



OXFORD

Justification without Awareness

MICHAEL BERGMANN

JUSTIFICATION WITHOUT AWARENESS

FOR EDUCATIONAL PURPOSES ONLY

This page intentionally left blank

FOR EDUCATIONAL PURPOSES ONLY

Justification without Awareness

A Defense of Epistemic Externalism

MICHAEL BERGMANN

FOR EDUCATIONAL PURPOSES ONLY

CLARENDON PRESS • OXFORD

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 trademark of Oxford University Press
in the UK and in certain other countries

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

© Michael Bergmann 2006

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

First published 2006

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 this same condition on any acquirer

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging in Publication Data
Data available

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

ISBN 0-19-927574-2 978-0-19-927574-8

1 3 5 7 9 10 8 6 4 2

For my father,
Abram Isaac Bergmann
1939–2001

FOR EDUCATIONAL PURPOSES ONLY

This page intentionally left blank

FOR EDUCATIONAL PURPOSES ONLY

Preface

The goal of this book is to defend externalism about epistemic justification. The first part of the book is devoted to a careful examination and refutation of externalism's main competitors. The aim there is to get the reader to see that these alternatives are dead ends and that the truth lies elsewhere. This will prepare the reader for Part II of the book where I argue in support of my favored externalist position and respond to some influential objections aimed at externalism generally.

The most prominent competitor to externalism is internalism, which is the primary focus of Part I. A crucial ingredient of internalist accounts of justification is the thesis that it is not enough for a belief's justification that it *has* some good-making feature; in addition, the person holding the belief must be *aware* of that good-making feature. The central point of contention between internalists and externalists is whether there is such an awareness requirement on justification.

In the first chapter, after explaining how I'll be thinking of justification, I present a dilemma for all versions of internalism. According to that dilemma, the awareness required for justification can be either strong or weak: if strong awareness is required, vicious regress problems arise; but if weak awareness is required, the main motivation for internalism is lost. Either way we should give up on internalism. This dilemma is similar to a dilemma proposed by Wilfrid Sellars ([1956] 1963) against *foundationalism*. Several internalists have been eager to formulate their internalist views on justification in such a way that the Sellarsian dilemma causes them no trouble. I argue in Chapter 2 that even views specifically designed to escape the similar-sounding Sellarsian dilemma are unable to escape the dilemma I propose in Chapter 1.

In Chapter 3, I focus on a view called 'mentalism' according to which a belief's justification supervenes solely on the believer's mental states in such a way that people who are the same mentally are the same justificationaly. Proponents of this view typically go on to point out that internalism is equivalent to mentalism, noting that this mentalist account of internalism conflicts with the account of internalism that I endorse in Chapter 1 according to which an essential ingredient of internalist positions is the endorsement of an awareness requirement on justification. My goal in this chapter is to argue against both the mentalist account of internalism and mentalism itself. In arguing against the *mentalistic account of internalism*, I accomplish two things: I defend the 'awareness requirement' account of internalism on which my Chapter 1 objection to internalism depends; and I argue that, contrary to the standard assumption that the internalism–externalism distinction is exhaustive, mentalism is distinct

from both internalism and externalism. In arguing against *mentalism*, which is externalism's other main competitor, I further prepare the way for my Part II arguments in support of externalism.

As I already noted, Chapter 1 concludes that internalists are faced either with vicious regress problems or with losing the main motivation for their view. In Chapter 4 I consider an important alternative motivation for internalism that can be used to replace its main motivation. That alternative motivation is *deontologism*, the view that justification should be thought of in terms of concepts like duty, obligation, and blame. I conclude that there is no good argument from deontologism to internalism. However, I go on to explain why it is tempting to be misled into thinking there is.

Having argued against the main competitors to externalism in Part I of the book, I turn in Part II to my arguments in support of externalism. I begin by defending a particular version of externalism, one according to which the two main requirements on justification are that it satisfies a *proper function* condition and that it satisfies a *no-defeater* condition. In Chapter 5 I suggest three ways to improve an evidentialist account of justification so that it becomes a proper function account of justification, one that differs from Plantinga's proper function account of warrant (1993a) in that it imposes no reliability requirement on justified belief.¹ After defending the proper function requirement in Chapter 5, I turn in Chapter 6 to explaining and then defending the no-defeater requirement on justification.

I conclude the book by considering two of the most prominent objections to externalism. In Chapter 7 I examine an objection according to which externalists are committed to approving of *epistemic circularity*, which is supposed to be a very bad thing for externalism. I turn that objection on its head, arguing that everyone—or at least everyone who endorses certain very plausible and widely held views on justification—is committed to approving of *some* kind of epistemic circularity, from which I conclude that epistemic circularity needn't be a bad thing after all. And in Chapter 8, I consider the complaint that externalist responses to skepticism are inadequate in various ways. By way of response, I argue first that opponents of externalism are saddled with the very same problems they attribute to externalists (or to even worse problems). I conclude by arguing that the alleged problems don't provide us with a good reason to reject externalism and, moreover, that we have good reason to prefer externalism to nonexternalist theories of justification.

The organization I've just described comes with a minor cost. In arguing against mentalism in Chapter 3, I draw at several points on my positive arguments in support of the version of externalism defended in Chapter 5.

¹ Warrant, by stipulative definition, is whatever is required to turn true belief into knowledge. In light of Gettier's widely accepted conclusion that justified true belief is not sufficient for knowledge (1963), we can see that warrant is distinct from justification.

