I am hungry' is evaluated with respect to the circumstance determined by the context of use in an assertion of 'John says, "I am hungry"' but with respect to a shifted circumstance in an assertion of 'John *said*, "I am hungry"'. The matrix of (2) does contain a circumstance shifter, namely the past tense operator. But tense operators are selective; they shift only the time feature of the circumstance of evaluation. The past tense operator shifts the time t to some time t^* before t . So the circumstance with respect to which the quoted sentence is interpreted is a $\langle w, t^*, \text{John} \rangle$ triple. The quoted sentence is true relative to this circumstance iff this chili will taste great the day after t^* by John's standards. So (2) is true iff Mary's assertion of 'this chili will taste great tomorrow' is false by John's standards, which is as it should be.²³

Our account of (2) can be extrapolated to account for any case of a sentence being evaluated for truth. For on Kaplan's notion of context any sequence of a speaker, addressee, world, time, and location counts as a context. A fortiori, a context in which the speaker is not saying anything counts as a context. Hence, a context at which a sentence is quietly evaluated for truth will count as a context. The present account can thus handle any case where a sentence is being evaluated for truth without relativizing sentence truth to a context of assessment.

²³ I assume that there is a difference between the truth-predicate for utterances when restricted to the meta-language and when it is not so restricted. When the truth-predicate for utterances is restricted to the meta-language, it functions in just the way suggested by Kaplan. See Brogaard (2008a).

Before returning to our discussion of epistemic value I would like to consider a couple of possible objections to the present proposal. One might be worried about the posited non-standard circumstances. One cannot just posit the existence of whatever circumstances one would like. The reason Kaplan included a time and world parameter in his circumstances was that those parameters could be shifted by circumstance-shifting operators. But, it may be said, there is no operator that can shift the value of the judge parameter (Stanley 2005).

By way of reply, I think that there *are* circumstance shifters that can shift the value of the judge parameter. 'According to so-and-so, p ' is true iff, for all worlds w compatible with the exact content of what so-and-so would assent to at the world of speech @ and the time of speech t^* , p is true at $\langle w, t^*, \text{so-and-so} \rangle$. But consider now my utterance of:

(5) According to John, this chili tastes great.

(5) is true iff for all worlds w compatible with the exact content of what John would assent to at $\langle @, t^* \rangle$, the proposition *this chili tastes great* is true at $\langle w, t, \text{John} \rangle$. In other words, for the proposition *This chili tastes great* to be true at a circumstance compatible with the exact content of what John would assent to here and now, someone who would assent to the very things John would actually assent to must be the judge in that circumstance. Attitude verbs function in much the same way. 'John believes that this chili tastes great' is true iff for all worlds w compatible with the exact content of what John believes at $\langle @, t \rangle$, the proposition *this chili tastes great* is true at $\langle w, t, \text{John} \rangle$.

Another problem is posed by factive propositional attitude verbs, such as 'realize,' 'know,' 'regrets that,' etc. Consider, for instance, my utterance at t of:

(6) John realizes that this chili tastes great.

(6), it seems, is true only if it is true at a circumstance that has me as the judge and a circumstance that has John as the judge. But it is straightforward to explain why we get this sort of double-indexing. As we have seen, propositional attitude verbs function as operator-forming expressions. For any mental state operator O , and any object s , ' s O s p ' is true only if for all worlds w compatible with the exact content of what s O s at $\langle @, t \rangle$, the proposition p is true at $\langle w, t, s \rangle$. But where O is a factive operator, Op is true at a circumstance of evaluation c only if p is true at c . So the truth of 'John realizes that this chili tastes great' requires that *this chili tastes great* be true for John. But 'John realizes that' is a factive operator. So (6) is true at a $\langle @, t, \text{speaker} \rangle$ circumstance only if *this chili tastes great* is true for the speaker.

A third problem is that there seem to be cases where the default value of the judge parameter is not the speaker (Lasersohn 2005: 672 ff.). There are two kinds

of problematic cases. The first kind of problematic case is presented by generics statements.²⁴ On the present account, it seems, a generic statement such as ‘if the wine is disgusting, people should spit it out’ is false if I dislike the wine, hardly a satisfactory result. The second kind of problematic case is constituted by cases of persuasion and cases of reporting. For instance, if you refuse to taste the chili I just made, I might say ‘c’mon it’ll taste great.’ What I say is true, it seems, if it is true by your standards, or perhaps by both of our standards. Or if someone asks me how Mary liked the rides at the theme park, I might reply ‘the roller coaster was fun, but the free fall was a bit scary’ (Lasersohn 2005: 672). Here again, it seems that what I said is true just in case the roller coaster was fun by Mary’s standards, and the free fall was a bit scary by Mary’s standards. So in these cases it seems that the default value of the judge parameter is not the speaker.

However, I think these cases are instances of what Recanati has called ‘free circumstance shift.’ A circumstance shift is free just in case it is not operator-controlled but is controlled instead by the speaker’s intentions. Here is a test of free shifts of the judge parameter. There is free circumstance shift just in case inserting an appropriate circumstance shifter in the front of the sentence does not change the truth-value of the sentence. As ‘For all x , if the wine is disgusting according to x , x should spit it out,’ ‘c’mon it’ll taste great, according to you,’ and ‘according to Mary, the roller coaster was fun, but the free fall was a bit scary’ have the same truth-values as the originals, these cases are plausible cases of free circumstance shifts.

Circumstance shifts are quite common in narratives. Consider, for instance:²⁵

- (7) John is completely nuts. Nothing tastes great anymore, not even chocolate ice cream.

Inserting ‘according to John’ at the front of the second sentence will not affect the truth-value in this case either.

My working hypothesis then is that sentences containing perspectivals are true relative to circumstances of evaluation that contain not only a world and a time parameter but also a judge parameter. The default value of the judge parameter is the speaker, but circumstance shifters that operate on the value of the judge parameter (e.g. ‘according to John’) can shift the value. Since the default value of the judge parameter is determined by the speaker parameter of the context of *use*, perspectivals are context-sensitive. But they are not context-sensitive in the same way as indexicals. Where indexicals have variable contents, perspectivals have constant contents but variable extensions.

²⁴ Thanks to Kent Bach here.

²⁵ This example is based on Recanati’s (2000) example: ‘John is totally paranoid. Everyone wants to kill him, including his own mother.’

The proposed semantics explains the same linguistic data as relativism but without relativizing sentence truth to a context of use and a context of assessment. The relativist takes the truth of any sentence, including unembedded sentences, to be relative to a context of use and a context of assessment. So, if I assert 'this chili tastes great,' what I asserted has no truth-value, except relative to a context of assessment. Perspectivalism, on the other hand, takes unembedded sentences to have a truth-value relative to the context of use. So, if I assert 'this chili tastes great,' then what I said has a truth-value relative to my context of use. Since the revisions to standard semantics required by perspectivalism are far less radical than those required by relativism, perspectivalism is preferable to relativism on methodological grounds.

Though I will not be able to argue for it here, it is quite plausible that any candidate for being a genuinely relative expression can be adequately accounted for as a perspectival. This is good news for the epistemic value monist, or anyone who is unconvinced by the trivial argument. For the hypothesis that there are perspectivals in the language does not create a demand for epistemic value pluralism. Perspectivalism is consistent with the following formulation of the epistemic goal:

- (EG''') For any proposition p that you are able to consider and any time t , the goal is for you to believe that p at t if and only if it is true that p at the circumstance of evaluation determined by your context of use C_U at t .

As (EG''') is a very meek revision of (EG'), the hypothesis that there are perspectivals is compatible with truth being the highest epistemic goal.

5. CONCLUSION

Relativism provides a compelling explanation of linguistic data involving predicates of personal taste (e.g. 'tasty' and 'fun'), moral expressions (e.g. 'right' and 'wrong'), future contingents (e.g. 'there will be a sea battle tomorrow'), color expressions (e.g. 'is the same color as'), epistemic modals (e.g. 'might'), and so on. But relativism is incompatible with epistemic value monism—the view that truth is the highest epistemic goal. The argument from relativism to epistemic value pluralism I called *the trivial argument for epistemic value pluralism*, because it threatens to show all too swiftly that epistemic value monism is false. I have argued that there is no way to formulate epistemic value monism in such a way as to take account of the assessment-sensitivity of the target expressions. Nor can we simply dismiss relativism without providing an alternative. I have argued, however, that a version of non-indexical contextualism can account for the same data as relativism without relativizing sentence truth to contexts of assessment. On this account, sentence truth is relative only to a context of use, but the

environments in which sentences are evaluated create shifted contexts at which to interpret the evaluated sentences.²⁶

REFERENCES

- Brogaard, B. (2008a). 'Sea Battle Semantics,' *Philosophical Quarterly*, 58: 326–35.
 — (2008b). 'In Defence of a Perspectival Semantics for "Know",' *Australasian Journal of Philosophy*, 86: 439–59.
 — (2008c). 'Moral Contextualism and Moral Relativism,' *Philosophical Quarterly*, 58: 385–409.
 — (ed.) (2008d). *Synthese: Special Double Issue on Relative Truth*.
 — (manuscript). 'Transient Truths—An Essay in the Metaphysics of Propositions.'
 Cappelen, H., and Lepore, E. (2003). 'Context-Shifting Arguments,' in Hawthorne and Zimmerman 2003.
 Douven, I. (2005). 'The Pragmatics of Belief,' Epistemic Value Work Shop, Stirling, October 2005.
 Dummett, M. (1991). *The Logical Basis of Metaphysics*. Cambridge, MA: Harvard University Press.
 Egan, A., Hawthorne, J., and Weatherson, B. (2005) 'Epistemic Modals in Context,' in G. Preyer and G. Peter (eds.), *Contextualism in Philosophy: Knowledge, Meaning, and Truth*. Oxford: Clarendon Press.
 Hawthorne, J., and Zimmerman, D. (2003). Philosophical Perspectives, 17: *Language and Philosophical Linguistics*. Oxford: Blackwell.
 Kaplan, D. (1989). 'Demonstratives,' in J. Almog, H. Wettstein, and J. Perry (eds.), *Themes from Kaplan*. New York: Oxford University Press.
 King, J. C. (2003). 'Tense, Modality, and Semantic Values,' in Hawthorne and Zimmerman (2003).
 Kölbel, M. (2002). *Truth without Objectivity*. London: Routledge.
 — (2003). 'Faultless Disagreement,' *Proceedings of the Aristotelian Society*, 104: 53–73.
 — (2008). 'The Evidence for Relativism,' in Brogaard (2008d).
 Lasersohn, P. (2005) 'Context Dependence, Disagreement, and Predicates of Personal Taste,' *Linguistics and Philosophy*, 28: 643–86.
 — (2008). 'Relative Truth, Speaker Commitment, and Control of Implicit Arguments,' in Brogaard (2008d).
 Lewis, D. (1980). 'Index, Context and Content,' in S. Kanger and S. Öhman (eds.), *Philosophy and Grammar*. Dordrecht: Reidel; reprinted in Lewis, *Papers in Philosophical Logic*. Cambridge: Cambridge University Press, 1998.
 Lynch, M. (2004). *True to Life*. Cambridge, MA: MIT Press.

²⁶ Thanks to Kent Bach, David Chalmers, David Dowty, Michael Glanzberg, John Greco, John Hawthorne, Carrie Jenkins, Jeff King, Jon Kvanvig, Ernie Lepore, Peter Ludlow, Michael Lynch, John MacFarlane, Alan Millar, Julien Murzi, Duncan Pritchard, Nathan Salmon, Joe Salerno, Jonathan Schaffer, Jason Stanley, and the audience at the Epistemic Value Conference in Stirling for discussion of these and related issues. Special thanks to two anonymous referees for OUP, Mikkel Gerken, and Adrian Haddock for written comments that much improved the paper and to Adrian Haddock, Alan Millar, and Duncan Pritchard for organizing a great conference.

- (2009 this volume). 'The Values of Truth and the Truth of Values,' 225–42.
- (manuscript). 'Expressivism and the Value of Truth.'
- MacFarlane, John (2003). 'Future Contingents and Relative Truth,' *Philosophical Quarterly*, 53: 321–36.
- (2005a). 'Making Sense of Relative Truth,' *Proceedings of the Aristotelian Society*, 105: 321–39.
- (2005b). 'The Assessment Sensitivity of Knowledge Attributions,' in T. Szabó Gendler and J. Hawthorne (eds.), *Oxford Studies in Epistemology*, 1. Oxford: Oxford University Press.
- (2007) 'Semantic Minimalism and Nonindexical Contextualism,' in G. Preyer and G. Peter (eds.), *Content and Context: Essays on Semantics and Pragmatics*. Oxford: Oxford University Press.
- (2008). 'NonIndexical Contextualism', in Brogaard (2008d).
- Prior, A. N. (1957). *Time and Modality*. Oxford: Oxford University Press.
- (1967). *Past, Present and Future*. Oxford: Clarendon Press.
- (1968a). *Papers on Time and Tense*. Oxford: Oxford University Press.
- (1968b). 'Now,' *Nous*, 2: 101–19.
- Recanati, F. (2000). *Oratio Obliqua, Oratio Recta*. Cambridge, MA: MIT Press.
- (2004). 'Indexicality and Context-shift,' Conference Paper, Workshop on Indexicals, Speech Acts, and Logophors. Harvard University.
- Richard, M. (2004). 'Contextualism and Relativism,' *Philosophical Studies*, 119: 215–42.
- Riggs, W. (manuscript). 'Insight, Open-Mindedness and Understanding.'
- Salmon, N. (1986). *Frege's Puzzle*. Cambridge, MA : MIT Press/Bradford Books.
- (1998). 'Tense and Singular Propositions', in J. Almog, J. Perry, and H. Wettstein (eds.), *Themes from Kaplan*. New York: Oxford University Press.
- Stanley, J. (1997a). 'Names and Rigid Designation,' in B. Hale and C. Wright (eds.), *A Companion to the Philosophy of Language*. Oxford: Blackwell Publishers.
- (1997b). 'Rigidity and Content,' in R. Heck (ed.), *Logic, Language and Reality: Essays in Honour of Michael Dummett*. Oxford: Oxford University Press.
- (2005). 'Against Knowledge Relativism', <http://pwp.netcabo.pt/0154943702/Handout2.pdf>. This is reprinted (with some revisions) as ch. 7 of his recent book *Knowledge and Practical Interests* (Oxford: Oxford University Press, 2005).
- Weatherson, B. (2008). 'Conditionals and Indexical Relativism', in Brogaard (2008d).
- Egan, A., and Hawthorne, J. (2004). 'Epistemic Modals in Context,' in G. Preyer and G. Peter (eds.), *Contextualism in Philosophy: Knowledge, Meaning, and Truth*. Oxford: Oxford University Press.

This page intentionally left blank

APPENDIX

SYMPOSIUM ON JONATHAN
KVANVIG'S

The Value of Knowledge and the Pursuit of Understanding

This page intentionally left blank

APPENDIX A

Precis of *The Value of Knowledge and the Pursuit of Understanding*

Jonathan Kvanvig

Reflection on the issues surrounding the value of knowledge and other cognitive states of interest to epistemologists can be traced to the conversation between Socrates and Meno in Plato's dialogue named after the latter. The context of discussion concerns the hiring of a guide to get one to Larissa, and the proposal on the table is that one would want a guide who knows the way. Socrates sees a problem, however, for it is not clear why a guide with merely true opinion will not be just as good.

Meno has difficulty understanding the counterexample, but once he sees it, perplexity ensues. Meno wonders why we prize knowledge more than true opinion, and indeed whether there is any difference between the two.

Socrates responds in two ways. First, in spite of his reticence to claim to know anything, he claims to know that knowledge is different from true opinion. He then suggests, noting that he is only speaking in metaphors, that true opinions are like the statues of Daedalus, which wander off if not tethered. Knowledge, according to Socrates, is tethered, and it is this feature that metaphorically points to its difference from true opinion and why it is prized more than true opinion.

There are two important lessons that can be gleaned from this conversation. The first concerns the interplay between questions about the value of knowledge and questions about the nature of knowledge. Our initial judgment agrees with Meno that knowledge is important, and if we transpose the discussion to a contemporary key that identifies components of knowledge, Meno's question becomes the more general question of how knowledge has value above that of any of its proper parts. The second point is that addressing the question of the value of knowledge can be in tension with an account of the nature of knowledge. As Meno noticed, if our account of knowledge identifies it with true opinion, then our initial judgment about the value of knowledge is mistaken: knowledge shouldn't be prized more than true opinion.

These two lessons land us on a train headed for a wreck, but one quite a distance off. If we identify the components of knowledge in terms of true belief plus an epistemic condition such as justification and a clause to handle the Gettier problem (an assumption that should not be confused with the stronger claim that these components provide the elements needed for an analysis of knowledge), we can address Meno's question in the contemporary key by asking how and whether knowledge is more valuable than these components.

The first part of the journey thus begins by asking whether and how knowledge is more valuable than true belief alone, and then proceeds to ask about the relationship between the value of knowledge and the value of justified true belief. A natural answer to the former issue appeals to the latter: knowledge is more valuable because it includes the satisfaction of a normative standard that true belief alone does not. So long as justified true belief is more valuable than true belief alone, we can explain the value of knowledge over that of true belief in terms of this inequality.

Defending the idea that justified true belief is more valuable than mere true belief, however, is a bit more difficult than it might initially seem. The problem here is that it is easy to confuse two distinct claims. The first claim is that justified belief is more valuable than unjustified belief, and the second claim is that justified true belief is more valuable than unjustified true belief. The former inequality is nearly trivially true on any normative conception of justification, but the latter claim is far from trivially true. The problem that shows that the latter claim is far from trivially true is the Swamping Problem. The simplest theory of justification that gives rise to the problem is one that identifies justification with objective likelihood of truth. Beliefs that are likely to be true are certainly preferable to beliefs that are unlikely to be true, but true beliefs that are likely to be true have no special axiological advantage over true beliefs that are unlikely to be true. There are many analogies that have been developed to make this point, and a very simple one concerns two good espressos that are duplicates of each other but which come from different machines, one likely to produce good espressos and the other quite unlikely to do so. Information about the quality of the machine is important prior to seeing the outcome, but once the outcome is given its importance swamps the importance of the prior information. Just so, once the truth of a belief is a given, the importance of likelihood of truth is swamped by the presence of truth itself.

What follows from the swamping problem, then, is that not just any theory of justification is capable of sustaining the claim that the value of justified true belief is greater than the value of true belief alone. I argue that the theory of justification most susceptible to the swamping problem is ordinary process reliabilism, but that two approaches to justification have resources for avoiding the problem. One view is a quite subjective version of internalism, according to which justification is an intentional means to the goal of getting to the truth and avoiding error, rather than an effective means. The other view is a version of virtue epistemology, according to which justification is a special case of reliability, a special case involving a display of an intellectual virtue.

The train wreck for the ordinary assumption that knowledge is more valuable than its proper parts occurs, however, when we consider the Gettier problem. It has long been noticed that an attempt to analyze knowledge in terms of justified true belief plus some condition meant to address the Gettier problem results in an account that is gerrymandered or ad hoc or both. It is this feature that causes the wreck, for what it shows is that the closer an account gets to rightly dividing cases into knowledge and non-knowledge groups, the fewer resources the account will have to explain why we should value un-Gettiered justified true belief over mere Gettiered justified true belief. A survey of approaches to the Gettier problem reveals that the better they are at addressing the value issue (by being simple and intuitive approaches to the problem, the best

example of which is the early suggestion that knowledge requires reasoning through no falsehoods), the worse they are at avoiding counterexamples; and the better they are at avoiding counterexamples, the worse they are at including features that have any hope of being value-enhancing.

We thus reach the pessimistic conclusion that our ordinary assumption that knowledge is more valuable than its subparts is simply mistaken. This conclusion should not be confused, of course, with the stronger and obviously false claim that knowledge has no value or that it is unimportant. The pessimistic conclusion is not so outlandish as this, but is rather the claim that knowledge does not have a unique and special value over and above its proper subparts. In particular, it is not more valuable than justified true belief.

All is not lost, however, for there is an additional ordinary assumption relevant here beyond the assumption that knowledge has such special value. This additional ordinary assumption is contained in the penetrating questions near the end of the opening stanza of T. S. Eliot's 'Choruses from *The Rock*,' where he asks, 'Where is the Life we have lost in living? Where is the wisdom we have lost in knowledge? Where is the knowledge we have lost in information?' The idea underlying these questions is that there may be cognitive achievements that have value beyond whatever value we find in true belief (information) or knowledge. One such achievement, I argue, is that of understanding. I argue that understanding has a special and unique value that exceeds the value of its subparts, and thus that it provides a solution to the value problem first discussed in the *Meno*. To have such a solution, an account of understanding is needed and a defense of the claim that the value of understanding exceeds the value of its proper parts.

The kind of understanding that has these features, I argue, is objectual understanding, the kind of understanding that one has of physics or one's spouse or some geographical region such as small-town Texas. Such understanding is composed of beliefs which, for the most part, must be true. Errors in matters of detail can be tolerated, but such understanding does not survive large amounts of error or error about matters central to the thing in question. Moreover, the beliefs in question must be justified in order to constitute objectual understanding. I argue, however, that there is no analogue for the Gettier problem regarding objectual understanding.

The argument for this latter claim comes from consideration of cases in which it is plausible to attribute understanding but which constitute classic Gettier cases. For example, one can have historical understanding of a particular period of time and place partly as a result of having one's dyslexia correct fortuitously for errors in one's sources. One's understanding is displayed by the capacity to answer correctly (from information one possesses) any question about the period in question, including explanatory questions about why events happened in that period in the way they did. Such cases fit the model for classic Gettier cases, but do not threaten objectual understanding. About such understanding, we might find quite a bit of luck in the existence of such understanding, but such luck does not undermine understanding as it undermines knowledge.

The result of this value-driven investigation, then, is a change in focus for epistemology. Instead of following a tradition which is in large part driven by the presence of the skeptic, thereby focusing nearly exclusively on the nature of knowledge, we have an argument for focusing attention on cognitive achievements themselves, which may or may not count

as knowledge or as components of knowledge. The argument is not that epistemologists should ignore knowledge, for such a suggestion is undermined by the fact that knowledge is important. It is rather that there is no good reason for a singular focus on knowledge and its parts, and good reason to focus on other cognitive achievements as well. In the process, we may develop an understanding of aspects of cognition that deserve pursuit even more than knowledge does.

APPENDIX B

The Value Problem

John Greco

In Plato's *Meno*, Socrates raises a question about the value of knowledge. Why is knowledge valuable? Or perhaps better, What is it that makes knowledge valuable? Jonathan Kvanvig (2003) argues that this question is as important to epistemology as Socrates' question about the nature of knowledge, or what knowledge is. Any adequate epistemology must answer both the nature question and the value question. In fact, Kvanvig argues, the two questions interact: If a theory of knowledge does a poor job answering the value question, then that counts against its answer to the nature question. Likewise, if a theory does a good job explaining the value of knowledge, that counts in favor of its answer to the nature question. This seems exactly right. Put another way, the value question is at the heart of the project of explanation. The task of explaining what knowledge is involves the task of explaining why knowledge is valuable.

Before looking more closely at Plato's question in the *Meno*, however, we should distinguish it from a different question: Why is the *concept* of knowledge valuable? A plausible answer to that question is that the concept plays valuable roles in the lives of information-using, information-sharing beings such as ourselves.¹ The human form of life demands good information and the reliable flow of that information. The concept of knowledge, along with related concepts, serves those needs. That is not yet to say, however, why knowledge is valuable. We may put things this way: The concept of knowledge is valuable because it allows us to identify and share reliable information. But why is *knowledge* valuable? That question remains to be answered.

1. WHY IS THERE A PROBLEM?

We have seen that Socrates raises a question, but why is answering the question a problem? Why can't we say, for example, that knowledge is a kind of information, and that knowledge is valuable because information is valuable? Here I am understanding the concept of information to be 'factive': to be information is to be true information. The present suggestion, then, is that knowledge is valuable because true information is valuable. For example, true information has practical value—it helps us to get things that we want.

Socrates rejects this kind of answer because we think that knowledge is more valuable than *mere* true information, or true information that is not knowledge. In an often cited

¹ This is the answer defended by Craig (1999).

passage from the *Meno*, Socrates points out that mere true belief seems to have the same practical value as knowledge—the man who truly believes that the road leads to Larissa is as well served as the man who knows that it does. The problem then is this: We think that knowledge has value over and above its practical value as useful information. How do we explain that extra value? This is something that a good theory of knowledge should do.

2. A SPECIAL PROBLEM FOR RELIABILISM?

The value question is ancient, and as we have described it, it is a problem for any theory of knowledge whatsoever. Recently, however, Linda Zagzebski has argued that the problem is especially difficult for reliabilism.² In her terminology, reliabilism can't explain the added value that knowledge has over true belief. This is because reliabilists conceive the difference between true belief that is knowledge and true belief that falls short of knowledge as a difference in the reliability of the source. But the reliability of a source, Zagzebski argues, cannot add value to its product. To make the point she draws an analogy to good espresso. The value of a good cup of espresso is not increased by the fact that it was made by a reliable espresso machine. Good espresso is valuable, and reliable espresso machines are valuable. But the value of the espresso is not increased by the value of its source. Consider: a cup of espresso with the same intrinsic qualities, but made by an unreliable machine, would have exactly the same value. The conclusion that Zagzebski draws is that simple reliabilism cannot solve the value problem for knowledge. If knowledge is more valuable than mere true belief, its value must be explained in some other way.

3. A PROBLEM FOR EVERYONE

Kvanvig, however, argues that the value problem is even more intractable. According to Kvanvig, 'We are left . . . with no decent answer to the question of the value of knowledge. . . . [T]here is no good answer to the problem of the *Meno*' (Kvanvig 2003: 184). Kvanvig's conclusion, it should be noted, depends on particular ways in which he conceives the problem. We have already seen that the question of the value of knowledge has a tendency to shift. For example, it shifts in Plato's discussion from 'Why is knowledge valuable?' to 'Why is knowledge more valuable than mere true belief?' We will see that the question shifts again, and a number of times, in Kvanvig's discussion as well. For example, at times Kvanvig wants an explanation of why knowledge is more valuable than any of its proper parts. Even more strongly, he sometimes asks why knowledge is more valuable than the sum of its parts. Clearly, each question in this series requires an increasingly demanding solution to the value problem.

I will defend two theses. The first is that some of Kvanvig's formulations of the value problem are too demanding. That is, we should not expect an answer to each of his

² See Zagzebski (1996; 2000).

questions, since we should not expect that knowledge really is valuable in the ways that his questions suppose. For example, and most obviously, there is no pre-theoretic reason for thinking that knowledge is more valuable than the sum of its parts. The second thesis I will defend is that, nevertheless, a virtue-theoretic account of knowledge answers all of Kvanvig's questions. In other words, a virtue-theoretic approach solves each of his formulations of the value problem. Putting these two theses together, we get the following result: A virtue theory gives us a better answer to the value problem than anyone should expect!

4. SHIFTING QUESTIONS

Consider again Kvanvig's two criteria for an adequate theory of knowledge: any such theory, he says, ought to give an account both of the nature of knowledge and the value of knowledge. Suppose that the question about the value of knowledge is the very general one: Why is knowledge valuable? The idea that a theory of knowledge ought to answer this question seems right, and precisely because we are confident that knowledge is indeed valuable. It is a reasonable criterion of success, therefore, that a theory preserves this pre-theoretical data and explains it.

As Kvanvig notes, however, Socrates' question soon shifts to a more specific one: Why is knowledge more valuable than true opinion? But once again, the demand for an answer seems appropriate. And once again, this is because we are pre-theoretically confident that knowledge has such a value. As Kvanvig writes,

part of the challenge of explaining the value of knowledge is in explaining how it has more value than other things, one of these other things being true opinion—as Meno claims after acquiescing to Socrates' point that true belief is every bit as useful as knowledge. . . . Meno expresses here a common presupposition about knowledge, one that is widely, if not universally, shared. Given this presupposition, an account of the value of knowledge must explain more than how knowledge is valuable. It must also explain why the value of knowledge is superior to the value of true opinion. (Kvanvig 2003: 3–4)

But now consider a second shift in Kvanvig's question about the value of knowledge. At another point in the book Kvanvig writes,

To explain the value of knowledge in a way that satisfies the constraints of the *Meno* requires showing that knowledge is more valuable than any proper subset of its constituents. (Kvanvig 2003: 107)

This new and stronger demand is illustrated by what Kvanvig says about a widespread position in epistemology—that knowledge is composed of justified true belief, plus some further condition to handle Gettier cases.

An adequate account of the value of knowledge must explain why it is more valuable than any subset of its constituents. If we assume that there is some property like justification that distinguishes knowledge from true belief, then an adequate explanation of the value of knowledge could be achieved by giving an adequate account of the value of justification. Because knowledge is more than justified true belief, such an explanation is only one part of a complete explanation. In addition, what is needed is an explanation of why knowledge is more valuable than justified true belief. (Kvanvig 2003: 112)

Finally, consider one more passage along the same lines.

The conclusion to which our investigation seems to be pointing is that ordinary thinking about knowledge is mistaken, that knowledge does not have the kind of value it is ordinarily thought to have. In particular, we seem to be heading for the conclusion that knowledge does not have a value that exceeds that of subsets of its constituents. (Kvanvig 2003: 157)

Is it true what Kvanvig is implying here? In other words, is it true that 'ordinary thinking' assumes that knowledge is more valuable than any subset of its constituents? That is far from clear. To illustrate, suppose that some JTB+ account of knowledge is correct—that knowledge is justified true belief plus something further to handle Gettier problems. Suppose also that the 'plus' part amounts to some minor, technical adjustment to the traditional idea that knowledge is justified true belief, and that this further condition adds no further value to knowledge over justified true belief. Would these suppositions conflict with ordinary thought? Would this show 'that knowledge does not have the kind of value it is ordinarily thought to have'? It is hard to see how that could be the case. Most people are not at all aware of Gettier problems, and we can suppose that almost no one was before 1963. But then how could ordinary thought include the idea that knowledge is more valuable than justified true belief?

Kvanvig might reply that ordinary thinking does include the idea implicitly. For once ordinary thinkers are exposed to Gettier problems, they quickly agree that knowledge is not equivalent to justified true belief. But that reply does not address the issue at hand, which is whether ordinary thinkers assume that knowledge is more *valuable* than justified true belief. Put more generally, the issue is whether ordinary thinkers assume that knowledge is more valuable than any subset of its constituents, whether this is understood in terms of the JTB+ account or in terms of some other account. And this is what I am saying is dubious. Ordinary thinking, I am suggesting, contains no convictions about that issue.

And now the point is this: If there is no pre-theoretical conviction that knowledge is more valuable than any subset of its constituents, then Kvanvig has placed an inappropriate demand on a solution to the value problem. The original demand to explain the value of knowledge was grounded in an ordinary and widespread conviction that knowledge is indeed valuable. That is why we accepted it as an appropriate criterion for an adequate account of knowledge. When the question about value shifts in the way that it has by this point in Kvanvig's discussion, the stronger demand on an explanation that emerges is not so grounded.

To sum up, there is no pre-theoretical conviction that knowledge is more valuable than any subset of its constituents, and therefore it is not appropriate to require that a theory of knowledge explain why knowledge has that sort of value. Kvanvig's criterion for an adequate solution to the value problem is too strong.

That being said, in some places Kvanvig seems to place even stronger demands on a solution to the value problem. Specifically, he seems to demand that a solution explain how the value of knowledge exceeds the value *of all its parts together*. Consider the following passages.

a satisfactory answer to the question of the value of knowledge will need to explain why knowledge is, by its very nature, more valuable than its parts. (Kvanvig 2003: p. xiv)

Thus, I will be arguing that knowledge is valuable, but that it fails to have a value exceeding that of its parts, thereby leaving us with no adequate answer to the problem of the value of knowledge first posed by Plato in the *Meno*. (Kvanvig 2003: pp. xv–xvi)

On the assumption we have been making in the past several chapters (that the value of knowledge is in some way a function of the value of its parts), the need to account for both the nature and value of knowledge requires that we identify a fourth condition that not only yields a counterexample-free account of knowledge but also provides some basis for explaining the value of knowledge over the value of its constituents. (Kvanvig 2003: 116)

There are aspects of Kvanvig's discussion that suggest that these latest passages are slips—that his considered position is to require only an explanation of why knowledge is more valuable than its *proper* parts. Nevertheless, it will be useful to explore the cogency of this newest demand on a solution to the value problem.

Someone might think that the demand is not even coherent. Consider, for example, an account of bachelorhood. Let us say that a bachelor is an unmarried male who is eligible for marriage. Now suppose we wanted a solution to 'the value problem for bachelorhood.' That is, suppose we want an account of why bachelorhood is valuable. Is it appropriate to demand that such an account explain why the value of bachelorhood exceeds the value of all its parts together? It can seem that this demand is not even coherent, for being a bachelor just *is* being an unmarried male who is eligible for marriage. But then how could the value of the former exceed the value of the latter?

We can take the question another way, however. For we can ask whether the value of being a bachelor exceeds the sum of the value of the parts. That is, we can ask whether the value of bachelorhood exceeds the value of being unmarried plus the value of being male plus the value of being eligible for marriage. We might answer in the negative, but the question is at least coherent. That is, at least sometimes the whole is worth more than the parts taken separately.

That is the way that we should understand the latest formulation of the value problem. It asks whether the value of knowledge exceeds the value of each of its parts taken alone. This might be the case if knowledge were some sort of organic whole, in the sense that its parts are organized in a particular way. Perhaps some relation among the parts adds value beyond that which the parts have of themselves.

The question is coherent, I take it, but there is no pre-theoretical reason for thinking that it should get a positive answer. And as such, it would be inappropriate to place such a demand on a solution to the value problem. Put another way, there is no pre-theoretical reason to think that knowledge is more valuable than its parts taken separately, and therefore no reason to expect that a solution to the value problem will explain why it is.

5. A SOLUTION TO THE VALUE PROBLEM

There is no reason to expect such an explanation. However, I want to argue, a virtue-theoretic approach to the value problem gives us one. In this section I will articulate a solution to the value problem that is consistent with an account of knowledge that

I have defended elsewhere.³ In fact, we will see, the solution falls out of the account straightforwardly. In the next section I will argue that the account satisfies all of Kvanvig's demands for an adequate solution—even the ones that are unreasonable.

According to the account I have in mind, knowledge is a kind of success through virtue. Put another way, knowledge is a kind of success through virtuous agency. The intellectual virtues that give rise to knowledge are best understood as intellectual abilities, and therefore knowledge is a kind of success through one's own abilities. This sort of success can be juxtaposed to mere lucky success: When S has knowledge, S gets things right as the result of her own abilities, as opposed to getting things right as the result of blind chance or dumb luck, or something else. Put yet another way, in cases of knowledge S gets things right *because* she is intellectually able and because she has exercised her abilities.

But now an answer to the value problem falls out of this account straight away. In the *Nicomachean Ethics* Aristotle makes a distinction between (a) achieving some end by luck or accident, and (b) achieving the end through the exercise of one's abilities (or virtues). It is only the latter kind of action, Aristotle argues, that is both intrinsically valuable and constitutive of human flourishing. 'Human good,' he writes, 'turns out to be activity of soul exhibiting excellence' (*Nicomachean Ethics*, § 1.7). In this discussion Aristotle is clearly concerned with intellectual virtue as well as moral virtue: his position is that the successful exercise of one's intellectual virtues is both intrinsically good and constitutive of human flourishing.

If this is correct, then there is a clear difference in value between knowledge and mere true belief. In cases of knowledge we achieve the truth through the exercise of our own intellectual abilities, which are a kind of intellectual virtue. Moreover, we can extend the point to include other kinds of intellectual virtue as well. It is plausible, for example, that the successful exercise of intellectual courage is also intrinsically good, and also constitutive of the best intellectual life. And of course there is a long tradition that says the same about wisdom and the same about understanding. On the view that results, there is a plurality of intellectual virtues, and their successful exercise gives rise to a plurality of epistemic goods. The best intellectual life—intellectual flourishing, so to speak—is rich with all of these.

6. KVANVIG'S DEMANDS

We saw that Kvanvig places a series of increasing demands on an adequate solution to the value problem. These demands can be understood in terms of the series of questions articulated above. In effect, each demand is a requirement that one of the following questions be answered:

- A. Why is knowledge valuable?
- B. Why is knowledge more valuable than true belief?
- C. Why is knowledge more valuable than any subset of its constituents?
- D. Why is knowledge more valuable than the value of all its parts taken separately?

³ e.g. in Greco (2003). See also Sosa (1988; 1991; 2003; 2007), Zagzebski (1996; 1999), and Riggs (2002).

Finally, we may note that each question in the series, and each requirement that the question be answered, involves a supposition: that knowledge *is* valuable in the way to be explained.

We may now see that the solution proposed respects all the suppositions and answers all the questions. The answer to questions A and B is straightforward: knowledge is a kind of success through virtue, and in general success through virtue is both intrinsically valuable and constitutive of human flourishing, which is also intrinsically valuable. Therefore, knowledge has value over and above the practical value of true belief.

The proposed solution answers question C as well. Knowledge is a kind of success through virtue. And in general, success through virtue is more valuable than either success without virtue or virtue without success. In particular, virtuously produced true belief is more valuable than both true belief that is not virtuous and virtuous belief that is not true. Neither subset is intrinsically valuable, or constitutive of what is intrinsically valuable, in just the way that knowledge is.

Finally, the proposed solution answers even question D, respecting the supposition that knowledge is more valuable than all of its parts taken together. This is because success through virtue is more valuable than an act that is both successful and virtuous, but not successful because virtuous. Suppose, for example, that an athlete runs a race in a way that is clearly an exercise of her athletic excellence. Suppose also that she wins, but only because the other runners, some of who are equally excellent, get sick before the race. Or suppose that she wins, but only because the other runners were bribed. Clearly, neither sort of win is as valuable as it could be. What one really values as an athlete is to win as the result of ability.