Because of this, it would have been convenient in certain respects if the material from Chapter 5 were presented *before* the material from Chapter 3. But doing that would prevent me from clearing away externalism's main competitors in Part I before defending the version of externalism described in Part II. Thus, despite the cost, I opted to preserve the two-part structure outlined above.

Close ancestors of various parts of this book have previously appeared in print. Parts of Chapter 1 and Chapter 2 appeared in 'A Dilemma for Internalism' (2006). Parts of Chapter 5 are from 'Externalist Justification Without Reliability' (2004b) and parts of Chapter 7 are from 'Epistemic Circularity: Malignant and Benign' (2004a). In addition, a few paragraphs from 'Commonsense Naturalism' (2002) appear in Chapter 7 and a few paragraphs from 'BonJour's Dilemma' (forthcoming a) appear in Chapter 6. My thanks to Springer Science+Business Media for permission to reprint material from the first and last of these articles; for permission to reprint material from the other three, my thanks to Blackwell Publishers, *Philosophy and Phenomenological Research*, and Cornell University Press (respectively). I am also very grateful to the Center for Humanistic Studies at Purdue University for a fellowship and to Purdue University for a sabbatical leave; together they provided me with much appreciated financial support while I worked on this project.

I received help from many people in writing this book. My thanks to Evan Fales, Richard Feldman, Patrick Kain, and Timothy McGrew for helpful conversations and to Joseph Cruz, John Greco, Kevin Meeker, Matthias Steup, and Christopher Tucker for their advice on various parts of the manuscript. I'm especially grateful to Jeffrey Brower, Ram Neta, and Alvin Plantinga (each of whom read more than half the manuscript) and to Trenton Merricks and Michael Rea (each of whom read the entire manuscript) for insightful and careful comments from which I benefited greatly. Finally, I wish to thank the two anonymous readers for Oxford University Press for their excellent suggestions and criticisms.

MB

This page intentionally left blank

FOR EDUCATIONAL PURPOSES ONLY

Contents

I. AGAINST INTERNALISM

1. A Dilemma for Internalism	3
1 Epistemic Justification	3
2 Understanding Internalism	9
2.1 What It Is	9
2.2 Why It Is Held	11
3 The Dilemma	13
3.1 An Initial Statement of the Dilemma	13
3.2 Premise (III): Vicious Regress Problems	14
3.3 Premise (IV): The Subject's Perspective Objection	19
3.4 Premise (V): Rejecting Internalism	21
2. No Escape	25
1 Applying the Dilemma to Particular Cases	26
1.1 A Perceptual Belief	26
1.2 An Introspective Belief	28
2 Fumerton	29
3 Infallibilism	32
4 BonJour	35
5 Fales	38
3. Mentalism	45
1 Conee and Feldman	48
1.1 What is CF-Mentalism?	48
1.2 Is CF-Mentalism Equivalent to Internalism?	49
1.3 Is CF-Mentalism True?	57
1.4 Conclusion on Conee and Feldman	64
2 Pollock and Cruz	65
2.1 What is PC-Mentalism?	65
2.2 Is PC-Mentalism Equivalent to Internalism?	70
2.3 Is PC-Mentalism True?	71
2.4 Conclusion on Pollock and Cruz	75

4. Deontologism	77
1 Subjective and Objective Epistemic Duty	78
1.1 Deontologism and Objective Duty	80
1.2 Deontologism and Subjective Duty	81
2 Ginet's Argument	83
2.1 Ginet and Alston	83
2.2 Another Objection to Ginet's Argument	86
3 The Blamelessness Argument	89
3.1 Blameworthiness, Blamelessness, and Justification as Blamelessness	90
3.2 Some Versions of the Blamelessness Argument	93
3.3 Why Deontologism Seems Relevant to Internalism	98
4 Steup's Arguments	99
4.1 The Direct Argument	100
4.2 The Indirect Argument	101
 II. DEFENDING EXTERNALISM 	
5. Proper Function	109
1 From Evidentialism to Proper Function	111
1.1 Evidentialism	111
1.2 The First Improvement: Drop Necessity	114
1.3 The Second Improvement: Add Proper Function	130
1.4 The Third Improvement: Drop Mentalist and Internalist Implications	131
2 A Proper Function Account of Justification	132
2.1 The Proposed Analysis	132
2.2 The Virtues of the Analysis	137
3 Objections	143
3.1 Proper Function and Naturalism	144
3.2 The Supervenience Thesis	146
3.3 Swampman	147
3.4 Skeptical Scenarios	149
6. Defeaters	153
1 Kinds of Defeaters	153
1.1 Propositional vs. Mental State Defeaters	154
1.2 Actual vs. Believed Defeaters	160
2 No-Defeater Conditions	162
2.1 Kinds of No-Defeater Condition	162
2.2 The No-Believed-Defeater Condition	163