Likewise in the case of intellectual virtue. One's belief can be virtuously formed and true, but not true because virtuously formed. This is just the structure of Gettier cases, where true and virtuous belief falls short of knowledge.

It is apparent, then, that a virtue-theoretic approach to the value problem meets all the demands that Kvanvig requires of an adequate solution. So why is Kvanvig dissatisfied with this kind of answer? I can only speculate that it is because he misses the force of the proposal. For example, when he first introduces the virtue-theoretic approach he writes,

Recently, several epistemologists have proposed such an idea, to the effect that credit accrues to the agent who has intellectually virtuous beliefs. . . . All three share a common theme about the value of the virtues, for they think of this value in terms of some kind of credit due to the agent whose belief is virtue-based. (Kvanvig 2003: 81)

Just as actions that result from virtues yield credit for the actor, beliefs resulting from faculties that count as virtues generate credit for the believer. (Kvanvig 2003: 82)

Notice that there is no mention here of *success* through virtue. The central idea of the proposal, which is that knowledge is valuable because it is a kind of success through ability or virtue, is entirely absent.

Even when Kvanvig talks about true virtuous belief, he seems to miss the distinction between (a) a belief's being true and virtuously formed, and (b) a belief's being true because virtuously formed. At places Kvanvig does seem aware of the distinction. For example, he quotes Ernest Sosa (2003) as follows:

The grasping of the truth central to truth-connected reliabilist epistemology is not just the truth that may be visited upon our beliefs by happenstance or external agency. We desire rather truth gained

through our own performances, and this seems a reflectively defensible desire for a good preferable not just extrinsically but intrinsically. What we prefer is the deed of true believing, where not only the believing but also its truth is attributable to the agent as his or her own doing. (Kvanvig 2003: 95)

But when Kvanvig describes the view in question, he says this:

The basic idea of a virtue approach to the value of knowledge over that of its subparts is that there is a special value for beliefs that arise out of intellectual virtue. When true belief is a product of the virtues, the claim is that there is epistemic credit due to the agent in question and hence that virtuous true belief is more valuable than true belief. (Kvanvig 2003: 106)

In the first sentence he refers to the special value of 'beliefs that arise out of intellectual virtue,' whereas in the second sentence he says that credit is due when 'true belief is a product of the virtues.' It isn't clear that he sees the importance of the distinction, and so it isn't clear whether the phrase 'virtuous true belief' in the second sentence means (a) belief that is both true and virtuous, or (b) belief that is true because virtuous.

In the following passage, however, it seems clear that Kvanvig is missing the distinction.

It is equally true, however, that knowledge is more than intellectually true virtuous belief. Goldman's fake barn case discussed earlier is a well-known example that reveals a difference between knowledge and such virtuous belief, for impressive perceptual abilities count as intellectual virtues and could be displayed in the fake barn case. The reason the display of such virtues falls short of knowledge is that perception can be an impressive ability and still be unable to distinguish real barns from well-designed fake ones, and so a true belief could result that still was only accidentally true. (Kvanvig 2003: 107)

Kvanvig's claim in the last sentence is problematic for independent reasons. Specifically, it is unclear how perception can count as an ability relative to S's environment and yet be unable to distinguish real barns from fake ones. But suppose we agree for the sake of argument that S's belief is not only true but also formed from a virtue. Still, the belief is not true *because* it is formed from a virtue. Put more carefully, the belief's being so formed does not explain why S has a true belief rather than a false belief. On the contrary, S believes the truth because she happens (luckily) to be looking at the one real barn in the area. If she had been looking anywhere else nearby, excellent perception or no, she would have a false belief.

The fake barn case looks to be a counterexample, then, only if we ignore the distinction between (a) belief that is true and virtuous, and (b) belief that is true because virtuous. But as we have seen, it is just this distinction that is crucial to the proposed solution.

REFERENCES

- Craig, E. (1999). *Knowledge and the State of Nature: An Essay in Conceptual Synthesis*. New York: Clarendon Press.
- Greco, J. (2003). 'Knowledge as Credit for True Belief,' in M. DePaul and L. Zagzebski (eds.), *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Oxford University Press.
- Kvanvig, J. (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge, MA: Cambridge University Press.

- Riggs, W. (2002). 'Reliability and the Value of Knowledge,' *Philosophy and Phenomenological Research*, 64: 79–96.
- Sosa, E. (1988). 'Beyond Skepticism to the Best of our Knowledge', *Mind*, 97: 153–89.
- (1991). *Knowledge in Perspective*. Cambridge, MA: Cambridge University Press.
- (2003). 'The Place of Truth in Epistemology,' in M. DePaul and L. Zagzebski (eds.), *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Oxford University Press.
- (2007). *Apt Belief and Reflective Knowledge*. Oxford: Oxford University Press.
- Zagzebski, L. (1996). *Virtues of the Mind*. Cambridge, MA: Cambridge University Press.
- (1999). 'What is Knowledge?' in J. Greco and E. Sosa (eds.), *The Blackwell Guide to Epistemology*. Oxford: Blackwell Publishing.
- (2000). 'From Reliabilism to Virtue Epistemology,' in G. Axtell (ed.), *Knowledge, Belief and Character*. Lanham, MD: Rowman and Littlefield.

APPENDIX C

Is Understanding Factive?

Catherine Z. Elgin

In *The Value of Knowledge and the Pursuit of Understanding*, Jonathan Kvanvig (2003) sets out to do two things. First, he seeks to show that there is no feasible way to account simultaneously for the nature of knowledge and for the value of knowledge. Second, he argues that understanding should be a central epistemological concern. In his book these two contentions are connected. At least one reason why epistemology should take understanding more seriously, Kvanvig believes, is that it cannot adequately explain what makes knowledge valuable. I do not plan to say anything about the argument concerning knowledge, for I think that the positive reasons for treating understanding as central stand on their own. We should treat understanding as epistemologically central because, if we do not, we fail to do justice to important aspects of cognition. The bulk of my paper will be devoted to a disagreement with Kvanvig about what the proper scope of epistemology should be. But before I turn to that, I want to say something about areas in which we agree.

The term ‘understanding’, as Kvanvig rightly points out, is used in a variety of ways. Some are irrelevant to epistemology. I can say ‘I understand’ to hedge an assertion or to moderate its force. ‘I understand that you are angry with me’ may be a mild overture that gives you space to politely demur. This is a moderating use. Or I might say ‘I understand that you are angry with me’ when I am not certain that you are angry, but have some reason to think so. Then ‘I understand’ serves as a backing away from a full-fledged claim to epistemic entitlement. This is hedging. These are not the sorts of usages that Kvanvig and I are interested in. We are concerned with cases where understanding is a sort of epistemic success. So for the remainder of this paper I shall restrict the term ‘understanding’ to the sort of understanding that should be of interest to epistemology—the sort that manifests epistemic success. In such cases the understander has a claim to epistemic entitlement. Two questions arise: what is the bearer of the entitlement and what is the claim to it?

There are two obvious candidates for the bearer of epistemic entitlement: individual propositions and more comprehensive bodies of information. I can say, ‘I understand that the Comanches dominated the southern plains of North America in the eighteenth century.’ Or I can say, ‘I understand the power relations among the tribes in the southern plains in the eighteenth century.’ If the primary unit of understanding is the proposition, then the difference between knowledge and understanding seems slight. If the proposition ‘I understand that the Comanches dominated the southern plains’ is supposed to be a stand-alone proposition (and is not supposed to be a hedge), it is hard to see how it differs from ‘I know that the Comanches dominated the southern plains.’ But if my understanding that the Comanches dominated the southern plains depends in a suitable

way on my overall understanding of the Comanches, or of the political forces and power relationships in the North American plains in the eighteenth century, or some such thing, then the situation is different. The epistemological standing of 'The Comanches dominated the southern plains' then derives from its place in a more comprehensive general understanding of the history of North America (or some part of it). That is, the proposition derives its epistemological status from a unified, integrated, coherent body of information. This is the conception of understanding that Kvanvig takes to be central. I agree.

Understanding, then, is in the first instance a cognitive relation to comprehensive, coherent sets of cognitive commitments. The understanding individual propositions express derives from an understanding of larger bodies of information they belong to. I understand that the Comanches dominated the southern plains because I grasp how that proposition fits into and is justified by reference to a more comprehensive understanding that embeds it.

As Kvanvig rightly insists, to understand the Comanches' dominance of the southern plains involves more than knowing the various truths that belong to a comprehensive, coherent account of the matter. The understander must also grasp how the various truths relate to each other. This is an important point. One might think that the comprehensive body of information is just a large collection of propositions. I suggest that understanding involves more. The understander should be able (and perhaps be aware that she is able) to use that information—for example, to reason with it, to apply it, to perhaps use it as a source of working hypotheses about other related matters. Someone who knows geometry, for example, knows all the axioms, all the major theorems and their proofs. You can acquire such knowledge by rote. Simply memorize the axioms and proofs and you have it. But someone who understands geometry can reason geometrically about new problems, apply geometrical insights in different areas, assess the limits of geometrical reasoning for the task at hand, and so forth. Understanding something like the Comanche dominance is obviously not exactly like understanding geometry. The applications and extensions are more tentative. The range to which insights can reasonably be applied is more restricted. The evidence for a successful application is empirical (and may be hard to come by). And so on. But in both cases understanding involves a *falicity* for using the information one has, not merely in an appreciation of how things are. Kvanvig does not discuss this aspect of understanding. But it is something he could easily assimilate into his account, either by saying that this is part of grasping or by saying that, in addition to grasping connections, an understander has to have an ability to use the information at his disposal.

According to Kvanvig, understanding consists of coherent bodies of (mostly) true beliefs. Coherence alone is not enough. A coherent body of beliefs that is largely false, such as astrology, does not constitute an understanding. Understanding is cognitively valuable on Kvanvig's view, because (a) true beliefs are good and (b) a grasp of the coherence of constellations of true beliefs affords subjective justification, which is also good. This is the basis for his conviction that understanding is *factive*.

Knowledge is *factive* in that it is impossible to know that *p* unless '*p*' is true. Kvanvig maintains that understanding is *factive* as well. But understanding concerns subject matters rather than individual propositions. So what it means to claim that understanding is *factive* is a bit harder to make out. Perhaps understanding is *factive* if it is impossible

to understand a subject—say, the history of the American Plains Indians—unless some identifiable, suitably comprehensive proposition is true. That proposition might be the long conjunction of all the shorter propositions that belong to the coherent body of information. (This parallels the interpretation of coherence theories of knowledge as requiring the truth of the conjunction of the propositions in the coherent system.) On such an account, understanding would be a sort of knowledge, namely the knowledge of long, subject-matter-connected, conjunctive propositions.

There are several problems with this proposal. The first is that it does not accommodate the requirement that the understander grasp the relations among the propositions—that the understander appreciate how they bear on one another. Although the body of information understood must be coherent, if the understander need only know the conjunction, there is no requirement that she grasp the coherence. Second, it does not accommodate the insight that the student who understands geometry can do more with it than the student who just knows all the axioms, the main theorems, and their derivations. Third, it does not accommodate the fact that not all of the propositions that comprise a genuine understanding of a subject need to be true. We would be inclined to say that a historian understood the Comanche dominance even if he harboured a few relatively minor false beliefs about the matter.

Kvanvig agrees. He does not believe that understanding a topic consists in believing a long conjunction. He does not insist that every proposition in the comprehensive body of information be true. Rather, he maintains, we do not understand a topic unless most of the propositions and all of the central propositions that constitute our coherent take on that topic are true. He allows that a few peripheral falsehoods might degrade one's understanding of a subject matter, but not destroy it. That understanding is *factive* in this sense is the thesis I want to dispute.

I do not deny that there is such understanding or that it should be of interest to epistemology. What I deny is that epistemology should limit itself to this sort of understanding. If epistemology is to accommodate science, I maintain, it needs a wider scope.

Antirealists contend that truth is not the goal of science. Hence fully successful scientific theories can be largely false. If understanding is *factive*, they must conclude that science does not yield understanding. Realists too should find Kvanvig's position uncongenial. Whatever may be the case at the end of inquiry, neither current nor previous science consists largely of truths, with a few relatively insignificant falsehoods at the periphery. Maybe, someday science will yield understanding. But if Kvanvig is right, it hasn't done so yet. This is implausible.

Unlike knowledge, understanding admits of degrees. A freshman has some understanding of the Comanche dominance, while her teaching fellow has a greater understanding and her professor has an even greater understanding. So epistemology should explain what such differences in degree consist in. Kvanvig recognizes two dimensions along which understanding can vary: breadth and depth. The professor might have a broader understanding of the Comanche, being able to embed his coherent body of true beliefs into a more comprehensive understanding of American history. He might also have a deeper understanding. In that case, his web of beliefs is more tightly woven. It contains more information. But both the student and the professor understand the Comanche dominance because they grasp coherent bodies of predominantly true propositions, and

believe the propositions that belong to those bodies. There is another dimension along which the student and the professor might differ. The professor and the student might weigh the facts differently. Even if both believe a given proposition, and both incorporate it into a coherent body of beliefs about the topic, the professor might consider it highly significant, while the student considers it just another fact about the Comanches. If the fact really is significant—if, for example, it is central to explaining why the Comanches allied with one tribe but were antagonistic to another—then the professor's better understanding consists in his appreciating the significance of the fact, not merely in his recognizing that it is a fact. Again, this is something that Kvanvig could easily concede.

However, there is another dimension along which we can measure degrees of understanding that Kvanvig cannot take on board. For it involves conceding that some bodies of information, even though they are not true, nonetheless display a measure of understanding. The growth of understanding often involves a trajectory from beliefs that, although strictly false, are in the right general neighbourhood to beliefs that are closer to the truth. The sequence may terminate in true beliefs. But, I contend, the earlier steps in the sequence should fall within the ambit of epistemology. An 8-year-old's understanding of human evolution might include as a central strand the proposition that human beings descended from apes. A more sophisticated understanding has it that human beings and the other great apes descended from a common hominid ancestor who was not strictly an ape. The child's opinion displays some understanding of evolution. It is clearly cognitively better than the belief that humans did not evolve. But it is not strictly true. And since it is central to her take on human evolution, it follows from Kvanvig's theory that her take on human evolution does not qualify as understanding. Epistemology need give no account of what makes the child's view of evolution cognitively valuable or more valuable than a view that takes humans to have evolved from butterflies. But the pattern exhibited in this case is endemic to scientific education. We routinely start with crude characterizations that properly orient us toward the phenomena, and then refine the characterizations as we become scientifically more sophisticated. Think of the trajectory from naïve folk physics through Newtonian mechanics to relativity and quantum mechanics.

When we construe such a take on a subject as understanding, Kvanvig believes, we use the term 'understanding' in an honorific sense, just as we use the term 'knowledge' in an honorific sense when we speak of 'the current state of scientific knowledge', while conceding that some of what belongs to the current state of scientific knowledge is false. Such honorific usages of epistemic terms are, he believes, extended usages that fall outside the scope of epistemology.

Perhaps we could accept Kvanvig's dismissal of such uses of 'understanding' as merely honorific if they applied only to young children and novices. I tend to think otherwise, however, for I think epistemology should have something to say about what makes the views of the child who thinks humans evolved from apes better than the views of a child who thinks humans did not evolve or evolved from butterflies. But the main problem with Kvanvig's contention that understanding is factive is that the pattern displayed by the student as he moves from the naïve view of human evolution up to the view held by the professor of evolutionary biology is the same pattern as science displays in the sequence of theories it develops.

A central tenet of Copernicus's theory is that the Earth travels around the sun in a circular orbit. Kepler improved on Copernicus by contending that the Earth's orbit is

not circular, but elliptical. With the abandonment of a commitment to absolute space, current astronomers can no longer say that the Earth travels around the sun *simpliciter*, but must talk about how the Earth and the sun move relative to each other. Despite the fact that Copernicus's central claim was strictly false, the theory it belongs to constitutes a major advance in understanding over the Ptolemaic theory it replaced. Kepler's theory is a further advance in understanding, and the current theory is yet a further advance. The advances are clearly cognitive advances. With each step in the sequence, we understand the motion of the planets better than we did before. But no one claims that science has as yet arrived at the truth about the motion of the planets. Should we say that the use of the term 'understanding' that applies to science should be of no interest to epistemology?

Again Kvanvig might contend that the use of 'understanding' here is honorific. We apply the term in these cases only because we think that the scientific advances are on the way to the truth—the comprehensive, general account of celestial motion that gets it right. In effect, current science borrows its epistemic status from its descendants. Wilfrid Sellars (1963) argued that in a mature science, later theories should show why their predecessors were right to the extent that they were. So the later theories should at least partially vindicate their predecessors. When this does not happen, we are apt to conclude that the earlier scientists did not understand the phenomena that their theory purported to explain. We do not, for example, consider phlogiston theorists to have had any understanding of combustion. Suppose we concede this point. Then in saying that the various astronomical theories embody an understanding, we take out a lien on the future of science. Still, I would urge, the cognitive achievements embodied in such theories should be a central concern for epistemology. Even if we do not yet have (and may never get to) the truth, we have made real cognitive progress. We understand the motions of celestial bodies better than our predecessors did. Epistemology should explain what makes current understanding better. If we say that the uses in question are honorific, epistemology should explain why certain attitudes toward certain subject matters are worthy of honour.

There is another aspect of science that is even more troublesome for Kvanvig's view. That is science's penchant for idealization. Science streamlines and simplifies. It devises and deploys simplified models that diverge from the phenomena it seeks to explain. The ideal gas law, for example, accounts for the behaviour of gases by characterizing the behaviour of a gas composed of dimensionless, spherical molecules that are not subject to friction and exhibit no intermolecular attraction. There is no such gas. Indeed, there could be no such gas. Nonetheless, scientists purport to understand the behaviour of actual gases by reference to the ideal gas law.

Scientists do not consider idealization an unfortunate expedient. Rather they value it as a powerful tool. They do not expect that in the fullness of time idealizations will be eliminated from scientific theories. So the 'promissory note-ishness' that we saw in talking about the progress in our understanding of celestial motion seems not to be in place here. Elimination of idealizations is not a desideratum. Nor is consigning them to the periphery of a theory. It is simply not the case that the bodies of information that constitute scientific understanding are, or that their ultimate successors can be expected to be, composed of truths, with any residual falsehoods only occurring at the periphery. The ideal gas law lies at the core of statistical mechanics, and some such law is likely to lie at the core of any successor to current theories.

I concede that many of the propositions that fall within the scope of 'the current state of scientific knowledge' are not strictly *knowledge* because they are not true. In ordinary usage we withdraw a claim to know a proposition if we discover that the proposition is false. So it is reasonable to construe 'knowledge' as a factive. If we are being scrupulous, we should probably not speak of the current state of scientific *knowledge* unless we are convinced that the propositions we are speaking of are true. But 'understanding' is different. Since 'understanding' applies to large, often somewhat inchoate bodies of information, it takes a direct object that is not a proposition. Ann understands the Comanche dominance of the plains. Pat understands the Krebs cycle. Eve understands the Peruvian tax code. And we typically acknowledge that people can have some measure of understanding even if the contentions making up the bodies of information they endorse diverge somewhat from the truth. So our ordinary use of 'understanding' as applied to bodies of information does not favour a factive analysis. Nevertheless, 'understanding' is some sort of a cognitive success term. If I am going to reject the factive analysis, I need some way to identify or characterize the cognitive success.