3	Objections	168
3.1	Concerns About Explanatory Circularity	168
3.2	Necessity Revisited	169
3.3	Fumerton on Conceptual Regresses	172
3.4	BonJour on Quasi-Externalism	175
7.	Epistemic Circularity	179
1	The 'Epistemic Circularity' Objection to Externalism	179
1.1	Epistemic Circularity	179
1.2	An Objection to Externalism	183
2	The First Argument in Support of Epistemic Circularity	184
2.1	In Defense of Foundationalism	185
2.2	A Foundationalist Defense of Epistemic Circularity	187
2.3	A Dilemma for the 'Epistemic Circularity' Objection to Externalism	192
3	The Second Argument in Support of Epistemic Circularity	193
4	Why Epistemic Circularity Seems Like a Bad Thing When it Isn't	197
4.1	Malignant and Benign Epistemic Circularity	198
4.2	Why Benign Epistemic Circularity is Ignored	200
5	Realistic Benign Epistemic Circularity	206
8.	Responding to Skepticism	213
1	The Externalist Response to Skepticism	213
2	Nonexternalist Responses to Skepticism	218
3	Defending Externalism	225
3.1	First Complaint: 'Uncomfortable Moving Up a Level'	225
3.2	Second Complaint: 'Conditional Answer'	227
3.3	Third Complaint: 'Anything Goes'	229
3.4	Fourth Complaint: 'Philosophically Irresponsible'	233
3.5	The Superiority of Externalism	238
	<i>References</i>	241
	<i>Index</i>	249

This page intentionally left blank

FOR EDUCATIONAL PURPOSES ONLY

PART I

AGAINST INTERNALISM

FOR EDUCATIONAL PURPOSES ONLY

This page intentionally left blank

FOR EDUCATIONAL PURPOSES ONLY

1

A Dilemma for Internalism

For the last thirty years, the debate between internalists and their opponents has been a prominent feature on the epistemological landscape. Traditionalists defend the internalist position whereas proponents of more recent epistemological theories (such as reliabilism, tracking accounts, and proper functionalism) insist that it is false. It is easy, I think, to see the attraction of the central internalist thesis. That thesis may be stated, very roughly, as follows: in order for a person's belief to be epistemically justified, it is not enough that it *has* certain virtues (such as being in accord with the evidence or being caused by the fact that makes it true); in addition, that person must be in some sense *aware* of those virtues. The thought is that if the belief has such virtues but the person holding the belief has absolutely no idea that it has them, then those virtues won't yield justification. However, despite its initial appeal, the idea that there is such an awareness requirement on justification is, I will argue, a mistake. In Section 2 of this chapter I will explain more precisely what I take epistemic internalism to be and what the main motivation for holding it is. Then, in Section 3, I will lay out and defend a troubling dilemma confronting all versions of internalism. But first, in Section 1, I will say something about how I'm thinking of epistemic justification.

1 EPISTEMIC JUSTIFICATION

One of the most persistent debates in contemporary epistemology concerns the requirements for epistemic justification. One explanation of the persistence of this debate is that there is a single epistemic property about which epistemologists are arguing and that there is genuine disagreement about its nature (with the consequence that many contributors to the debate have simply failed to understand that property correctly). Another, more conciliatory explanation is that there are as many epistemic properties as there are positions taken in the debate (and hence that what contributors have failed to understand is they are talking about different properties).¹

I'd like to propose yet another explanation, one that steers a middle course between both extremes. I'd like to admit more than one kind of epistemic justification but not as many as there are positions taken in the debate about what

¹ See Alston (1993a, 2005a) for an amplification of this latter explanation.

justification is. In this section I will try to demarcate one kind of epistemic justification about whose nature there can be and, I think, have been genuine disputes. My view is that although there are different kinds of justification, there is genuine disagreement between internalists and externalists concerning at least some of them.

Before proceeding, however, I want to note that although I will here be making clear what I have in mind in this chapter and the rest of the book when I use the term 'justification', I believe that my argument in this chapter against internalism is applicable to other epistemic properties as well, including knowledge and other kinds of justification that have been of interest to epistemologists. Thus, my purpose in explaining how I use the term 'justification' is *not* to strike a conciliatory note by saying: 'while externalism is true of the sort of justification I have in mind, internalism is true of the sort of justification *internalists* have in mind'. Indeed, I think the sort of justification I'll be highlighting (and with respect to which I'll be arguing that externalism is true and internalism false) just *is* the sort of justification in which many internalists are interested.

Let me begin my attempt to isolate the sort of justification on which I'll be focusing by saying that it is a form of *doxastic* justification, not of *propositional* justification.² Doxastic justification is a property of beliefs. Nearly everyone agrees that it isn't sufficient for a belief's justification that one *has* good reasons or adequate evidence for it. In addition, the belief must be *based on* those reasons or that evidence if it is to count as justified. This latter requirement is called a 'basing requirement'.³ The reason it is so common to think there is a basing requirement on doxastic justification is that it is extremely plausible to think that it matters, for a belief's doxastic justification, that it was formed in the right way.⁴ Propositional justification, by contrast, is a property that is had by a proposition relative to a person. A proposition can have such justification for a person even if the person doesn't believe it or even if she believes it but not for the right reasons. Very roughly, a proposition *p* is propositionally justified for a person *S* just in case *S* has reasons or evidence for *p* such that if *S* were to believe *p* and base her belief that *p* on those reasons or that evidence, her belief that *p* would be doxastically justified.⁵ Externalists tend to use the term 'justification', without any qualifier preceding it, to refer to doxastic justification. But it isn't uncommon for internalists to use the term 'justification', without a modifier, to refer to propositional

² I take this distinction between doxastic and propositional justification from Roderick Firth (1978: 217–20).

³ See Korcz (1997) for a nice review of the literature on the basing requirement.

⁴ In Chapter 3 (Section 1.3.3) I explain why I deny that there is a basing requirement on (doxastic) justification even though I endorse the 'formed in the right way' requirement that motivates it.

⁵ What I call 'doxastic justification', Feldman and Conee call 'well-foundedness' and what I call 'propositional justification' they call 'justification'. See Feldman and Conee (1985: 24) and Feldman (2004: 147–8; 2005: 274–5).

justification, especially when they speak of the justification a proposition has for a person. In this book, I'll use the unmodified term 'justification' to refer to doxastic justification.

Some philosophers think there is a sort of doxastic justification—which we can call 'subjective deontological justification'—that is more or less equivalent to epistemic blamelessness.⁶ They think that if a person isn't *blameworthy* for holding a belief, then that belief has subjective deontological justification: epistemic blamelessness is both sufficient and necessary for such justification. I'm willing to grant that that is one sort of epistemic justification. But it is not the kind in which I'm interested. The kind of justification I have in mind is such that Jones's belief that *p* could have it and Smith's could lack it even if both Jones and Smith are equally epistemically blameless in believing that *p*. Bruce Russell gives the following example to illustrate the point: 'Someone who grows up in a religious society and is taught to listen to the deliverances of an oracle can be epistemically blameless in believing those deliverances even though her belief may not really be supported by the evidence and so is objectively *unjustified*' (2001: 36). As the quotation makes clear, Russell is thinking of a sort of epistemic justification that is more objective than epistemic blamelessness. Feldman and Conee give another example:

A paranoid man might believe without supporting evidence that he is being spied on. This belief might be a result of an uncontrollable desire to be a recipient of special attention. In such a case the belief is clearly epistemically unjustified even if the belief is involuntary and the person cannot alter the process leading to it. . . . The person who believes that he is being spied upon as a result of an uncontrollable desire does not deserve to be blamed for that belief. But there is a fact about the belief's epistemic merit. It is epistemically defective—it is held in the presence of insufficient evidence and is therefore unjustified. (1985: 17)

Or consider a demon victim who is given an irresistible disposition to believe on the basis of *q* and (*q* v *p*) that *p*. Beliefs so formed will be epistemically blameless and yet they seem to be unjustified in some important sense. In each of these examples, we have an illustration of the absence of a sort of justification that is distinct from—and in some sense *more objective than*—subjective deontological justification. It is this more objective sort of justification that I am interested in. But how are we to understand it?

One suggestion might be that it is warrant—that property, whatever it is, that makes the difference between knowledge and mere true belief. For warrant is more objective than subjective deontological justification as defined above.⁷

⁶ See for example Plantinga (1993b: 11–25). Feldman (2001: 89) thinks that this way of understanding epistemic justification is not as common or natural as Plantinga and Alston (1988) say it is.

⁷ And there is some precedent for using the word 'justification' to refer to what I here call 'warrant'. See Chisholm (1982: 43, 1986b: 89).

But the justification I'm interested in was shown by Gettier to be insufficient for warrant.⁸ Thus, justification, as I will be thinking of it, is more objective than subjective deontological justification but still insufficient for warrant. Furthermore, justification as I'll be thinking of it is very widely thought to be necessary for warrant. This isn't a matter of stipulative definition. It is an open question whether this sort of justification is necessary for warrant (though, for the record, I think it is). The sort of justification I have in mind, then, is viewed by most as necessary but not sufficient for warrant. It certainly isn't the same thing as warrant.

Perhaps the best way to think of the justification I have in mind is that it is the sort of *doxastic* justification in which evidentialists like Conee, Feldman, and Russell are interested. Often when evidentialists talk of justification, they have in mind *propositional* justification, the sort of thing a proposition could have (relative to a person) even if it isn't believed and even if it is believed but doesn't satisfy a basing requirement. But they also recognize that there is a property of doxastic justification (or 'well-foundedness' as Conee and Feldman have called it) which is what S's belief B has when its content is propositionally justified for S and S bases B on whatever it is that makes its content propositionally justified. That is the sort of justification that is the focus of this book. Nevertheless, I *disagree* with evidentialists about what is required for that sort of justification. They say such justification is a function of the evidence the subject has and they often add that only mental states count as evidence and that for a subject to *have* evidence, the subject must be aware of it. But I'm assuming that they don't make these things a matter of definition. I'm assuming that I can sensibly talk about that sort of justification even though I don't give an evidentialist analysis of it.

To sum up, the sort of justification I have in mind is *doxastic justification that is more objective than subjective deontological justification (understood as epistemic blamelessness) and yet insufficient for warrant*. When I need to distinguish this sort of justification from other sorts, I'll refer to it as 'ordinary justification', since as I understand the contemporary literature on justification it is this type that much (though admittedly not all) of it is concerned with. And even some that is focused more on some other sort of justification (for example, on propositional justification) has a bearing on ordinary justification. Furthermore, as I read the epistemological literature, some contributors to it are disagreeing about what is necessary and sufficient for ordinary justification. Some think ordinary justification is subject to a reliability requirement; others disagree. Some think it requires epistemic blamelessness; others disagree. Some think it requires propositional justification; others disagree. Some think it has an internalist element; others disagree. Some think it is required for knowledge; others disagree. Some think

⁸ One can read Gettier's paper (1963) thinking of justification as subjective deontological justification or as the sort of justification I'm describing here. Either way his conclusion is persuasively established.

it should be understood deontologically; others disagree. But the main point I want to emphasize is that although I acknowledge that there is more than one kind of justification (and, of course, that people can stipulatively define an endless variety of justification types), there is actual disagreement in the literature about particular sorts of justification, including the sort of justification I've highlighted here.