As a very crude first approximation, I suggest that understanding is a grasp of a comprehensive general body of information that is grounded in fact, is duly responsive to evidence, and enables non-trivial inference, argument, and perhaps action regarding that subject the information pertains to. Obviously this is hideously rough. Some of the roughness is inevitable, if understanding comprehends everything from the second grader's very shallow take on evolution to the mature scientist's broad, deep, textured grasp of the subject. But some of the roughness can be smoothed out with a bit more work.

Let us start by looking at scientific idealizations. These are both central and ineliminable. We understand the behaviour of actual gases by reference to the alleged behaviour of a so-called ideal gas. There is no such gas. So how can it figure in our understanding of the world? I suggest that effective idealizations are felicitous falsehoods. That they are false is evident. They are felicitous in that they afford epistemic access to matters of fact that are otherwise difficult or impossible to discern. Idealizations are fictions expressly designed to highlight subtle or elusive or otherwise nearly inaccessible matters of fact. They do so by exemplifying features they share with the facts.¹

To make this out requires saying a bit about exemplification. Let us look at a pedestrian example. Commercial paint companies supply sample cards that exemplify the colours of paints they sell. The cards contain instances of those colours, and refer to the colours they instantiate. Such cards have a large variety of other properties as well. They consist of sequences of splotches of colour, each captioned with a name or number. They are a few inches long, and perhaps an inch and a half wide. They make good bookmarks. They were manufactured somewhere, on some particular date, were shipped via some means. They are a certain distance from the Eiffel Tower. Most of the properties of the cards are utterly irrelevant to their function. Some nonfunctional elements facilitate but do not figure in the card's function. None of these properties is exemplified. To exemplify a property, an exemplar must both instantiate and refer to it. The function of the cards, in their standard use, is to display and hence afford epistemic access to the paint colours. By at once instantiating and making reference to the colours then, the cards perform their function.²

¹ See Elgin (2004).

² See Goodman (1968: 52–68) and Elgin (1996: 170–83).

Other samples and examples function in the same way. A blood sample exemplifies its antibodies. A sample problem worked out in a textbook exemplifies a reasoning strategy that the students are supposed to learn. Each sample highlights some of its own properties, makes them manifest, draws attention to them.

Exemplification is selective. An exemplar exemplifies only some of its properties. It brings those properties to the fore by marginalizing, downplaying, or overshadowing others. What a given exemplar exemplifies depends on how it functions. The paint sample cards could be used as bookmarks. In that case, they would exemplify their positions in books rather than the colours of the patches.

In principle an item can exemplify any property it literally or metaphorically instantiates. But doing so is not always easy. The Bhutan lizard's snout is a distinctive shade of greenish gold. But a paint company would be ill advised to recommend that potential customers look at the lizard's nose order to see that colour. Bhutan lizards are among the rarest, most elusive animals on Earth. We are unlikely ever to see one, and any glimpse we get of one is unlikely to make the colour manifest. We could not see it long enough or well enough, and would be unlikely to attend to the colour carefully enough to decide whether it was the shade we wanted to paint the den. It is far better to create a lasting, readily available, easily interpretable sample of the colour—one whose function is precisely to manifest the colour. Such a sample should be stable, accessible, and have no properties that distract from attention to the colour. Effective samples and examples are carefully contrived to exemplify particular features. Factors that might otherwise predominate are omitted, bracketed, or muted. If the property is at all subtle or difficult to discern, a good deal of stage setting may be required to bring it to the fore. Similarly in scientific cases. The conductivity of water is hard to determine in nature, because the liquid in lakes, puddles, rivers, and streams contains impurities. By eliminating the impurities in the lab, the scientist can contrive a sample of pure water, thus gaining epistemic access to the property she is interested in studying.

But if the cognitive contribution of an exemplar consists in the exemplification of select features, then anything that exemplifies exactly those features can, in a suitable context, make the same contribution. Return to the sample cards mentioned above. Like just about everyone else, I spoke of the cards as though they comprised paint samples, telling instances of the stuff you might use to paint the den. This is not true. The sample on the card does not consist of paint, but of an ink or dye of the same colour as the paint whose colour it exemplifies. If the sample were supposed to exemplify other properties of the paint, such as its durability or resistance to fading, this divergence would be objectionable. But since it purports only to exemplify the paint's colour, and is in fact the same colour as the paint, the divergence is unproblematic. The card affords epistemic access to the property we want epistemic access to.

I suggest that idealizations in science function similarly. The ideal gas is a fiction that exemplifies features that exist, but are hard to discern in actual gases. The idealization affords epistemic access to those features, and enables us to explore them and their consequences by prescinding from complications that overshadow the features in real cases. The reason why it is valuable is that it equips us to recognize these features, appreciate their significance, and tease out subtle consequences that might be obscured in the welter of complicating factors that obtain in fact. It serves as a focus that facilitates indirect comparisons, where direct comparisons are unilluminating or intractable. We

understand the phenomena in terms of their deviations from the ideal. Such idealizations are not, do not purport to be, and do not aspire to be replaced by truths. But it is hard to deny that they are cognitively valuable, and hard to deny that epistemology should attempt to explain what makes the theories they figure in cognitively valuable.

What should we say about the false factual propositions that occur in the scientific understanding of both scientists and novices? I said that we might concede with Kvanvig that there is something honorific about calling these cases of understanding. At least their claim to be genuine understanding depends on their relation to some (real or anticipated) future account that is cognitively better. I suggest that they too are felicitous falsehoods. The child who thinks that humans descended from apes embeds that contention in a general account that reflects both a cognitive commitment to evolution and an idea that humans and other apes are closely related. So although there is a falsehood involved, it is a falsehood that enables her to connect, synthesize, and grasp a body of information that is grounded in the biological facts, and is supported (to an extent) by her available evidence. It may not be a lot, but it is something. Similarly in the case of Copernicus. The Earth's orbit is not circular. But the Earth can be accurately represented as going around the sun in an orbit that is not all that far from circular. So the falsehood is felicitous in that it figures in, and enables Copernicus to unify, a body of information in a way that answers to the evidence better than his predecessors could.

These felicitous falsehoods are not fictions. Fictive sentences neither are nor purport to be true. They function in other ways. So it is no defect in ideal gas descriptions that there are no gases that instantiate them. But it is a defect in Copernicus's view that the Earth's orbit is not circular, and it is a defect in the child's view that humans did not descend from apes. So understandings that embed propositions like these are in need of improvement. They are just way stations toward a better understanding of the subjects they concern.

Although idealizations per se are not defective, they are epistemically parasitic. The only reason to accept them is that they figure in theories that make sense of the facts. If those theories are discredited, we lose our reason to accept the idealizations they contain. The theories in question are answerable to evidence. So there is no danger that by acknowledging that genuine understanding may involve felicitous falsehoods, epistemology loses touch with facts. For duly accommodating the evidence is answering to the facts. But answering to the evidence is a requirement on the entire theory or comprehensive body of information, not on each individual element of it.

I have argued that the sort of understanding displayed in science falls within epistemology's purview, and that that sort of understanding cannot plausibly be construed as factive. Kvanvig does not discuss scientific understanding. His example is drawn from history. It seems far more plausible that historical understanding is factive than that scientific understanding is. History, for example, does not resort to idealizations or simplifying assumptions. It does not go in for thought experiments. If this is right, then one question that arises for an epistemology that comprehends understanding is how do the understandings afforded by different disciplines differ? By insisting that epistemology should concern itself with understanding, Kvanvig makes such questions salient.³

³ I am grateful to Jonathan Adler, John Greco, Jonathan Kvanvig, Duncan Prichard, and Wayne Riggs for useful comments on an earlier draft of this paper.

REFERENCES

- Elgin, Catherine Z. (1996). *Considered Judgment*. Princeton: Princeton University Press.
- (2004). 'True Enough', *Philosophical Issues*, 14: 113–31.
- Goodman, Nelson (1968). *Languages of Art*. Indianapolis: Hackett.
- Kvanvig, Jonathan (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge, MA: Cambridge University Press.
- Sellars, Wilfrid (1963). 'The Language of Theories', in his *Science, Perception, and Reality*. London: Routledge.

APPENDIX D

Understanding, Knowledge, and the *Meno* Requirement

Wayne D. Riggs

1. INTRODUCTION

Jonathan Kvanvig's book *The Value of Knowledge and the Pursuit of Understanding* (2003), is a wonderful example of doing epistemology in a style that Kvanvig himself has termed 'value-driven epistemology.' On this approach, one takes questions about epistemic value to be central to theoretical concerns, including the concern to provide an adequate account of knowledge. This approach yields the demand that theories of knowledge must provide, not just an adequate account of the *nature* of knowledge, but also an account of the *value* of knowledge. Given the near-universal assumption that knowledge has a special kind of value, this demand seems reasonable, though surprisingly hard to satisfy. Another consequence of this approach to doing epistemology is that certain assumptions about epistemic value, like what sorts of things have it and what sorts of things don't, and where such value comes from, become much more salient to the epistemic enterprise. In his book, Kvanvig challenges the assumption that knowledge has some unique store of epistemic value. And he investigates the matter by asking questions about what the bearers of epistemic value are and where they get it. He concludes, of course, that knowledge as we have come to conceive it in twenty-first century epistemology has no such special value.

Kvanvig frames his task in the book by taking up the question from the *Meno*, 'why is knowledge more valuable than true opinion?' This way of asking the question acknowledges that knowing is valuable, at least in part, because it implies that one has a true belief. It is uncontroversial that we value having true beliefs, so it is uncontroversial that knowledge is valuable at least in virtue of the true belief component of it. But the *Meno* question is prompted by a conviction that there is something particularly good about knowing, beyond thereby possessing a true belief. It is this value that Kvanvig is after, if it exists.

To pursue this quarry in the terrain of contemporary epistemology, Kvanvig is forced to reformulate the *Meno* problem slightly. In these post-Gettier days, we are aware that there might be more than one property that is required in addition to true belief to adequately account for knowledge. Thus, it is not quite sufficiently general to ask why knowledge is more valuable than true belief, because knowledge might require both properties X and Y in addition to true belief. If so, then true belief + X might be more valuable than mere true belief, but true belief + (X & Y) might be no more valuable than true belief + X. In that case, knowledge would be more valuable than mere true

belief, but only because of the value of a proper subset of its components. In such a case, Kvanvig argues, it is not knowledge per se that is more valuable than true belief, but rather true belief + X, which falls short of knowledge. As a consequence, Kvanvig sets the requirement that must be met for a theory of knowledge to properly answer the *Meno* question to be that knowledge must be more valuable than any combination of its subparts. Hereafter, I will call this the *Meno requirement*.

I have explained that the need for generality seems to be what has led Kvanvig to this formulation, but I will argue that it has led him astray. This way of construing the *Meno* question is too demanding to be a plausible constraint on a theory of knowledge. Ironically, the strongest argument for this arises from some of the radical proposals he makes himself at the end of the book. In these brief remarks, I will make the case that Kvanvig's formulation of the *Meno* problem makes demands on a conception of knowledge that are too strong, and I will propose an alternative that serves the purpose just as well, while avoiding the problems of his original one.

Let us take a characteristic statement of Kvanvig's formulation of the *Meno* problem: 'An adequate account of the value of knowledge must explain why it is more valuable than any subset of its constituents' (Kvanvig 2003: 112). There are at least two reasons why this formulation is too strong. First, it begs the question against a certain way of thinking about the value of true belief. Second, it does not square with what Kvanvig himself says about the nature and value of understanding at the end of his book. I'll take these considerations in order.

But first, a quick note about what I am not claiming here. I am not claiming that by weakening the *Meno* requirement, some of Kvanvig's arguments against attempts to account for the value of knowledge thereby fail. It is nonetheless worth correcting the requirement for at least two reasons. One is that it sorts out a tension that I claim is present in the book itself. Kvanvig's formulation of the *Meno* requirement simply doesn't sit well with his later claims about the nature and value of understanding, and my formulation represents a friendly amendment to bring these two elements of his view into harmony. Another reason is that I hope some version of the credit theory of knowledge will be able to answer Kvanvig's charges and account for the value of knowledge.¹ Though I can't make that case today, I at least want to make it easier by watering down—I mean, getting right—the *Meno* requirement.

2. HOW MUCH FOR THAT TRUTH?

It has been noted in passing at various points in the literature that there is some intuitive pull to the idea that not all truths are created equal. That is, the value of some true beliefs seems to exceed that of others, based on the contents of those beliefs. Ernest Sosa expresses a view in this neighborhood at least, in 'The Place of Truth in Epistemology.'

At the beach on a lazy summer afternoon, we might scoop up a handful of sand and carefully count the grains. This would give us an otherwise unremarked truth, something that on the view before us is at least a positive good, other things equal. This view is hard to take seriously. The number of

¹ John Greco has defended a version of such a theory in Greco (2003). I have also defended a different version in Riggs (2007).

grains would not interest most of us in the slightest. Absent any such antecedent interest, moreover, it is hard to see any sort of *value* in one's having that truth. (Sosa 2003: 156)

Sosa here points out that the value of a true belief about the number of grains of sand in one's hand is, intuitively (and absent any specific reason to find out such a thing), not very valuable despite its truth. Compared to a true belief about, say, the nature of reality or the moral law, Sosa seems right. Taking for granted for the moment that there is an intuitive value-discrepancy between true beliefs of these sorts, how shall we account for it?

It is possible, of course, that this value-discrepancy is purely a matter of a difference in non-epistemic value, in which case it would be irrelevant to a discussion of the *Meno* problem. After all, it is uncontroversial that a true belief about the correct way to disarm a bomb has considerably more prudential value for me when said bomb is about to go off than does a true belief about tomorrow's weather at that moment. But once we begin to raise questions directly about the sources and bearers of epistemic value, it is at least reasonable to ask whether the intuitive value-discrepancy between some true beliefs and others can be explained, at least in part, by some difference in *epistemic* value.

One might think, for example, that certain subjects are intrinsically and epistemically more worthy of contemplation, and even belief, than others. Bob Roberts and Jay Wood argue for this in their book on intellectual virtues. About epistemic goods in general, they say, '[e]pistemic goods are not all created equal; in fact, some of them are so far down the value scale as hardly to be goods at all' (Roberts and Wood 2007: 208–9). They list several criteria according to which a virtuous believer will discriminate the relative value amongst various epistemic goods. One such criterion is the evidential supporting role the belief plays in one's overall system of beliefs.

If beliefs often need support, then beliefs can gain in value by being supports for other beliefs. So not only is *having* support a sometime desideratum in beliefs; a belief's *providing* support for other beliefs can also make it more interesting than some others, and a belief can be trivial because of its utter lack of a supporting role. (Roberts and Wood 2007: 156)

But some beliefs can be more valuable than others simply on the basis of their content. Roberts and Wood discuss this under the criterion of the 'worthiness' of beliefs. Some beliefs are more worthy than others because of their bearing on important issues beyond the epistemic—beliefs about the guilt or innocence of a defendant at a trial, for instance (Roberts and Wood 2007: 158). But not all differences in value between beliefs is explicable by appeal to non-epistemic considerations.

[W]e are not saying that knowledge is valuable only if it is 'practical.' Some things are worth knowing even if the knowledge has no 'application.' Why is it worth knowing how old the universe is, while it is not worth knowing how many grains of sand are in a particular cubic centimetre of the Sahara Desert (assuming that both truths are 'useless')? . . . The universe, with all its processes, is worthy of respect. And this worthiness of the objects of knowledge is tied to their particular character—their particular complexity and simplicity, the particulars of their structure and composition and functions. The human genome is interesting because of what it is, whereas the cubic centimetre of the Sahara, simply as so many grains of sand together, is uninteresting because of what *it* is. (Roberts and Wood 2007: 158)

I am not endorsing any of these views, but I think the fact that they are possible views, and, in my opinion, not obviously false, militates against placing a requirement on theories of knowledge that it be impossible for any true belief to be as valuable or more

valuable than any instance of knowledge. Such a requirement seems to assume without argument that knowledge exemplifies all possible epistemic value. Otherwise, it would be possible that a belief, true or false, might derive sufficient value from some source other than by being an instance of knowledge, that it becomes as valuable or even more valuable than some instances of knowledge. I see no reason to preclude this possibility until it is shown that either (a) any such value is necessarily going to be less valuable than an instance of knowledge would be, or (b) there is no other such possible source of value.

It may be, though, that we can modify Kvanvig's formulation of the *Meno* requirement in a way that preserves the generality that prompted the problematic formulation, yet allows for the possibilities mentioned above. Perhaps we could say that an adequate account of the value of knowledge must explain why

$$(\forall s)(\forall p)[\text{Value}(sKp) > \text{Value}(sRp)] \text{ (where } R \text{ is some relation comprising elements of } K, \text{ and } R \neq K)$$

This formulation keeps fixed both the believer and the proposition believed, eliminating possible value-discrepancies between instances of knowledge and states that fall short of knowledge that are due to differences in the specific contents of the propositions believed. It still requires, however, that any instance of S's knowing that p must be more valuable than any instance of S's truly-believing-but-not-knowing that p. Despite the promise of this formulation, it will not quite do either. To see this, we need to look at what is arguably the most original contribution to epistemology in Kvanvig's book—his discussion of understanding.

3. UNDERSTANDING

After doing some necessary spade work, Kvanvig narrows down our commonsense usages of the term 'understanding' to two that are particularly relevant to theoretical epistemology. 'My suggestion . . . is that we focus on understanding in two central uses: when understanding is claimed for some object, such as some subject matter, and when it involves understanding that something is the case' (Kvanvig 2003: 189). This leaves out common uses of the term to indicate that one has grasped the meaning of someone's utterance, or that one understands *how to do* something. I do not take issue with Kvanvig's choices here, so I will continue on the assumption that these two uses are the ones that are of primary epistemological interest.

Kvanvig has a lot of interesting things to say about understanding, but there are three points of particular relevance for my purposes. First, he claims that understanding, like knowledge, is factive. Thus it makes no sense to say that someone understands X, though X is false. But it is not entirely clear what this means when applied to the notion of understanding a subject matter. Kvanvig acknowledges this awkwardness, but dismisses it quickly,

Objectual understanding is, of course, not straightforwardly factive, for only propositions can be true or false. Still, the uses I wish to focus on are ones in which factivity is in the background. For example, to understand politics is to have beliefs about it, and for this objectual understanding to be the kind of interest here requires that these beliefs be true. (Kvanvig 2003: 191)

This appears to explain the sense in which one's understanding, which comprises, perhaps among other things, some beliefs, is factive in a derivative sense. Whatever beliefs are implicated in one's understanding of X must be true.