Given the above-mentioned connection between justification and knowledge (i.e. that justification seems to be necessary for warrant, which together with true belief gives us knowledge), it should come as no surprise that, in addition to disagreeing about what is required for justification, epistemologists also disagree about what is required for knowledge. Here too the question arises: is there genuine disagreement between internalists and externalists about what is required for a single property, knowledge, or is there only the appearance of disagreement because the two sides unwittingly have in mind different properties when speaking of knowledge? I'm inclined to think that, except in cases where philosophers invent special kinds of knowledge by stipulatively defining specific varieties of it, internalists and externalists are often disagreeing about a single property, knowledge. But I won't argue for that admittedly controversial conclusion here. Instead, I will restrict my comments to the sort of knowledge (ordinary knowledge) for which internalists tend to think ordinary justification (highlighted above) is necessary but, due to Gettier problems, insufficient.

As I see it, then, internalists and externalists disagree about what is required for knowledge. As I noted in the introduction to this chapter, internalists think there is an awareness requirement on *justification*. Given how common it is for internalists to think justification is required for knowledge, it is extremely common for internalists to think there is an awareness requirement on *knowledge*. Externalists are more apt to focus their work on knowledge than on justification. And in doing so, they are, as I understand them, very often speaking of ordinary knowledge. It seems, therefore, that one salient point of disagreement between internalists and externalists has to do with whether there is an awareness requirement on ordinary knowledge. In fact, given that externalists are sometimes not very forthcoming about their views on justification, it might be best to think of the main difference between internalists and externalists as follows: internalists think there is an awareness requirement on *knowledge* and externalists think there isn't.⁹ Nevertheless, in order to ensure a strong point of contact with the internalist literature, I will continue to speak of their views on *justification* because that is where the internalists place their focus.

⁹ This is only a rough characterization of the *main* difference between internalists and externalists. See Section 2 of this chapter for a more detailed account of internalism. See also Section 1.2.3 of Chapter 3 where I explain why I think there are views on justification that are neither internalist nor externalist views. And see the introductory paragraphs of Chapter 3 for an account of why I'm optimistic about identifying *some* necessary conditions of being an internalist even though I doubt we can settle definitively what is necessary and sufficient for being an internalist.

Before turning to a more precise description of internalism than I have given so far, I want to consider the objection that, by focusing on ordinary justification, I've set internalism up for failure. After all, doxastic justification, in virtue of including a (causal) basing requirement, seems less likely than propositional justification to be correctly understood in an internalist manner. Likewise, given that ordinary justification is more objective than justification understood as epistemic blamelessness, internalism seems less likely to be true of it. Doesn't this just go to show that, by attacking internalism as it applies to ordinary justification, I am attacking a straw man?

The answer is 'no'. Given what my objection to internalism is, the worry about my focus on doxastic justification is misguided. To see this, note first that the *propositional* justification of the content of a belief B is typically viewed by internalists as being necessary for B's *doxastic* justification. For they tend to think of doxastic justification as requiring propositional justification plus something extra, namely, the satisfaction of a basing requirement. Thus, the concern about my focus on doxastic justification, rather than propositional justification, is produced by the thought that, because doxastic justification requires the satisfaction of an extra *noninternalist* constraint (i.e. the basing requirement), internalism will be false of doxastic justification even though it is true of propositional justification. But this difference between doxastic and propositional justification is irrelevant as far as my objection to internalism is concerned. My objection to internalism about doxastic justification won't be avoided by focusing on a sort of justification that requires the satisfaction of fewer *noninternalist* conditions. For my objection to internalism about doxastic justification has to do with the fact that it requires the satisfaction of a certain *internalist* condition—namely, an awareness condition. By dropping the noninternalist basing requirement while keeping all the internalist constraints on doxastic justification, internalists about propositional justification won't avoid imposing the internalist constraint to which I'll be objecting. So my objection to internalism about doxastic justification applies equally to internalism about propositional justification.

As for the worry about my focus being on a sort of justification that is more objective than subjective deontological justification (thought of as epistemic blamelessness), it is either misguided or needless. It is misguided for those who think subjective deontological justification, like propositional justification, is *necessary* for ordinary justification and, together with the satisfaction of some extra noninternalist conditions, *sufficient* for it. For, given those assumptions, remarks exactly parallel to those made in the previous paragraph in connection with propositional justification apply to subjective deontological justification as well: even if subjective deontological justification isn't subject to the extra noninternalist constraints that apply to ordinary justification, it still requires the satisfaction of the awareness requirement that is the basis of my objection to internalism.

For those who think subjective deontological justification isn't even necessary for ordinary justification, the concern about my focus isn't misguided; even so, it

is needless. For the dilemma for internalism presented below applies equally well to subjective deontological justification understood in terms of epistemic blamelessness, even if that sort of justification isn't necessary for ordinary justification. Moreover, in Chapter 4, I give special attention to justification thought of as epistemic blamelessness, in part to allay worries of the sort alluded to here. Thus, there is no need to be concerned that, by focusing on the more objective doxastic justification described above, I will unfairly aid my case against internalism.

2 UNDERSTANDING INTERNALISM

Now that we have some understanding of what justification and knowledge are, we can turn our attention to internalism, which is a view *about* justification (and, by implication, about knowledge). In this section I'll explain what internalism is and why it seems so plausible to many.