This strikes me as an implausibly strong requirement for understanding, especially when conjoined with the second of Kvanvig's claims about understanding that are relevant to my present point—that understanding comes in degrees. This second point seems right, since both the degree of explanatory coherence (which Kvanvig takes to be a likely way to account for understanding) as well as the amount of information present in someone's understanding can vary. Why not, then, allow that one can have objectual understanding that includes false beliefs, while acknowledging that such false beliefs lower the *degree* of understanding one has?

Kvanvig himself actually appears to soften the factivity requirement a bit later in his discussion of understanding. For instance, he says,

There is still something of a problem here, though, for it is hard to resist the view that understanding may be correctly ascribed even in the presence of some false beliefs concerning the subject matter . . . When the falsehoods are peripheral, we can ascribe understanding based on the rest of the information grasped that is true and contains no falsehoods. In such a case, the false beliefs are not a part of the understanding the person has, even though they concern the very material regarding which the person has understanding. So in this way, the factive character of understanding can be preserved without having to say that a person with false beliefs about a subject matter can have no understanding of it. (Kvanvig 2003: 201–2)

My objection here may just be a quibble, but I am not sure that we can make sense of what Kvanvig here proposes. First of all, there seem to be cases in which one can have understanding even when one believes something false that is not 'peripheral,' but rather central. Secondly, I'm not sure how to understand the sense in which we can talk about the understanding one has of a subject matter once one has dropped out of consideration some perhaps fairly key beliefs that bear evidential and explanatory connections with other beliefs, your grasp of which constitutes your understanding. Consider an example.

Suppose I understand my wife. I can reliably predict her behavior in a wide range of circumstances, I possess a lot of facts about her past and about her present dispositions, and so on. Included in this set of my beliefs about her past is that she suffered a very traumatic boating accident when she was 3 years old that resulted in her refusal ever to travel on water. Her lifelong attempts to conquer her fear have failed over and over, and all this has caused a certain amount of guilt and pain and humiliation, all of which I have witnessed, and I have a deep appreciation of how all this has affected her character, etc. But, as it happens, she never had a boating accident. Instead, her parents neglectfully let her wander off while they were on vacation in Mexico when she was 2 (not 3), and she fell off a pier into the ocean. As she got older, her parents explained her fear of the water by making up a 'boating accident' that did not render them culpable for the resulting psychological damage.

Obviously, my belief about her boating accident (as well as her belief) is false, and if understanding is strongly factive, it cannot be included in my understanding of my wife. Yet this belief is central to my understanding of my wife's history, experience, and personality. In other words, it hardly seems peripheral. Moreover, even if I was willing to grant that it should be 'deleted' from what constitutes my understanding of my wife,

it is hard for me to imagine what it would be like to 'subtract' this one belief from my total understanding of her. What are we to do with all the evidential and explanatory connections that would be 'left dangling'? Are they to be abandoned as well? If so, it looks as though my understanding of my wife suffers a terrible blow once we subtract my false belief about her alleged boating accident. But it does not seem plausible to me that such a judgment is reasonable. The lesson I take from this example is that my understanding of my wife survives the falsehood of even this central belief, and so even the more lenient constraint of factivity that Kvanvig admits must be loosened some more.

The final, and most significant, of Kvanvig's claims about understanding that I want to highlight is that it is not a species of knowledge. Kvanvig at first acknowledges the intuitive pull of the idea that it is, but then offers several arguments against it. The main argument here is that the primary determinants of knowledge and of understanding are quite different.

Note that the crucial features just discussed concerning understanding draw attention to things other than what is central to knowledge. . . . [O]nce we move past its factivity, the grasping of relations between items of information is central to the nature of understanding. By contrast, when we move past the factivity of knowledge, the central features involve nonaccidental connections between mind and world. (Kvanvig 2003: 197)

Kvanvig buttresses this general argument by appeal to some specific examples, but since I want to grant this point, I need not belabor it here.

The two conclusions I want to leave this section of the essay with are that, by Kvanvig's own lights, he ought to acknowledge that some, and perhaps even a significant, departure from factivity is consistent with even a very high degree of understanding, and Kvanvig's avowal that one can have an understanding of X even in the absence of knowledge about X.

4. THE VALUE OF UNDERSTANDING

There is one more point about understanding that needs making before I can return to the issue of how to properly formulate the *Meno* requirement. I take it to be uncontroversial that understanding is epistemically valuable. But what Kvanvig calls objectual understanding, at least, is a complex and, more to the point, composite thing. Even remaining neutral with respect to whether understanding M is exhausted by a list of beliefs or known information, it is typical that understanding M requires that S have some beliefs about the subject matter that alone do not constitute the whole of S's understanding. Therefore, understanding has parts. And it is valuable. Though the literature on extrinsic value is sparse on the subject, it is customary to suppose that an object that is a contributing part of a valuable whole derives value from that relationship. The fuel injection system in my car is valuable to me because it contributes to the ability of my car to provide transportation, which is valuable to me. Similarly, the beliefs that partly constitute my understanding of M are valuable to me because they contribute to that understanding. Therefore, being a part of one's understanding of some subject matter, M, can be a source of epistemic value for one's beliefs.

5. BACK TO THE *MENO* REQUIREMENT

When we left our hero, the *Meno* requirement, it looked like this:

$(\forall s)(\forall p)[\text{Value}(sKp) > \text{Value}(sRp)]$ (where R is some relation comprising elements of K , and $R \neq K$)

This is a reformulation of Kvanvig's original version so as to avoid the objections raised previously. We are now in a position to see why this reformulation will not work either. The lessons learned about understanding will be crucial here.

All we need to show that the above formulation is problematic is that it is possible that S 's believing but not knowing that p could have equal or greater epistemic value for S than S 's knowing that p . And the things we learned about understanding will allow us to show just such possibilities. Imagine that S understands some subject matter, M , to a very high degree. Included as a part of this understanding is S 's belief that p . Let us further suppose that p is fairly central to S 's understanding of M . As we have seen, S 's belief that p will derive some epistemic value from its inclusion in S 's understanding of M . It is at least possible that this value could be equal to or greater than the value of S 's knowing that p , in a case where that item of knowledge is *not* a part of S 's understanding of M or of anything else. If so, then we have a case in which S 's belief that p has equal or greater value than S 's knowing that p .

Of course, a lot has been left unexplained here. In particular, I have given you no indication of how we are supposed to measure and compare the 'total epistemic value' present in each of these two possible cases. Indeed, there is no guarantee that these two different kinds of epistemic value are even commensurable, allowing a meaningful comparison to be made. But all I am trying to show here is that even the reformulated *Meno* requirement begs the question against such possibilities as described above. As long as there is a possible source of epistemic value for a belief that is independent of whether that belief constitutes knowledge, then there will be the possibility that the belief could derive enough value from that alternative source that it is equally or more valuable than it would be if it were known. Being a part of one's understanding of a subject matter has been shown, by Kvanvig's own lights, to be just such a possible alternative source of epistemic value. So it is unreasonable to require that S 's knowing that p always be more valuable than S 's bearing some relationship short of knowledge to p , because S 's bearing that other relationship to p is compatible with p deriving value from an alternative source. As I indicated before, the only way such a requirement would make sense is if it were already determined either that there was no other source of epistemic value for beliefs besides being true or being an instance of knowledge, or that such alternative value can never rival the value derived from the more traditional sources. I see no reason to accept either of those views at this point.

As a brief aside, notice that it is also possible, at least, that a false belief could derive value from being a part of S 's understanding of M , assuming I am right that the factivity requirement on understanding should be liberalized somewhat. Though it is less plausible, it is still possible that even such a false belief could derive enough value from being part of S 's understanding that M , that it is more valuable than at least some instances of S 's knowing that p .

We need yet another reformulation of the *Meno* requirement, preferably one that is still sufficiently general to satisfy the circumstances of contemporary epistemology, but that does not impose unreasonable demands on a theory of knowledge. The problems raised for the previous formulations arose from the fact that truth and knowledge may not be the only sources of epistemic value out there. If not, then formulating the *Meno* requirement in terms of the relative *amount* of value exhibited by a mere belief on the one hand and an item of knowledge on the other seems a poor way to proceed. It will always be possible that the mere belief could derive enough value from elsewhere to compare favorably with the item of knowledge in terms of the amount of value involved. I think the better way to construe the *Meno* requirement is that an item of knowledge must have a kind of epistemic value, or perhaps a source of epistemic value, that no belief that fails to count as knowledge has. This does not guarantee that knowledge is always more valuable than whatever falls short of knowledge, but it does insure that knowledge always has a special kind of value that nothing that falls short of it can have.

Formulating the requirement in this way avoids all the problems raised for previous versions. All those problems stemmed from the fact that the requirement was put in terms of relative quantity of epistemic value between what is known and what is believed but not known, which is not guaranteed always to come out in favor of the former. Yet I believe that it captures the spirit of the original *Meno* requirement. For instance, one still has an answer to the question, 'why should I prefer knowing to not-knowing?' The answer is that only by achieving knowledge can I attain epistemic value of a certain sort. Thus, all else being equal, knowing will be more valuable than not-knowing. But we have to acknowledge that all else is not always equal. We could, perhaps, have gotten to this point more quickly by simply adding the 'all else being equal' caveat to the second formulation of the *Meno* requirement, but that would have left obscure the kinds of things that must be equal, and so we would have been left with a poorer understanding of the value of knowledge. Also, putting the *Meno* requirement in terms of a unique kind or source of epistemic value that only comes with knowing highlights the fact that knowing still has an exalted position for a reason. It gives us something we can't get anywhere else.

REFERENCES

- Kvanvig, J. (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge University Press, Cambridge.
- Greco, J. (2003). 'Knowledge as Credit for True Belief,' in Michael DePaul and Linda Zagzebski (eds.), *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Oxford University Press.
- Riggs, W. D. (2007). 'Why Epistemologists Are so Down on their Luck,' *Synthese*, 158: 329–44.
- Roberts, Robert, and Wood, W. Jay (2007). *Intellectual Virtues*. Oxford: Oxford University Press.
- Sosa, Ernest (2003). 'The Place of Truth in Epistemology,' in Michael DePaul and Linda Zagzebski (eds.), *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Oxford University Press.

APPENDIX E

Responses to Critics

Jonathan Kvanvig

I begin by expressing my sincere thanks to my critics for taking time from their own impressive projects in epistemology to consider mine. Often, in reading their criticisms, I had the feeling of having received more help than I really wanted! But the truth of the matter is that we learn best by making mistakes, and I appreciate the conscientious attention to my work that my critics have shown.

My responses here will focus on two main areas of contention by my critics. In the first part, I'll address some of the questions and criticisms concerning the nature of understanding and the quasi-factive account I gave of it; and in the second part, I'll talk more about the *Meno* problem and the prospects for a successful response to it.

Prior to doing so, however, I want to address briefly Wayne Riggs's remarks about the value of truth. There is a growing sense among some epistemologists (e.g. Ernest Sosa (2002), Robert Roberts and Jay Wood (2007), Michael DePaul and Stephen Grimm (2007), and perhaps now Riggs as well) that some value for truth applied uniformly and universally to all truth may not be defensible. Instead of counselling the love of truth for inquirers, such critics recommend a more circumspect love for deserving truths only. Perhaps they are correct, but I do not think anyone has provided the necessary argumentative support for the view, and I'll explain why here briefly.

Let me start with an analogy. I think pain is bad, even though pain often contributes to well-being. When I put my hand on a hot stove, the pain causes me to move my hand quickly. If you leave everything the same and remove the pain, things would be much worse.

There's no good reason to explain such examples by saying that the pain in question is good. Instead, the language of defeat helps us understand the case. Pain is a bad-making feature of a case, always and everywhere, but sometimes it occurs in circumstances where its intrinsic badness is overridden or defeated by other factors. We often use the term 'prima facie' to capture this defeasible character of the badness in question.

The view that truth is valuable should be understood in a similar, defeasible way. When philosophers point out that the truth about the number of grains of sand regarding a particular stretch of beach is hard to see as worthy of discovery, that point by itself doesn't undermine the prima facie value of truth. Instead, we can explain this perception as involving defeat of the prima facie value in question with other features: the claim in question doesn't do much explanatory work for us, it doesn't help our overall understanding of the universe much, taking the time to learn it will interfere with other learning, etc.

So it is not hard to see how to resist the claims of lovers of limited truth, but it would be nice to have a strategy for getting around the impasse thus created. I believe a strategy can be outlined that makes it very difficult to defend the limited view. The strategy is to address the value of a posteriori truth first by imagining that you are an angel in Plato's heaven.¹ You have available to you all the information that doesn't depend on concretia, and you are trying to determine which of the truths about concretia one ought to attend to once plucked from Platonic heaven and placed in the world of concrete reality. I think the answer to this question is that you'd have no reason to classify certain truths about concretia as unimportant, prior to placement in the world. You wouldn't be able to classify some as more basic from an explanatory perspective or more important for a complete understanding of any given phenomenon, for these kinds of claims depend on the nature of the concrete universe, not on some features detectable from what is known in Plato's heaven. Once so placed, however, you'd learn things that would immediately generate a list of priorities in terms of important and unimportant truths. That's how defeasibility functions, however, and leaves intact the view that all truths of the sort in question begin from the same level of importance.

If I'm right about this thought experiment, the only route left is to argue that some of the truths available to one within Plato's heaven are themselves unimportant. It is interesting to note, however, that none of the critics of the universal value of truth use examples of truths that would be available in such a setting. Instead, they always appeal to empirical truths about the concrete world that, from our perspective, are fairly worthless (memorizing the phone book, counting grains of sand, etc.). I suspect, then, that any argument for a limitation on the value of truth won't be extendable to the truths available in Plato's heaven. On the simplest picture of the truths here, there are logical truths and metaphysical truths (i.e. necessary truths that aren't logical truths). Which of these are unimportant? The logical truths are all derivable from no premises whatsoever, and are interderivable; the metaphysical truths are truth in all metaphysically possible worlds and mutually entail each other. For any system in which some such truths are proposed as basic, there is another system in which other truths are proposed as basic. What, then, could account for the importance of some such truths and the unimportance of others? I can think of one such property some might wish to appeal to: some such truths are more complex than others. I have no idea, however, why anyone would think that such a property is a mark of importance or unimportance. If no such account can be found, however, then an account of the value of truth in universal yet defeasible terms is able to withstand any of the purported counterexamples to the view, just as the similar view about the badness of pain can withstand counterexamples to it.

None of my critics here, including Riggs, actually endorses the arguments for the limited perspective, however, so I won't pursue this issue any further here. Instead, I want to focus on the questions concerning understanding and the precise nature of the *Meno* problem. I turn first to the issues surrounding the nature of understanding.

¹ I would hope that the following goes without saying, but to prevent any misunderstanding, I will point it out anyway: the thought experiment in question doesn't assume that there is such a thing as Plato's heaven or even that it is possible that there is such a thing. It is a rhetorical device only.

1. ON UNDERSTANDING

Let's call the view I defend the 'quasi-factive view.' The form of understanding that I am interested in, and which I claim escapes the *Meno* problem, is objectual understanding, the kind of understanding in which the content of the attitude is an object of some sort (person, theory, part of reality, etc.). Such understanding is related, presumably constitutively, to various pieces of information, and on the quasi-factive view, the pieces of information that are central to the understanding in question must be true.

My critics raise important concerns about the defensibility of the quasi-factive view, but in my opinion many of these objections can be explained away in one of two ways. Some concerns can be explained away in terms of honorific attributions of understanding and others in terms of the vagueness and ambiguity of ascriptions of understanding. Let me say a few words about the latter issue first. Concerning vagueness, the lesson is clear. On any view, there will have to be some connection to the facts in order for objectual understanding to obtain, but it will always be vague precisely what the connection needs to be. On the quasi-factive view, vagueness is found twice over. First, there is the vagueness of exactly what amount of information can be tolerated as false and yet understanding be retained. Second, there is vagueness in the distinction between central and peripheral pieces of information. When falsehoods are found within the noetic system of a particular person, it will often be vague whether that falsehood is sufficient to undermine understanding. This result is just as it should be, since, as Aristotle informed us, we should not look for more precision than the subject matter allows. More relevant to the cases presented by my critics, however, is that ambiguity is found as well in attributions of understanding. I think something like this is occurring in Riggs's example concerning his wife (pp. 335–6 above). He claims to understand his wife, but when you look at the example, it looks to me like the object of his understanding is his wife's psychological constitution. He understands her moods, her fears, her motives, how she thinks, what her interests are, and the like. For such understanding, it is not a central piece of information precisely how she came to have these features. Had the claim to understanding been one about his wife's life story, her history, I'm inclined to say that the story doesn't sustain that attribution, since what appears to be a major event in her life is misunderstood. Perhaps not, however, since what caused a central feature of her psychological makeup need not be central from the point of view of her life's story. I suspect that is not the case, since any decent biography would seem to need an account of how she came to have the particular aversion in question. In any case, once we disambiguate the ascription of understanding, we can say correctly that Riggs understands his wife's psychological constitution even though he doesn't have a good grasp of some central biographical facts about her.

The other point of note in responding to apparent counterexamples to the quasi-factive view of understanding is to recognize that epistemic terms are often used in an honorific sense. The clearest example of such a use is when we talk about the present state of scientific knowledge about a certain phenomenon. If pressed, most such ascriptions would be retracted: we speak this way to express some laudable epistemic standing for a certain conception of the phenomenon in question, but we would deny in many cases that such a conception is literally known to be true. Such honorific uses can be found using the

notion of understanding as well. On such case is Catherine Elgin's example of the second grader's understanding of evolution. Elgin notes rightly that I will treat such cases on the model of language describing the current state of scientific knowledge, and she objects that such extended usages fall outside the scope of epistemology, on my view. She says,

When we construe such a take on a subject as understanding, Kvanvig believes, we use the term 'understanding' in an honorific sense, just as we use the term 'knowledge' in an honorific sense when we speak of 'the current state of scientific knowledge', while conceding that some of what belongs to the current state of scientific knowledge is false. Such honorific usages of epistemic terms are, he believes, extended usages that fall outside the scope of epistemology. (p. 325 above)

I think this objection is not quite accurate, however. If I were to offer a theory of knowledge, I would not expect it to answer to locutions involving the current state of scientific knowledge, and when we are talking of understanding we should not expect the theory to answer to honorific ascriptions of understanding of evolution to second graders. This point doesn't mean, however, that honorific uses of epistemic terms are not suitable items of epistemological theorizing. They would fall under what I would metaphorically refer to as the pragmatic dimension of epistemic terminology, and it is a worthwhile epistemological project to determine exactly what epistemic reality underlies such uses. My view isn't that such uses fall outside the scope of epistemology, but rather only that we should not confuse real knowledge and understanding with whatever epistemic reality underlies such honorific uses.