2.1 What It Is

What all forms of internalism have in common is that they require, for a belief's justification, that the person holding the belief be aware (or at least potentially aware) of something contributing to its justification.¹⁰ There are different views on what the subject must be aware of. And there are different views on what sort of awareness is required (though all agree that the awareness must be the kind that involves only armchair reflection).¹¹ But the common denominator is the insistence that if there is no (actual or potential) awareness of anything that might contribute to the belief's justification, the belief is not justified.¹² In what follows, I shall take the following to be the canonical formulation of this requirement:

The Awareness Requirement: S's belief B is justified only if (i) there is something, X, that contributes to the justification of B—e.g. evidence for B or a truth-indicator for B or the satisfaction of some necessary condition of B's justification—and (ii) S is aware (or potentially aware) of X.¹³

¹⁰ I noted above that internalists who think justification is required for knowledge will also think there is an awareness requirement on knowledge. That awareness requirement can be spelled out in terms of warrant (that which makes the difference between knowledge and mere true belief) rather than in terms of justification. According to that requirement, a belief counts as knowledge only if the person holding the belief is aware (or at least potentially aware) of something contributing to its warrant.

¹¹ It is typically assumed that armchair reflection involves only introspection, rational intuition, and memory.

¹² Not all awareness that it is metaphysically possible for a person to have counts as potential awareness. Potential awareness is just awareness that a person is able to have by reflection alone.

¹³ Why require awareness of only *some* justification-contributor rather than of *all* justification-contributors? To require awareness of all justification-contributors is to require (for your belief's

A view counts as a version of internalism only if it endorses this awareness requirement.¹⁴

A full defense of this characterization of internalism will have to wait until Chapter 3.¹⁵ Here I will simply draw attention to the fact that this way of understanding internalism is consistent with what several influential authors say on the topic. According to Laurence Bonjour, internalism may be summarized as the 'idea that the justifying reason for a basic belief, or indeed for any belief, must somehow be cognitively available to the believer himself, within his cognitive grasp or ken'.¹⁶ This appears to be just another way of saying that internalists are those who endorse the awareness requirement. William Alston considers two ways of defining internalism: in terms of perspective and in terms of access. Perspectival internalists, according to Alston, are those who think that what contributes to a belief's justification 'must be something that falls within the subject's ken, something of which the subject has taken note' (1986: 186). And he says that access internalists are those who think that what contributes to a belief's justification 'must be accessible to the subject in some special way, for example, directly accessible or infallibly accessible' (1986: 186). Thus, both access and perspectival internalism, as defined by Alston, seem to involve endorsement of something like the awareness requirement identified above.

Finally, consider Richard Fumerton's characterization of internalism. Unlike Bonjour and Alston, Fumerton doesn't think internalism should be understood in terms of requiring access to the fact that justification conditions have been satisfied.¹⁷ Nevertheless, he does think justification depends on direct

justification) awareness of the fact that your belief has satisfied each of the conditions necessary for its justification. But that means that for every necessary condition C of justification there is another necessary condition C* requiring awareness of the fact that C is satisfied. But that means there will be another necessary condition, C**, requiring awareness of the fact that C* is satisfied, and so on. Thus, to require awareness of *all* justification-contributors will lead automatically to a vicious regress of increasingly complex necessary conditions for justification. See Fumerton (1995: 81).

¹⁴ Three things should be noted here. First, it is important to keep in mind that I've given only a necessary condition of being an internalist, not a sufficient condition. Second, contrary to common practice, I do not think of externalism as being simply the denial of internalism. There are views on the nature of justification that I think of as neither internalist nor externalist. (See Chapter 3, especially Section 1.2.3, for a defense of this way of viewing the matter.) Third, this account of internalism is subtly different from the account of internalism proposed in Bergmann (1997b) and employed in Bergmann (2000a). For a discussion of this difference, see Bergmann (2006: n. 7).

¹⁵ My main targets in that chapter will be Earl Conee, Richard Feldman, John Pollock, and Joseph Cruz, all of whom claim to be defending internalism while denying that there are any awareness requirements on justification. For now, those who are suspicious of my characterization of internalism can think of the argument to be presented in this chapter as an objection to all views according to which there is an awareness requirement of some sort on justification. Even if my argument were to show only that there are no awareness requirements of any sort on justification, it would still establish a significant conclusion.

¹⁶ Bonjour and Sosa (2003: 24). There are numerous passages of a similar nature throughout Bonjour's works.

¹⁷ He thinks of having access to a fact as having (or being able on reflection alone to have) a justified belief that the fact in question obtains. See Fumerton (1995: 63–73).

acquaintance with certain facts that are crucial to a belief's justification. And, in characterizing the most important feature of internalism, he says:

The fundamental difference between externalism and one historically prominent and important form of internalism [i.e. the sort Fumerton is interested in defending] is that the internalist wants to ground all justification on a 'direct confrontation' with reality. In the case of noninferentially justified belief, the internalist wants the fact that makes the belief true 'there before consciousness.' (1995: 83)

According to Fumerton, it's not enough for justification that the belief is caused by the fact that makes it true. In addition, that fact must be 'there before consciousness'. Clearly here too we have a sort of awareness requirement being imposed, even if he is careful to avoid some of the problems connected with requiring certain sorts of access.

Now that we have a fairly good idea of what internalism is, let's return to the question of why the view seems so appealing.