This treatment of understanding, as falling in some safe harbor protected by appeal to honorific ascriptions on the one hand and the phenomena of vagueness and ambiguity on the other also helps with the case of scientific understanding that Elgin raises. The heart of the objection, Elgin points out, involves the idealization characteristic of scientific models. Since these models are essential for scientific understanding, and since idealization results in models which are inaccurate in central ways, it appears that scientific understanding is not quasi-factive in the way I claimed. Quoting from her paper,

There is another aspect of science that is even more troublesome for Kvanvig's view. That is science's penchant for idealization. . . . Elimination of idealizations is not a desideratum. Nor is consigning them to the periphery of a theory. . . . I suggest that effective idealizations are felicitous falsehoods. That they are false is evident. They are felicitous in that they afford epistemic access to matters of fact that are otherwise difficult or impossible to discern. Idealizations are fictions expressly designed to highlight subtle or elusive or otherwise nearly inaccessible matters of fact. They do so by exemplifying features they share with the facts. . . . These felicitous falsehoods are not fictions. Fictive sentences neither are nor purport to be true. They function in other ways. (pp. 326–9 above)

I think Elgin is onto something quite important here, but it is not, I think, the idea that understanding is not quasi-factive in the sense I maintained. The issue here concerns the object of understanding. One might understand the model or theory itself, as when one understands phlogiston theory. One does not thereby understand combustion, however. Understanding the world scientifically is not simply a matter of understanding the given model but involves, rather, some relationship between the model and reality. Scientific understanding of the sort in question consists in the possession of a model and a realization of the extent to which it is an idealization and what aspects of reality the model is intended to shed light on. Here Elgin agrees, I believe: she says that the falsity of the idealizations are 'evident.' Of course, she might mean that the falsities in question are evident only to

enlightened epistemologists, but I suspect she realizes that scientists themselves are aware of the idealizing that is occurring and that the understanding made possible by the models they construct involves grasping the nature of the connection between these models and reality. What follows from these points is that the understanding made possible by the sophisticated models of modern scientific inquiry cannot be identified with the information in the model, but must advert to the modeling relationship itself and details about it, including information about the extent to which the model idealizes reality.

Notice as well that this last point is one that Elgin quite correctly stresses. Idealizations are felicitous in virtue of 'exemplifying features they share with facts.' So the understanding generated by the model depends not only on an understanding of the model itself, but the ways in which the model mirrors, and does not mirror, the facts. Once we move past the model itself as being the body of information, however, the point about idealization doesn't yield non-factivity for understanding, but rather affirms it. There is much more to be said here about the precise nature of scientific understanding, but we don't need the details to see the point that the quasi-factive view of understanding tells the right story here once we recognize the fact that the object of understanding is not simply the model itself but some more complex thing involving a relationship between the model and reality.

2. ON THE PROPER UNDERSTANDING OF THE *MENO* PROBLEM

The other major issue I want to respond to concerns the attempts by my critics to undermine the claim that there is a *Meno* problem of the sort I describe. John Greco, for example, takes the heroic line that there simply is no presumption that knowledge is more valuable than its parts, and Riggs holds that a suitably subtle rendering of the problem leaves room for a decent answer to the problem in terms of the notion of credit for true belief. He says, 'Another reason is that I hope some version of the credit theory of knowledge [e.g. the one defended by John Greco] will be able to answer Kvanvig's charges and account for the value of knowledge' (p. 332 above).

Before addressing these issues, however, let me clear up a couple of issues that Greco and Riggs talk about that do not accurately reflect the position I argued for. Greco worries that there is some ambiguity in what I wrote about what the *Meno* problem really is, and Riggs is concerned that what I say about understanding conflicts with how I describe the *Meno* problem. There is some truth in both charges, but a proper understanding of the position I was arguing for will show that there is nothing theoretically worrisome here.

A brief summary of the view I defended will help avoid some side issues. First, the Socratic worry about the value of knowledge over true belief, transposed into the key of contemporary epistemology, is the worry about the value of knowledge over any proper subset of its parts. To avoid repetitiveness, I often put the issue in the book using less precise language, but that was for stylistic effect only. At no point did the work focus on any other question. In characterizing the worry in this way, I did not endorse the idea that knowledge really is, in fact, more valuable than any proper subset of its parts. Instead, my approach was as follows. The history of epistemology focuses on the nature and extent of knowledge (and elements such as justification taken to be partially constitutive of knowledge). Such a focus involves a narrowing of focus from the more

general issues concerning cognitive achievements and excellences of a purely theoretical sort, and I wondered why an interest in such achievements and excellences should focus so exclusively on one particular kind of excellence or success. The beginning of an answer to this question points out that we want to get to the truth but we want more besides. Think of a purely hypothetical meeting on marketing strategies for, say, some pulp journalistic publication. The marketers are trying to decide which human motivation to tap into. One marketer suggests belief itself: he's a fan of fundamentalism and remarks that holding beliefs and being certain of them is so satisfying to human beings that they should use the slogan, 'Inquiring minds want *opinions*.' A second marketer scoffs, suggesting instead, 'Inquiring minds want *the truth*.' The third and fourth marketers are former students of Richard Feldman and Sosa, respectively, and they resist both suggestions, offering instead, 'Inquiring minds want to justifiably believe the truth,' and 'Inquiring minds want to virtuously believe the truth.' Talk of firing the marketing firm begins to circulate, since these proposals are so hopeless, until the bright young former Gettier student (turned marketer) exclaims, 'I know what we should use! We should use "Inquiring minds want to *know*!"'

The publication makes millions relying on that ad campaign.

What is instructive about the example is not only the intuitive superiority of the slogan to its competitors and the central role that the concept of knowledge plays in it, but the natural way in which the proposal is first offered. Notice the force of saying, 'I know what we should use!' Why appeal to the concept of knowledge here? This is something we do all the time, by the way, and doing so serves as an inquiry-stopper. We either reject the claim to knowledge or we acquiesce with the suggestion. Notice that we don't do that for any of the other suggestions in the neighborhood. If someone says, 'I have a true belief about what we should use as a slogan,' such a remark doesn't function in this way. Nor does adding that the belief is either justified or virtuous in the appropriate way. An appeal to knowledge does.

When you conceive of yourself as knowing a given claim, you will be puzzled by any counsel to investigate the matter further. 'Do you know who that is on the stage talking?' 'Yes, I do; it's Kvanvig.' 'Shouldn't you check to make sure?' 'Excuse me?' Knowledge is the kind of thing that licenses closure of inquiry from a purely theoretical point of view. Mere belief does not license such closure, and neither does mere true belief. Moreover, neither does justified true belief, as is shown by the lottery paradox. One can be justified in believing that one's ticket will lose without that justification legitimating the closure of inquiry. It is for reasons such as this that I claim that knowledge is ordinarily thought of, or assumed to be, more valuable than its proper subparts.

This position on the unique value of knowledge is buttressed by some fairly reasonable proposals about the connections between knowledge and assertion. According to Williamson, for example, knowledge is the norm of assertion, formulable pithily as the advice not to say what you don't know to be true. Slightly weaker is the view defended by Max Black, G. E. Moore, Robert Shope, and Peter Unger that, in asserting a claim, one represents oneself as knowing that the claim is true. By forming an acronym of the last names of this group of philosophers, we can affectionately refer to this position as the BUMS view. Whether one agrees with the knowledge norm view or the BUMS view, it is clear that these philosophers are onto something important about our ordinary conception of knowledge. One of the platitudes about the functional role of knowledge ascriptions

is that it is a legitimator of inquiry closure. Nothing similar can be said about belief, true belief, or justified or virtuous true belief. In each case, such theorists could quite naturally display the phenomenon of metalinguistic negation at such proposals: if you suggest that in assertion you represent yourself as believing correctly, a natural response by one of the BUMS would be, 'no you don't; you represent yourself as *knowing*!'

The argument that there is a *Meno* problem concerning the value of knowledge derives from these considerations. It is because attributions of knowledge function to legitimate the closure of inquiry that it ought to have a value that exceeds that of its proper subparts, for these parts do not legitimate such closure. It is banal to remark that Bo believes something that he ought to investigate further, or even that Joe ought to investigate further even though his believing is correct, a display of cognitive excellence, and justified or rational. Things are different, however, when we utter the perplexing 'Bo knows . . . but ought to check further.'

One thing that seems to bother Greco here is the idea that ordinary folk haven't thought enough about knowledge to think of it in this way.

Suppose also that the 'plus' part amounts to some minor, technical adjustment to the traditional idea that knowledge is justified true belief, and that this further condition adds no further value to knowledge over justified true belief. Would these suppositions conflict with ordinary thought? Would this show 'that knowledge does not have the kind of value it is ordinarily thought to have'? It is hard to see how that could be the case. Most people are not at all aware of Gettier problems, and we can suppose that virtually no one was before 1963. But then how could ordinary thought include the idea that knowledge is more valuable than justified true belief? . . .

To sum up, there is no pre-theoretical conviction that knowledge is more valuable than any subset of its constituents, and therefore it is not appropriate to require that a theory of knowledge explain why knowledge has that sort of value. Kvanvig's criterion for an adequate solution to the value problem is too strong. (p. 316 above)

Greco voices two criticisms here. The first is about how the ordinary conception of knowledge could include features that seem to require awareness of the Gettier problem. This issue arises in part because I wrote carelessly regarding the value that knowledge is 'ordinarily thought to have.' Two points are worth recognizing here, however. The first point is a subtle one about the relationship between thoughts, ideas, assumptions, presuppositions, and the like. Our behavior in ascribing knowledge can reveal assumptions and presuppositions about knowledge that are not in our thoughts at all, just as one's behavior can reveal that one is a racist even though one would sincerely disavow it. Our linguistic behavior can show that we are assuming or presupposing that knowledge has some special value even if no one had ever reflected on the issue of the value of knowledge.

One might resist here on grounds that assumptions at least have to be non-occurrent mental states akin to dispositional beliefs, and that ordinary folk not only have never thought about the Gettier problem, they are in no mental state whatsoever that involves a content concerning that problem. I doubt, however, that this view is sustainable. I have two reasons for doubt here. The first questions the claim that assumptions and presuppositions need to be mental states. Why wouldn't it be enough to assume something that one be strongly disposed to believe it, as opposed to believe it already (albeit only dispositionally)? It is one thing to believe a claim dispositionally, and a different thing to be disposed to believe it. In order to sustain the claim that assumptions must be mental states, we'd need an argument against the suggestion that dispositions to be in a relevant mental state could be enough to count as having made an assumption.

The second point makes the stronger point that, even if assumptions are mental states, they are not plausibly thought of as dispositional beliefs. In many cases where our assumptions are pointed out to us, we experience chagrin at the realization. Upon thinking about the particular propositional content in question, we do not embrace it. Instead, we reject it. So what explains the chagrin? To experience chagrin in such an immediate fashion, it would seem that we need some propositional attitude in place that runs contrary to the assumption in question. Suppose, for example, that you are strongly averse to racism, but respond on a given occasion in terms that you agree are accurately described in terms of assuming that people of a particular race are more dangerous. You experience chagrin upon having this fact pointed out to you.

Your chagrin depends upon your aversion to racism, and such aversion involves, I would expect, cognitive commitments. You believe, or are committed to, lots of claims, including the exact opposite of the assumption underlying your response. But in general we don't want to try to make sense of ascribing both the belief that *p* and the belief that not-*p* to an individual (except in cases where modes of presentation explain away the absurdity in question). Nor do we want to say, I think, that some of your beliefs about the races somehow went out of existence during the period of your assumption. Neither will appeal to degree or strength of belief help, so long as we are still willing to countenance the reality of the distinction between beliefs and non-beliefs (even if the threshold varies by context).

The above is a bit of philosophical overkill, since most of it isn't needed to make the basic point that the fact that most people have not considered the Gettier problem and have not reflected on the value problem regarding knowledge does not show that there is no assumption or presupposition that knowledge is more valuable than any proper subset of its parts. So when Greco claims that there is no 'pre-theoretical conviction' that knowledge has some special value, we should perhaps grant the point but deny its relevance. If convictions are conscious, occurrent mental states, then I agree that there is no such generally shared conviction. That leaves untouched, however, the point about assumptions and presuppositions, as well as the data about the functional role of knowledge ascriptions in terms of legitimating the closing of inquiry.

Even worse, if Greco were right that there is no assumption that knowledge is more valuable than any proper subset of its parts, the focus in the history of epistemology on the nature and extent of knowledge would be downright indefensible. Recall that the more general issue for reflection is one concerning cognitive successes and excellences from a purely theoretical point of view (i.e. from a point of view abstracting away from other purposes such as purposes which are practical, moral, aesthetical, religious, or political in nature). Why should the history of reflection on such successes and excellences focus so centrally on knowledge? If knowledge had no special value, there would be no hope of vindicating the history of this philosophical subdiscipline.

This point ties directly into Riggs's worry that what I say about the *Meno* problem may not fit well with what I say about understanding (Riggs, Section 2). He is right about this, but the result is one that should be expected, given the nature of the project I undertook. It is important here that the characterization of the *Meno* problem with which we began is only a working hypothesis. It involves a plausible characterization of the special value knowledge is assumed to have, and it characterizes a viewpoint that has the power to substantiate the obsession the history of epistemology shows concerning

the concept of knowledge. In the end, I do not think that this assumption will survive scrutiny. The point is, rather, that one should approach the theory of knowledge with the working hypothesis that knowledge is more valuable than its proper subparts. If I'm right, this hypothesis cannot withstand careful scrutiny, so it should be no surprise if, in defending the special value of understanding, I say things that do not fit well with the working hypothesis with which I began. The working hypothesis is false precisely because the myopic focus in the history of epistemology on the nature and extent of knowledge cannot be defended. The result may be that some special value can be found for knowledge, but it won't be a special value of the sort that would be needed to justify the singular attention to the concept of knowledge that the history of epistemology displays.

So there is a *Meno* problem and it needs a good solution. Perhaps Riggs and I are right that a better understanding of the problem will show that its proper construal is not enough to justify the history of epistemology, but we should look first to see whether a construal of the problem that has hope of vindicating that history has any chance of success. Here Greco's work is important, for he thinks his virtue account can give us everything we want, and Riggs indicates some agreement on this point. As I argue in the book, we should expect such heroism to fail either on the issue of the nature of knowledge or on the issue of the value of knowledge. In my book, I argued in favor of the idea that beliefs formed through virtue display or indicate a special kind of value, and hence that virtuous true belief is more valuable than true belief. What I argued is that such a view can't explain the value of knowledge, since knowledge is more than virtuous true belief.

This point plays out in expected fashion in some of Greco's own remarks. Greco's own proposal to the *Meno* problem is as follows:

Knowledge is a kind of success through virtue. And in general, success through virtue is more valuable than either success without virtue or virtue without success. In particular, virtuously produced true belief is more valuable than true belief that is not virtuous and virtuous belief that is not true. Neither subset is intrinsically valuable, or constitutive of what is intrinsically valuable, in just the way that knowledge is. (p. 319)

Note that Greco and I agree that success through virtue is more valuable than success without virtue. That is, we agree that virtuous true belief is more valuable than mere true belief. What I deny is that this point solves the *Meno* problem. To do that, the account would have to be an adequate account of the nature of knowledge, and it's not. Familiarity with the Gettier literature reveals a variety of cases to which it succumbs: Carl Ginet's fake barn case, Gilbert Harman's assassination case, the Tom/Buck case offered by Keith Lehrer and Tom Paxson, and even Gettier's original inferential cases. In the fake barn case, the traveler's belief that the thing in the field is a barn is a product of excellent perceptual abilities.² In Harman's case, Jill believes what the reporter tells her in virtuous fashion, and the reporter reports sincerely what he saw, where the perceptual equipment

² The canonical formulation of the case is in Alvin Goldman's 'Discrimination and Perceptual Knowledge' (1976: 772–3), and runs as follows. The local inhabitants have been up to mischief, attempting to replace all the real barns in the area with fake barns. They unwittingly failed to replace one barn, however, and it is this one barn that the traveler looks at and comes to believe that it is a barn.

in question can be as superior as one cares to stipulate.³ In the Tom/Buck case, you know that Tom stole the book because you saw him do it, and the facts about belief production remain constant whether or not the testimony of Tom's mother amounts to a misleading defeater.⁴ In the bad case, the police have no record of contact with her, and so the testimony implies a lack of knowledge on your part. In the good case, the police have a long file on her and don't need to take the story seriously, and so the testimony counts only as a misleading defeater. What is telling is that the facts about the nature of your belief and whether it counts as an instance of success through virtue remain the same across the cases, so success through virtue cannot account for the possible variants of the case. Finally, in Gettier's original inferential cases, there is no reason to suspect that there is some defect of intellectual character or failure to display it. On any ordinary reading of the notions in question, success is achieved through a display of positive character traits.

Greco holds, however, that such an understanding of the key notion of success through virtue is not the one he intends, and uses the fake barn case as a test case for clarifying that notion in a way that provides some hope of escape from the apparent counterexamples above. He says,

[I]t is unclear how perception can count as an ability relative to S's environment and yet be unable to distinguish real barns from fake ones. But suppose we agree for the sake of argument that S's belief is not only true but also formed from a virtue. Still, the belief is not true *because* it is formed from a virtue. Put more carefully, the belief's being so formed does not explain why S has a true belief rather than a false belief. On the contrary, S believes the truth because she happens (luckily) to be looking at the one real barn in the area. If she had been looking anywhere else nearby, excellent perception or no, she would have a false belief. (p. 320 above)

There are two key claims here. The first sentence suggests the possibility of taking a truly heroic path of trying to explain away every Gettier case as involving a person who simply does not have the requisite abilities or virtues in question. Jill's believing that the President was assassinated couldn't count, on this path, as a display of cognitive abilities; your belief that Tom stole the book couldn't either; etc. The trick to this approach is the relativization move: it is not success through virtue that is to be identified with knowledge, but rather success through virtue-relative-to-environment. It is clear that the relativization move can be used to carve off all the cases of knowledge from the cases of non-knowledge: just gerrymander the individuation of environments enough to get that result. The problem, however, is that such gerrymandered individuation will not be useful in the project of trying to explain the value of knowledge but will instead undermine any such account. That, as I have argued, is the price of gerrymandering and ad-hocery.

³ The assassination case is found in Gilbert Harman's *Thought* (1973). The case involves the assassination of the leader of a country, and a report of it by an eyewitness reporter. Jill hears the original report before the CIA puts out a false story that it wasn't the leader, but only a CIA agent who was assassinated. The false story is carried by all major news organizations, but Jill doesn't see or hear any such report, even though they are readily available. So she believes the original, lone report.

⁴ The Tom Grabit case is found in Keith Lehrer and Thomas D. Paxson, Jr. 1969. The case involves a perceptual experience of your close friend, Tom, stealing a book from the library. The police are called in to handle the crime, and they go to his house to find him. His mother tells them that he's gone on a trip and has a twin, Buck, who must have stolen the book. This defeating information doesn't undermine your knowledge, however, because the woman is obviously deranged and known to have recently been released from a mental institution.

To see how this charge plays out in a given case, consider variants on the Tom/Buck case. In one variant, you know that Tom stole the book because the police have this long track record with Tom's mother, so that her testimony counts only as a misleading defeater. In another variant, you don't know that Tom stole the book because the police have no such track record and have to take her story seriously. The question for the heroic path here is how the relativization is supposed to work. The answer is that the two 'environments' have to differ, in spite of the fact that they can be as far away from you geographically as you please. There is no interesting intuitive notion of 'environment' on which the facts about the prior relationship between Tom's mother and the police constitutes any part of the environment in which you observed Tom steal the book and flee from the library. One can, of course, concoct a technical notion to get the desired result, but one does so only at the expense of drawing distinctions that clearly have no relevance whatsoever to the value problem, leaving unexplained why one should value knowledge over its lack in the case in question.

It is not clear, however, that John intends to take the heroic path, for after suggesting it he quickly moves to a different approach. On this approach, the crucial issue is a causal/explanatory one. On it, we acknowledge the obvious point that in a broad range of classic Gettier examples a true belief results from a display of cognitive virtues or excellences. What is missing, according to Greco, is the idea that the true belief is achieved *because of* the virtues, the virtues do not *explain why* the person in question has a true belief rather than a false one.

I think there are three fundamental difficulties with this proposal, and I'll take them in increasing order of importance. The first point concerns the evidence that Greco cites for concluding that the virtues do not explain, in the fake barn case, why one has a true belief rather than a false one. Notice that John infers that the belief's being the product of a virtue does not explain why the belief is true rather than false from the counterfactual claim that if she had looked anywhere else nearby with the same virtues, she would have had a false belief. This counterfactual claim, however, is, in general, false. It is only true when we restrict the claim to what she takes to be a barn. Regarding trees and shrubs, cows and horses, and people and houses, everything would be fine. So the counterfactual has to be very specific to the content of the belief in question to support the conclusion Greco wishes to draw, and even then the claim need not be true because of the possibility of Frankfurt cases. Let the demon be interested in your having true beliefs about barns. The landscape is still littered with fake barns, but the demon will do nothing unless you direct your attention to a fake, in which case he'll zap it into a real barn. But you don't look at the fakes, you look at the one real barn. The demon does nothing. You don't know, however, that the object in question is a real barn, and this even though had you looked at any other barn-like thing in your environment you would have had a true belief. In short, one cannot use the truth or falsity of the counterfactual in question as decisive evidence for the explanatory claim Greco uses to try to avoid the fake barn case.

Moreover, the strategy instanced here for the fake barn case is hardly capable of being exported to all the other cases. In particular, it won't be any help in the variants of the Tom/Buck case noted earlier. The difference in the two variants concerns differences in the past interaction between the police and Tom's mother, but it is simply implausible to think this difference implies any difference in the explanation of why your belief is true rather than false.

The second point is that it is unclear that there is any notion of explanation that will generate the desired result. In the fake barn case, the perceptual abilities of the observer do explain, in part, why a true belief is achieved. Those abilities are not the only factors relevant to a complete explanation, but the obvious fact is that our abilities never constitute a total explanation of why we get to the truth rather than fail. Only infallible abilities would be good enough to yield a complete explanation on the basis of the abilities alone. So the critical question such an account must face is why it will focus on one part of the complete explanation in one case but not in another. Here the philosophical investigation of the term 'the cause' is probably instructive. The result of this investigation is pretty clearly the following: what counts as the cause of a given event is not any set of semantically invariant features. Instead, what counts as the cause of an event seems to vary with the interests and purposes involved in a given context or conversation. For example, what was the cause of the forest fire, the campfire not completely put out, the trail of trash strewn by the bear, or the decade-long drought that allowed the smouldering campfire to move to nearby leaves, then to the trash, and finally to the entire forest? Environmentalists will cite the campfire and Republicans will favor the drought, while the scientist will, in all likelihood, abstain on the project of selecting a distinguishing factor from the set of causal factors underlying the forest fire.

The pessimistic conclusion grounded in the literature on the notion of the cause is that there is no interest-independent notion here to be found.⁵ If that conclusion is correct, then the best Greco could hope for is that our language of explanation tracks our interest in knowledge over non-knowledge, and even if we are interested in knowledge, there is no reason to suppose that the language of explanation will always favor this interest over other interests. Since that assumption is surely implausible, Greco's proposal will have the result that attributions of knowledge have to track whatever interests are dominant, *even when these interests are clearly trumping our interest in knowledge*. That's interest-relativity in the theory of knowledge that is indefensible.

The third problem with Greco's proposal is that it is hard to reconcile with the variety of kinds of knowledge that are possible. One kind is testimonial knowledge of a mundane sort, where you simply take the word of someone about a given topic. Children learn in this way much before they acquire the critical perspective to assess their sources, and any suitably nuanced attempt to sort explanatory factors properly should put more stress on the virtues of the source instead of on the virtues of the child in explaining why the child has acquired a true rather than false belief. There is, of course, the possibility of stipulating some notion of explanation that places greater stress on the child's virtues, but by now we should be familiar with the objection that results: the degree of stipulation is inversely proportional to the plausibility such an account possesses for addressing the value problem.

An even more difficult possibility than knowledge by testimony, however, is the possibility of innate knowledge. It may be that we have no innate knowledge, but in some epistemologically interesting sense, the possibility of such knowledge exists. Any account of the value of knowledge should be able to explain the value of innate knowledge as well, but the idea that innate knowledge involves some kind of success through virtue, at

⁵ See e.g. Mackie (1980), as well as the papers collected in Sosa (1975) and Collins, Hall, and Paul (2004).

least on the part of the individual in question, is difficult to imagine. As before, I'm sure we can stipulate some understanding of success through virtue to yield this result, but stipulations don't explain value.

3. CONCLUSION

I believe, therefore, that there is still an unsolved *Meno* problem and that future attempts to solve the problem as I have stated it will fail. I have no compelling argument for the prediction, however, and the discussion here is a good illustration of the direction discussion must take about this prediction. Those less pessimistic than I will continue to propose solutions, and those sharing my pessimism will continue to try to demonstrate that the solutions do not work.

About the quasi-factive view of understanding I defended, I believe there is much more to be said both by way of elucidation and by way of criticism. In one way, this conclusion is wholly unsurprising, since there has been so little discussion in the history of epistemology concerning the nature of understanding. I will end by expressing the hope that future work in the field will remedy this deficiency.

REFERENCES

- Collins, J., Hall, N. and Paul, L. (eds.) (2004). *Counterfactuals and Causation*. Cambridge, MA: MIT Press.
- DePaul, M. and Grimm, S. (2007). 'Review Essay: Jonathan Kvanvig's *The Value of Knowledge and the Pursuit of Understanding*,' *Philosophy and Phenomenological Research*, 74: 498–514.
- Goldman, A. (1976). 'Discrimination and Perceptual Knowledge,' *Journal of Philosophy*, 73: 771–91.
- Harman, G. (1973) *Thought*. Princeton: Princeton University Press.
- Lehrer, K., and Paxson, Jr., T. D. (1969) 'Knowledge: Undefeated Justified True Belief,' *Journal of Philosophy*, 66: 225–37.
- Mackie, J. L. (1980). *The Cement of the Universe*. Oxford: Oxford University Press.
- Roberts, R. and Wood, J. (2007). *Intellectual Virtues: An Essay in Regulative Epistemology*. Oxford: Oxford University Press.
- Sosa, E., (ed.) (1975). *Causation and Conditionals*. Oxford: Oxford University Press.
- (2002). 'The Place of Truth in Epistemology', in M. DePaul and L. Zagzebski (eds.), *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Oxford University Press.

This page intentionally left blank

Index

- Adler, J. 150–1, 329 n.
- alethic correctness 10–11
- Alston, W. 24 n., 139 n., 243–51, 252, 253, 254 n., 257, 258 n., 261, 266
- analysis:
 - and definition 133–4
 - and immunity to intuitive counterexamples 124–5, 130
 - conclusive reason analysis 119
 - defeasibility analysis 119–21
 - gerrymandered analysis *see* ugly analyses
 - of concepts vs theories of things 123–5
 - of goodness 133–5
 - of knowledge *see* analysis of knowledge
 - of true belief 115, 121
 - ugly analysis *see* ugly analyses
- analysis of knowledge 5, 22, 112–37
 - and value problem 42–58
 - non-conceptual/metaphysical analysis 7
 - ugly analysis of knowledge *see* ugly analyses
- archer analogy 38
- Aristotle 247, 318, 341
- Armstrong, D. 30 n.
- assertability conditions 70
- Audi, R. 251 n., 254 n., 262 n.
- Ayers, M. 143, 159 n.

- Bach, K. 302 n.
- bachelorhood example 317
- Bachr, J. 5–6, 262 n.
- Baier, A. 274
- bank example 191, 197–8
- barn façade case *see* fake barns
- Barnes, B. 60 n., 61, 69, 77–9, 83–5, 89
- Barnes, B., Bloor, D. and Henry, J. 89
- Barnes's social theory of free will and determinism 84
- Bartaly, H. 262 n.
- Baumann, P. 282 n.
- Beckermann, A. 21 n.
- belief:
 - and voluntary control 254 n.
 - as desire 187
 - Bayesian/functionalist conception of 186–8
 - degrees of belief/subjective probability 186–8, 194
 - dispositional theory of 186
 - meta-ethics of 231–7
 - pragmatic picture of 186
 - reasons to believe *see* reasons, to believe
 - vs. acceptance 140
- Ben Ze'ev, A. 274
- Bennett, J. 81
- Bird, A. 91 n.
- Black, M. 344
- Blackburn, S. 233–5
- Bloomfield, P. 240 n.
- Bloor, D. 73, 82
- boating accident example 335–6, 341
- Boer, S. and Lycan, W. 96 n.
- Boghossian, P. 229
- BonJour, L. 42 n., 174, 243 n.
- Bontly, T. 240 n.
- Boyle, R. 85, 87
- Brady, M. 11–12, 183
- Brenner, S. 91 n.
- Brewer, B. 152
- Brogaard, B. 12–13, 37 n., 100 n., 182 n., 201, 287 n., 298 n., 299 n., 300 n.
- Broncano, F. 91 n.
- burglar example *see* practical reasoning, and burglar example
- buried treasure example 208, 214–15

- Cappelen, H. and Lepore, E. 297 n.
- causal/functional interpretation
 - function 115–16, 123 n., 135
- chocolate analogy 24 n.
- Clifford, W. K. 141, 259
- Code, L. 64 n.
- Cohen, S. 191 Comesaña, J. 175
- Collins, J., Hall, N., and Paul, L. A. 350 n.
- Colorado Rally analogy 165–6
- communitarian value-driven epistemology *see* epistemology, value driven approach to conceptual synthesis *see* genealogy
- conditional probability solution 4–5, 27–31
- Conee, E. and Feldman, R. 191, 196
- context/circumstance shifts 299–304
- contextualism 295–300
- Copernicus 325–6, 329
- Craig, E. 6, 61, 64–7, 68, 69, 70–3, 74, 75, 76, 80, 81, 88, 90, 139 n., 159 n., 259, 313 n.
- credit/credit-worthiness 3, 8, 23, 36–9, 81, 82, 140, 142, 156–60, 219–21, 319–20, 332, 343
- crystal healing example 270
- Cuneo, T. 254 n.

- curiosity 12
 about truth *simpliciter* 272
 and normativity of interest 276–7
 and response-dependent account of what is boring/interesting 276
 and selective attention 273–4
 and teleological approach to epistemic appraisal *see* teleological accounts, and curiosity
 and value of truth 265–82
 as alerting us to intrinsically interesting/fascinating questions 275–82
 as alerting us to intrinsically interesting/fascinating truths 281–2
 as emotion 274–5, 277–8
 as non-open-ended 274–5
 compulsive curiosity 270–1
 gap between curiosity and intrinsic value of truth 270–5
 idealized rational curiosity 271–2
- Daedalus 309
- Dancy, J. 114 n., 145 n.
- Darwall, S., Gibbard, A., and Railton, P. 133 n.
- David, M. 115 n., 135 n., 226 n., 228 n., 229 n., 230, 243 n., 248 n., 266
- decision theory 187–8, 194 n.
- defeasibility 119–21
- deference-emotion system 78
- DePaul, M. 7, 229 n., 246 n., 262 n.
- DePaul, M. and Grimm, S. 7, 104–9, 117 n., 135 n., 185 n., 339
- DePaul, M. and Zagzebski, L. 50 n.
- DeRose, K. 21 n., 73, 174, 182 n., 191, 197
- Deutscher, 30 n.
- Doris, J. and Stich, S. 34 n.
- double counting problem *see* swamping problem
- Douven, I. 182 n., 284 n.
- Dowe, D. 101 n.
- doxastic goods 7–8, 139–60
 in relation to reasons and motivation *see* reasons, and motivation; reasons, and goods
 list of possible 140–1
 surreptitious *see* surreptitious goods
- Driver, J. 155 n.
- Dummett, M. 236, 287 n.
- Durkheim, E. 73, 82
- Dutant, J. 194, 201 n.
- eclipse example 216–19
- Egan, A., Hawthorne, J. and Weatherson, B. 291 n.
- Elgin, C. 13, 64 n., 73 n., 327 n., 342–3
- Eliot, T. S. 311
- Engel, M. 164 n., 182 n.
- Engel, P. 9, 189 n.
- enquiry *see* inquiry; exploration
- epistemic appraisal 1, 4, 11
 and normativity 253–62
 and truth 225–304
 teleological approach *see* teleological accounts
- epistemic evaluation *see* epistemic appraisal
- epistemic good 285–303
- epistemic injustice 81
- epistemic luck *see* luck
- epistemic value monism/pluralism 12–13, 284–304
- epistemology:
 and ordinary language/intuitions 7, 34, 107
see also guiding intuition; value problem, and ordinary thinking/linguistic behaviour
 naturalistic 35
 social turn in 64
 value turn in 139, 155 *see also* value-driven epistemology
- espresso analogy 24, 28, 310, 314 *see also* green tea analogy
- eternal truth values 286–8
- Evans, S. 91 n.
- evidentialism 141, 149–51, 190–200
- evolution example 325, 342
- exploration 139–42
 the aim of 158–60
see also inquiry
- expressivism 10–11, 231–40
 about truth as a goal of enquiry *see* expressivism, epistemic
- alethic 231–7, 238
 and deflationism/minimalism 233, 235–7
 engaged and disengaged stances of 10–11, 233–5, 237, 238, 239
- epistemic 238–40
- fake barns 174, 213–14, 320, 347–50
- family resemblance 71
- Fantl, J. 193 n., 195 n., 201 n.
- Fantl, J. and McGrath, M. 9, 184–5, 190–7, 198, 199
- Farland, D. 160 n.
- Feller, D. 91 n.
- Field, H. 233 n., 236 n., 238
- Finn, M. 91 n.
- firefighter example 286–9, 290
- Foley, R. 109 n., 153–4, 243 n., 266
- Forrester, J. 91 n.
- Frankena, W. 77

- freedom:
 and honouring *see* honouring, and freedom
 and knowledge *see* knowledge, and freedom
 and testimony *see* testimony, and freedom
 and trustworthiness *see* trustworthiness, and freedom
 protofreedom 69, 83, 85
 social theory of 83
 free-riding 73–5
 'Fregean' semantic thesis 285–90
 Fricker, M. 64 n., 81
 Fricker, M. and Hornsby, J. 64 n.
 Friedman, J. 160 n.
 functionalism *see* normative determinism
- Gagnier, R. and Dupré, J. 60 n.
 Gelfert, A. 91 n.
 genealogy 6, 60–91
 imaginary and real 75, 87
 Craig's 61, 64–7, 69, 71
 Gettier, E. 2, 118, 163
 Gettier cases 2, 6, 38, 42 n., 63, 87, 90–1,
 118–21, 99–100, 103–10, 173, 174,
 180–2, 204, 211, 213, 214, 309, 310,
 311, 315–16, 319, 331, 345,
 347–9
 Gibbard, A. 145, 189
 Ginet, C. 347
 Goldman, A. 21 n., 22 n., 24 n., 37, 43 n.,
 64 n., 174, 243–51, 252, 253, 257, 258
 n., 261, 266, 268–9, 347 n.
 Goldman, A. and Olsson, E. 4–5
 Goodman, N. 327 n.
 Greco 3, 5–6, 13, 14, 37, 37 n., 42 n.,
 43 n., 44 n., 47 n., 53 n., 55, 64, 67, 81,
 103 n., 156 n., 183, 220 n., 259, 262 n.,
 318 n., 329 n., 332 n., 343, 344–50
 Green, S. 182 n.
 Green, J. and Haidt, J. 34 n.
 Greenough, P. 240 n.
 green tea analogy 62, *see also* espresso analogy
 Grice, P. 73
 Grimm, S. 11, 135 n., 250 n., 266 n., 267–9,
 273 n., 281 n.
 guiding intuition 5, 42–3, 44–6
 alternative conception of 55–7
 content of 47–55
 formality of 51–4
 generality of 47–51
see also value problem, and ordinary
 thinking/linguistic behaviour
- Haack, S. 248 n.
 Haddock, A. 46 n., 201 n., 240 n., 262 n.
 Hanfling, O. 64 n.
- Hardwig, J. 89
 Harman, G. 250 n., 347, 348 n.
 Hawthorne, J. 8, 19, 163–71, 184, 192, 197,
 199
 Hawthorne, J. and Stanley, J. 184 n., 200 n.
 Hazlett, A. 259 n., 262 n.
 Hempel, C. 266, 268–9
 Hetherington, S. 21 n., 213
 Higginbotham, J. 96 n.
 Hintikka, J. 96 n.
 honouring 69, 77–9, 80, 82, 88
 and freedom 83–4
 dishonouring 79, 81
 Hookway, C. 266 n.
 Horwich, P. 9, 185, 188–90, 233 n., 236 n.,
 247 n.
 Hume, D. 112, 134–5, 269 n.
 Hurka, T. 265 n.
 Hyslop, J. 91 n.
- ignorance, *see* knowledge, and ignorance
 individualism 78, 82
 inquiry:
 and epistemic windfalls 280
 and the value of truth 265–82, *see also*
 teleological accounts
 for its own sake 267–8, *see also* curiosity
 normative failings in inquiry 272
see also exploration; curiosity; teleological
 accounts
- instrumentalism 31–2, 156, 158, 184
 intellectualism/purism, *see* pragmatic
 encroachment, vs. purism/intellectualism
 interactionism 78
 invariantism 295, 297
 Irvine, B. 91
 Izard, C. and Ackerman, B. 274
- James, W. 82, 141, 226, 228
 Jardine, N. 91 n.
 Jenkins, C. 240 n.
 Jones, R. 51 n.
 Jones, W. 7–8, 22, 23, 24 n., 30 n., 31 n.,
 42 n., 64 n., 141 n., 144 n., 155 n., 159 n.
 Joyce, J. 187 n.
 judge parameters 302
 justification 9, 103, 105, 117–18
 and stakes 191–7, *see also* pragmatic
 encroachment, and justification
 coherence theories of 64
 enlarged justification/justification* 193–7
 pragmatic encroachment upon, *see* pragmatic
 encroachment, and justification
 subjective 118
 value of 76