2.2 Why It Is Held

The most influential case for internalism is found in the works of Laurence Bonjour, particularly in his presentation of his famous clairvoyant examples and his commentary on them. In what he considers to be the most powerful example, we are asked to consider a clairvoyant named Norman:

Norman, under certain conditions which usually obtain, is a completely reliable clairvoyant with respect to certain kinds of subject matter. He possesses no evidence or reasons of any kind for or against the general possibility of such a cognitive power or for or against the thesis that he possesses it. One day Norman comes to believe that the President is in New York City, though he has no evidence either for or against this belief. In fact, the belief is true and results from his clairvoyant power under circumstances in which it is completely reliable. (BonJour 1985: 41)

BonJour thinks that it is obvious that Norman in this example lacks justification for his clairvoyant beliefs. And he thinks the problem has to do with Norman's failure to satisfy the awareness requirement on justification. But exactly *why* is this supposed to be a problem? Here is BonJour's answer:

One reason why externalism may seem initially plausible is that if the external relation in question genuinely obtains [i.e. the reliable connection between the belief and the fact that makes it true], then Norman will in fact not go wrong in accepting the belief, and it is, *in a sense*, not an accident that this is so: it would not be an accident from the standpoint of our hypothetical external observer who knows all the relevant facts and laws. But how is this supposed to justify Norman's belief? From his subjective perspective it *is* an accident that the belief is true. And the suggestion here is that the rationality or justifiability of Norman's belief should be judged from Norman's own perspective rather than from one which is unavailable to him. (BonJour 1985: 43–4)

In a more recent work, BonJour expresses the same point by saying that externalists make the mistake of saying a belief is justified even when its truth is due to ‘what appears to be, from the standpoint of the believer, mere cognitive luck’. (BonJour and Sosa 2003: 27) He also notes there that a subject (who is like Norman) is being epistemically irrational ‘in accepting beliefs whose provenance can only be a total mystery to [him], whose status is as far as [he] can tell no different from that of a stray hunch or arbitrary conviction’ (BonJour and Sosa 2003: 32).

I’ll call the objection expressed in the previous paragraph ‘the Subject’s Perspective Objection’ (the SPO). We can state it succinctly as follows:

The Subject’s Perspective Objection: If the subject holding a belief isn’t aware of what that belief has going for it, then she isn’t aware of how its status is any different from a stray hunch or an arbitrary conviction. From that we may conclude that from her perspective it is an accident that her belief is true. And that implies that it isn’t a justified belief.

Notice that the SPO provides a perfectly understandable motivation for endorsing the awareness requirement on justification and, therefore, for being an internalist. The claim is this: if you are aware of the justification-contributor, then it isn’t an accident from your perspective that your belief is true; but if you aren’t aware of the justification-contributor, it *is* an accident from your perspective that your belief is true; and that difference (between its being an accident and its not being an accident) makes a difference in the justification of your belief. The SPO is often leveled against externalism, which doesn’t impose an awareness requirement on justification. But it is also used to motivate internalism because it suggests that the problem identified can be handled only by imposing an awareness requirement.

I believe that, though it is ultimately unsuccessful, the motivation for internalism provided by the SPO is the clearest and most compelling motivation for endorsing the internalist’s awareness requirement on justification.¹⁸ Indeed, my dilemma for internalism relies on the fact that there is no other adequate motivation for internalism to take the place of this one. Clearly, the question of whether I am right about *that* is one worthy of further discussion. But I will put that discussion off for the moment. In later chapters it will get the attention it deserves.¹⁹ We now have before us an account of what internalism is (i.e. an endorsement of the awareness requirement on justification) as well as an account of what makes it attractive (i.e. it avoids the SPO). With these things in mind, we can turn to my presentation of the dilemma for internalism.

¹⁸ Examples of internalists who, like BonJour, seem to approve of this sort of objection to externalism are Fumerton (1995: 116), Lehrer (1990: 162), and Moser (1985: 129).

¹⁹ In Chapter 4 I consider another positive motivation for internalism. And in Chapters 7 and 8 I consider some objections to externalism which could be viewed as motivations for internalism by those who think that internalism avoids those objections.

3 THE DILEMMA

3.1 An Initial Statement of the Dilemma

We should begin our consideration of the dilemma by noting that, because they endorse the awareness requirement, all internalists require for a belief's justification that the person holding the belief has some sort of actual or potential awareness of something contributing to that belief's justification. All such awareness will either involve *conceiving* of the justification-contributor that is the object of awareness as being in some way relevant to the justification or truth of the belief or it won't. Let's say that if it does involve such conceiving, it is *strong* awareness and that if it doesn't, it is *weak* awareness.²⁰ Very briefly, the dilemma facing internalists is this. If they require actual or potential *strong* awareness (i.e. if they are strong internalists), then they are faced with vicious regress problems implying the radical skeptical conclusion that we are literally incapable of forming any justified beliefs. But if they require only actual or potential *weak* awareness (i.e. if they are weak internalists), then the SPO can be used against their view, in which case the main motivation for imposing an awareness requirement in the first place is lost. Either way, the results are disastrous for internalism.

Some readers will note the similarity of the dilemma employed here to the one Paul Moser uses to attack internalist coherentism in his book defending internalist foundationalism (1989: 173–6). It also bears a resemblance to the dilemma used by BonJour against internalist foundationalism, when he was an internalist coherentist.²¹ But although the forced choice between two sorts of awareness is common to all these arguments, the dilemma presented here is different insofar as it is used against *all* versions of internalism.²²

In order to make it perfectly clear how my dilemma is supposed to work, it will be helpful to lay it out more carefully as follows:

A Dilemma for Internalism

- I. An essential feature of internalism is that it makes a subject's actual or potential *awareness* of some justification-contributor a necessary condition for the justification of any belief held by that subject.
- II. The awareness required by internalism is either *strong* awareness or *weak* awareness.