- Kant, I. 33, 61
Kaplan, M. 166 n., 180 n., 199, 286, 290 n., 294, 298, 299 n., 300
Kay, A. 91 n.
Kelly, T. 249, 258
Kepler, J. 325–6
King, J. 286 n., 287
Kitcher, P. 101
Klein, P. 120
knowledge:
 and contextualism 73
 and freedom 83–8
 and ignorance 19–21
 and moral obligation 82
 as achievement 43 n.
 as success through virtue/virtuous
 agency 318, 319–20, 347, *see also*
 virtue theory
 BUMS view of 344–5
 disvaluable 50–1
 factive vs non-factive 72
 honorific 82
 knowing-that 96–102, 104, 105, 106, 108
 knowing-why 100–1, 106, 107
 machine-product concept of 36
 protoknowledge 6, 65–7, 68, 69, 70, 71, 76, 79, 80, 84, 85, 88, 91
 semantics versus pragmatics of 73
 sociology of scientific knowledge 61, 88
 Swiss Army knife model of 8, 163–4, 180–2, 165–6
 trivial knowledge 50
 types of 96–102
 vs. psychological certainty 72
 weak vs. strong 19–21,
Kochan, J. 91 n.
Kölbel, M. 291 n.
Kornblith, H. 36 n., 245 n., 254
Korsgaard, C. 152, 265 n.
Kripke, S. 70–1
Kusch, M. 6, 60, 64 n., 70, 80, 89, 259 n., 260 n., 262 n.
Kutcher, F. 21 n.
Kvanvig, J. 3, 4, 6–7, 13–14, 22, 23, 24–7, 30, 30 n., 46 n., 47 n., 56 n., 61–4, 67–70, 72–3, 76, 82, 83, 88, 90, 95 n., 96 n., 103 n., 112–13, 115, 117–21, 121–8, 132 n., 133–4, 136–7, 139 n., 156 n., 163, 172 n., 183, 185 n., 204 n., 247 n., 265 n., 266, 267, 268 n., 272–3, 275 n., 309–51
Lackey, J. 10, 82 n., 135 n., 207–10, 211–12
Lane, M. 64 n., 91 n.
language games 71
Larissa, road to, *see* the Meno problem
Lasersohn, P. 295, 301–2
Latus, A. 206, 215–16
Le Morvan, 21 n.
Lehrer, K. 243 n.,
Lehrer, K. and Paxson, T. 347, 348 n.
Lenzen, W. 21 n.
Leontius 270 n.
Leopold, A. 260 n.
Levi, I. 21
Levi, P. 85
Lewis, D. 37 n., 96 n., 166 n., 187, 286 n., 287
Lipton, P. 91 n.
Lloyd, G. 80, 91 n.
lottery 8, 26, 214, 344, *see also* practical reasoning, and lottery problem
luck 6, 8, 9–10, 38, 63, 64, 66, 100, 105, 109, 118–19, 204–21
 and credit-worthiness 220–1
 and value of knowledge 220–1
 lack-of-control account of 10, 205, 206–7, 214–21
 reflective epistemic luck 210–13
 safety account of 206, 207–10, 210–14, 220
 value+chance model of 206
 veritic epistemic luck 210–13
Lycan, B. 213
Lynch, M. 10–11, 83, 225 n., 226 n., 227 n., 228 n., 229 n., 230, 231, 236 n., 238 n., 243–51, 252, 257, 265 n., 266, 268 n., 272–3, 275 n., 284, 285 n., 293
McDowell, J. 8, 145–6, 159
MacFarlane, J. 291, 294 n., 296 n., 298 n.
McGrath, M. 182 n., 201 n., 230
Macpherson, F. 282 n.
Mackie, J. L. 350 n.
Malcolm, N. 73
marketing example 344
Martinich, A. P. 182 n.
Menke, C. 91 n.
Meno problem 1–2, 3, 4, 14, 21–22, 28, 42 n., 49 n., 55, 105, 163, 183, 309, 311, 313, 314, 331–2, 333, 339, 340, 341, 343–51, *see also* swamping problem; value of knowledge
Meno requirement, *see* value of knowledge,
 Meno requirement
Milgram, E. 182 n.
Millar, A. 117 n., 132 n., 135 n., 152 n., 160 n., 168 n., 182 n., 226 n., 240 n., 254 n., 262 n.
misspeaking 72
Mitova, V. 160 n.
Moffet, M. 181 n.

- Moore, G. E. 73, 75, 76–7, 133–5, 265 n., 344
 Moore's method of absolute isolation 76–7
 Moore's open question argument, *see* open question argument
 Moran, R. 150 n.
- Nagel, T. 206
 Neta, R. 64 n., 182 n., 245 n.
 Neta, R. and Rohrbaugh, G. 175
 Nicomachean Ethics 318
 Nietzsche, F. 61
 normative determinism 78
- O'Hare, M. 280 n.
 Olmos, P. 91 n.
 Olympic gold medals analogy 37–8
 open question argument, 133–5
 Owens, D. 143, 195, 196 n.
- Pace, M. 58
 Padovani, F. 91 n.
 pain analogy 339
 Papineau, D. 231
 Parker, M. 160 n.
 Pascal, B. 149, 154–5
 Peachman, H. 86
 Pears, D. 141 n.
 Percival, P. 24 n., 30, 36, 187 n.
 perspectivalism 13, 298–304
 Piller, C. 182 n., 225 n.
 Plantinga, A. 24 n., 254 n.
 Plato 97, 163, 309, 313–14, *see also* Meno problem
 Plato's heaven 340
 Plato's Meno, *see* Meno problem
 practical reasoning 8
 and burglar example 171–3, 184
 and concept and value of knowledge 163–82
 and counterfactual success 173–6
 and Hawthorne's analysis of knowledge 164–71
 and lottery problem 166–71, 171–3, 175–9, 184, 199
 and probabilities 177–9
 and safety 173–6
 and subjective probabilities 178–9
 epistemic reasons as reducible to practical reasons 185, 188–90
 see also pragmatic encroachment; pragmatic accounts of value of knowledge
 pragmatic encroachment 9
 and contextualism 197
 and epistemic value 183–201
 and interest-relative invariantism 197
 and justification 190–7
 and lottery problem *see* practical reasoning, and lottery problem
 and value of knowledge 183–6, 200–1
 on belief and truth 186–90
 on knowledge 197–200
 vs. purism/intellectualism 185, 190, 197–200
 see also practical reasoning; pragmatic accounts of value of knowledge
 pragmatic accounts of value of knowledge 67–70, 141
 objections to 67–70
 see also practical reasoning; pragmatic encroachment
 pragmatism 82, 186, *see also* practical reasoning; pragmatic encroachment, pragmatic accounts of value of knowledge
 Princess Diana's dress example 36
 Prior, A. 290
 Pritchard, D. 3, 9, 10, 60 n., 109 n., 174, 182 n., 183, 194 n., 201 n., 206, 207–10, 210–14, 220, 240 n., 262 n., 273 n., 282 n., 329
 property parasitism, *see* reliabilism, and property parasitism
 protofreedom, *see* freedom, protofreedom
 protoknowledge, *see* knowledge, protoknowledge
 purism/intellectualism, *see* pragmatic encroachment, vs. purism/intellectualism
- Raatzsch, R. 91 n.
 Rabinowicz, W. and Roennow-Rasmussen, T. 37 n.
 racism example 211–13
 Railton, P. 269 n.
 Rampelt, J. 91 n.
 Ramsey, F. 186
 Ramsey's ladder 233
 rational choice theory 78
 Raz, J. 147
 reasons:
 and goods 146–8
 and motivation 145–6, 148–9, 149–53
 and norms 151–3
 for action 8, 145
 external 145, 146
 internal 145
 normative and motivating sides of 145
 to believe 145
 to believe and surreptitious doxastic goods 148–53
 to believe vs. reasons to act so as to gain belief 149, 158
 Recanati, F. 299, 302

- Reed, B. 90
 relativism/relative expressions 12, 285, 286, 291, 292, 293, 295–7, 298, 303
 reliabilism 4–5, 19–40, 42 n., 43, 62, 67, 103, 118, 314, 319
 and empirical regularities 29–30
 and objective probability 25, 28, 35, 62, 76
 and property parasitism 25–7
 reliabilist responses to the swamping problem 27–35
 swamping problem for reliabilism, *see* swamping problem
 value of reliabilist knowledge 21–3
 virtue reliabilism 36, 37–40
 Rescher, N. 206
 Reynolds, S. 64 n.
 Richard, M. 291 n.
 Riggs 3, 9–10, 14, 22, 23, 24 n., 30, 37, 42 n., 44 n., 60 n., 81, 103 n., 139, 156, 205 n., 215, 220 n., 262 n., 284 n., 318 n., 329, 332 n., 339, 340, 341, 343, 346–7
 Roberts, R. and Wood, W. 333–4, 339
 Rosenberg, J. 91 n.
 Rudwick, M. 91 n.
- safety 9, 38
 and practical reasoning, *see* practical reasoning, and safety
 and reliable reliability 175
 Salerno, J. 201 n.
 Salmon, W. 101, 287 n.
 Sartwell, C. 21
 Sayre, K. 97 n.
 Sayre-McCord, G. 134 n.
 scepticism about value of truth, *see* expressivism
 Schaffer, S. 81, 91 n., 96 n.
 Scheff, T. 78
 Schmitt, F. 64 n.
 Searle, J. 73, 89
 Sellars, W. 326
 Shah, N. 150, 189 n., 229–30, 251 n.
 Shah, N. and Velleman, D. 229, 251 n.
 Shapin, S. 61, 69, 83–5, 85–7, 89
 Shapin, S. and Schaffer, S. 85 n., 87
 Shope, R. 344
 Smartly the valedictorian example 209–10, 214–15
 Smith, M. 269 n., 271 n.
 soccer match example 104–9
 Socrates 2, 183, 309, 313–14, 315, *see also* Meno problem
 Söderqvist, T. 91 n.
 Sosa 3, 11, 22, 23, 24 n., 37–8, 39–40, 42 n., 81, 103 n., 174, 183, 220 n., 225 n., 226 n., 243, 246 n., 250, 251–7, 261, 262 n., 266, 267, 273 n., 318 n., 319, 332–3, 339, 350 n.
 Sparhawk, A. 160 n.
 Speak, D. 58
 Stalnaker, R. 186
 Stanley, J. 9, 184, 185, 192, 193, 194 n., 195, 197–200, 287 n., 301
 Stanley, J. and Williamson, T. 96 n.
 Stich, S. 112–13, 115–17, 121–8, 131, 133–5, 135–6, 245 n.
 Sturgeon, S. 201 n.
 subjective justification 62–3
 surreptitious goods 8, 141–60
 as motivators for action 154–5
 attenuated role of 153–5
 case against surreptitious goods as reasons 149–53
 of believing 142–4
 significance of surreptitious value 153–60
 value of knowledge as surreptitious 155–60
 swamping problem 2–3, 4, 6, 22–3, 23–7, 62, 67–8, 75–7, 88, 103, 118, 310, *see also* Meno problem
 Swampman 109–10
 Swinburne, R. 22, 23, 30
 Swiss Army knife model of knowledge, *see* Knowledge, Swiss Army knife model of
- teleological accounts 11, 243–62, 265–82
 and common good 259–62
 and curiosity 244–5, 247, 251, 265–82, *see also* curiosity
 and intrinsic epistemic value of truth/true belief 11, 243–53, 265, 266–9
 and normativity 253–62
 and trivial subjects/trivial truths
 objection 247, 248, 249–51, 257, 259, 273, 332–3
 definition of 246
 Goldman's, Alston's, and Lynch's
 teleological accounts 243–51
 restricted/qualified teleological account of the intrinsic value of true belief 248, 249–51
 Sosa's teleological account 251–7
 unrestricted teleological account of the intrinsic value of true belief 247–8, 249–51
 see also value of truth/true belief
 temporal contexts, time-indexed propositions, and tensed propositions 286–90, 294
 testimony 6, 69, 74, 75–7, 79, 80, 86–7, 88
 Theaetetus 97
 Timmons, M. 233–5
 train example 191–4, 196, 198
 Travis, C. 73

- trustworthiness 75–7
- and freedom 75
- truth:
 - and alethic correctness 231–7
 - and error theory 231, 232
 - and expressivism, *see* expressivism
 - as norm of belief 10, 228–30
 - as the (proper) goal of enquiry 1, 4, 10, 12, 158–60, 225–8, 229, 266–9, 284–304, *see also* epistemic value
 - monism; epistemic value pluralism; trivial argument for epistemic value pluralism
 - as value 10, 225–40, *see also* value of truth/true belief
 - concept of 116
 - deflationary view of 10–11
 - intrinsic value of, *see* value of truth/true belief
 - naturalism/non-naturalism about 231
 - scepticism about the value of, *see* expressivism
 - trivial, *see* teleological arguments, and trivial subject/trivial truths objection; value of truth/true beliefs, and trivial truths
 - trivial connection principle 230
 - value of, *see* value of truth/true belief
- truth-like concepts 116–17
- Tuske, J. 91 n.
- ugliness argument 121–28
 - and intuitive evaluation 131–2
 - conceptual version of 122–3, 124, 126–32
 - explanatory versions of 136–7
 - pure ugliness version of 126–7, 128, 136
 - realist version of 123–4, 131
 - recognition of value under analysis (RUVA) assumption 128–32, 133–5
 - reflective evaluation version 126, 127–8, 128–32, 136
- ugly analyses 7
 - and value 112–37, *see also* ugliness argument
- understanding:
 - analysis of and the relevance of ordinary language 97–9, 322–3, 327
 - and Meno requirement 336–8
 - and scientific
 - understanding/idealization 324, 326–9, 341–3
 - as distinct from knowledge 95–110
 - degrees of 324–5, 335, 336
 - explanatory, *see* understanding-why
 - factive vs. non-factive 13, 14, 63–4, 322–9, 334–6, 339, 341–3, 351
 - honorific 325–6, 329, 341–2
 - lucky 99–100, 103
 - objectual 6–7, 95–110, 311, 327, 334–6, 341
 - of indeterministic systems 101–2, 107, 109
 - quasi-factive account of 339, 341–3, 351
 - scientific understanding and idealization 13, 14
 - synonymy of ‘understanding’ and ‘knowledge’ 98–9, 106–7
 - types of 96–102
 - understanding-why 100–1, 104, 106, 107, 108
 - vagueness in understanding 14, 341, 342
 - value of 4, 95–110, 309–12, 322
- Unger, P. 344
- value:
 - automatization of 31–5
 - extrinsic 114, 122
 - fundamental 33
 - instrumental 32–3, 76, 122–3 n.
 - intrinsic 6, 74, 75, 76–7, 114, 122, 134–5, 135–6
 - of justification, *see* justification, value of
 - of knowledge, *see* value of knowledge
 - of true belief, *see* value of truth/true belief
 - of truth, *see* value of truth/true belief
 - of understanding, *see* understanding, value of
 - performance 39
 - practical 2
 - truth-independent 5, 43, 44, 52
 - see also* understanding, value of; value of knowledge; value of truth/true belief
- value-driven epistemology 6, 9–10, 60–91, 139, 155, 204, 311, 331
 - and genealogy 65
 - communitarian value-driven epistemology 60–91
 - Kvanvig’s value-driven epistemology 61–4
- value of knowledge 1, 4,
 - and guiding intuition 42–58
 - and luck 204–21
 - and Meno requirement 332, 334, *see also* understanding, and Meno requirement; value problem, demands on a solution to
 - and practical reasoning 163–82
 - and pragmatic encroachment 9, 183–201
 - and reliabilism 19–40, *see also* reliabilism
 - and surreptitious doxastic goods 139–60
 - and testimony 60–91
 - and trivial argument 284–304
 - and ugly analyses 112–37
- credit account of, *see* credit/credit-worthiness

value of knowledge (*cont.*)

- Kvanvig's 'The Value of Knowledge and the Pursuit of Understanding' 309–51
- vs. value of concept of knowledge 113–15, 313
- vs. value of understanding 95–110, 309–12
- see also* guiding intuition; luck; Meno problem; practical reasoning; pragmatic encroachment; reliabilism; surreptitious goods; swamping problem; testimony; trivial argument; ugly analyses; understanding; value problem; value of truth/true belief
- value of truth/true belief 4, 115–17, 186–190, 225–82
 - and curiosity 265–82
 - and epistemic appraisal 225–304
 - and normativity 253–62
 - and teleological accounts 243–62, *see also* teleological accounts
 - and trivial truths 332–3, *see also* teleological arguments; and trivial subject/trivial truths objection
 - as a common good 259–62
 - as defeasible 339
 - extrinsic 122
 - fundamental value, intrinsicness and domain transcendence 251–7
 - instrumental value of truth 265, 267–9
 - intrinsic value of truth 11, 243–53, 265, 266–9
 - see also* curiosity; reliabilism; swamping problem; teleological accounts
 - scepticism about, *see* expressivism
 - value principle 30
- value problem 155–60
 - and guiding intuition, *see* guiding intuition
 - and ordinary thinking/linguistic behaviour 316, 345–6, *see also* guiding intuition
 - and reliabilism 19–40, *see also* reliabilism
 - as a constraint on the analysis of knowledge, *see* analysis of knowledge
 - demands on a solution to 314, 318–20, 332, *see also* value of knowledge, and Meno requirement
 - primary value problem 3, 183
 - secondary value problem 3, 183, 201
 - shifting questions of 314, 315–16

- value pluralism conception of 57–8
- virtue-theoretic solution to, *see* virtue theory
- see also* Meno problem; swamping problem
- value scepticism 62–3
- van Frassen, B. 140
- van Gogh, V. 117 n.
- van Woudenberg, R. 160 n.
- Velleman, D. 142, 153 n., 229–30, 279–80
- Vendler, Z. 80
- veritism 24
- Vice, S. 160 n.
- virtue theory 2, 3, 7–8, 13, 23, 35–40, 74, 75, 77, 81, 82, 103, 105, 183, 201, 246, 310, 315, 317–20, 333, 347–51
- Vogel, J. 166, 179
- Vuilleumier, P., Armony, J., and Dolan, R. 274 n.
- water analogy 260
- Warfield, F. 130 n., 135 n., 262 n.
- Weatherson, B. 176 n., 195 n., 196 n., 291 n.
- Wedgewood, R. 228, 229, 230
- Weinberg, J. 64 n., 259 n.
- Weinberg, J., Nichols, S., and Stich, S. 181 n.
- Weiner, M. 8–9
- Welbourne, M. 80
- Whitcomb, D. 32, 39, 60 n., 90, 91 n., 182 n.
- Williams, B. 6, 8, 60, 61, 64, 69, 73–5, 76, 77–9, 85, 87, 88, 145–6, 149, 155 n., 233 n., 269 n.
- Williamson, T. 2, 3, 30, 30 n., 90, 103 n., 112–15, 127, 128 n., 133–4, 136–7, 158–9, 171, 172 n., 173 n., 174, 176, 180, 184, 185, 192, 201
- Winch, P. 70
- Wittgenstein, L. 60, 69, 70–2
- Wolterstorff, N. 253 n.
- Wrenn, C. 240 n.
- Wright, C. 240 n.
- Yim, D. 58 n.
- Zagzebski, L. 3, 22, 23–4, 36, 36 n., 42 n., 44 n., 52–3, 56 n., 81–2, 183, 314, 318 n.
- Zamora 91 n.