²⁰ Notice that weak awareness can involve conceiving too, so long as it doesn't involve the sort of conceiving that is distinctive of strong awareness. In Bergmann (2006), strong awareness is called 'conceptual₁ awareness' and weak awareness is called 'nonconceptual₁ awareness'.

²¹ See BonJour (1985: ch. 4). BonJour—in BonJour and Sosa (2003: 19)—and Fumerton (1995: 74–5, 93) both find a dilemma like this one in Sellars (1963: 131–2).

²² Stephen Hetherington (1990, 1991) has also attempted to use a dilemma like this one against all versions of internalism. However, for reasons given in the appendix to Bergmann (2006), his attempt to use his dilemma against internalism fails.

- III. If the awareness required by internalism is strong awareness, then internalism has vicious regress problems leading to radical skepticism.
- IV. If the awareness required by internalism is weak awareness, then internalism is vulnerable to the SPO, in which case internalism loses its main motivation for imposing the awareness requirement.
- V. If internalism either leads to radical skepticism or loses its main motivation for imposing the awareness requirement (i.e. avoiding the SPO), then we should not endorse internalism.
- VI. Therefore, we should not endorse internalism.

Given the above definitions of ‘strong awareness’ and ‘weak awareness’, premise (II) is completely unproblematic. Premise (I) could use some defense but, as already noted, I will defer that discussion until Chapter 3, taking its truth for granted in this chapter. This leaves only premises (III), (IV), and (V), which will be defended in the next three subsections.

3.2 Premise (III): Vicious Regress Problems

The claim to be defended here is that a strong awareness requirement on justification gives rise to vicious regress problems leading to radical skepticism. Strong awareness, you will recall, is awareness that involves conceiving of the justification-contributor that is the object of awareness as being in some way relevant to the truth or justification of the relevant belief. In defending premise (III) I will make use of the distinction between doxastic and nondoxastic versions of strong awareness as well as the distinction between actual and potential strong awareness. Doxastic strong awareness is strong awareness that involves the *belief* that the object of awareness is in some way relevant to the truth or justification of the relevant belief. Nondoxastic strong awareness is just strong awareness that isn’t doxastic. And an actual strong awareness requirement demands that the subject actually be aware whereas the potential strong awareness requirement demands only that the subject be able on reflection alone to be aware. I will first argue that actual doxastic strong awareness results in regress problems and skepticism. Then I will argue that moving to merely potential doxastic strong awareness or to nondoxastic strong awareness doesn’t help matters.

In order to require a specific type of awareness for justification, we need to make the second clause of the awareness requirement given above in Section 2.1 less general. Here is how it must be changed so that the awareness it requires is actual doxastic strong awareness:

Actual Doxastic Strong Awareness Requirement (ADSAR): S’s belief B is justified only if (i) there is something, X, that contributes to the justification of B and (ii) S

is actually aware of X in such a way that S justifiably believes that X is in some way relevant to the appropriateness of holding B.²³

The attentive reader will notice that, in addition to specifying that the awareness in question requires an actual belief, ADSAR stipulates that the required belief must be *justified*. The reason for this is that if the internalist has the intuition that merely having a justification-contributor isn't enough—that the subject must also *believe* that she has a justification-contributor—it seems highly doubtful that the internalist will be impressed by the *mere belief* (no matter how unjustified or insane) that the thing of which she is aware is a justification-contributor.

Now clearly ADSAR leads to regress problems.²⁴ In order for S's belief B to be justified, ADSAR says that S must have the further justified belief (with respect to something, X₁, that contributes to the justification of S's belief B) that:

P₁: X₁ is in some way relevant to the appropriateness of holding B.

And according to ADSAR, in order for her belief that P₁ to be justified S must have the further justified belief (with respect to something, X₂, that contributes to the justification of S's belief that P₁) that:

P₂: X₂ is in some way relevant to the appropriateness of believing that 'X₁ is in some way relevant to the appropriateness of holding B'.

And in order for her belief that P₂ to be justified, S must have the further justified belief (with respect to something, X₃, that contributes to the justification of S's belief that P₂) that:

P₃: X₃ is in some way relevant to the appropriateness of believing that 'X₂ is in some way relevant to the appropriateness of believing that 'X₁ is in some way relevant to the appropriateness of holding B''.

And so on. Given ADSAR, therefore, one has a justified belief only if one actually has an infinite number of justified beliefs of ever-increasing complexity. But most of us find it exceedingly difficult even to grasp a proposition like P₅ or P₆ in such a series, much less believe it with justification. Consequently, it's very difficult to see how a supporter of ADSAR could resist the conclusion that none of our beliefs are justified. The very ease with which this skeptical conclusion follows from ADSAR gives us a reason to reject it.²⁵

Let's consider whether it will help to switch from an *actual* strong awareness requirement to the following merely *potential* strong awareness requirement:

²³ Or that X is evidence for B or that X is a truth-indicator for B or that X contributes in some way to B's justification. I will often suppress these sorts of alternatives but they should be assumed.

²⁴ See Moser (1989: 173–6), Fumerton (1995: 64), and Alston (1986: 211) for discussions of this sort of regress problem.

²⁵ Even internalists shy away from imposing such exceedingly high standards for justified belief. See Fumerton (1995: 64) for this sort of reaction.

Potential Doxastic Strong Awareness Requirement (PDSAR): S's belief B is justified only if (i) there is something, X, that contributes to the justification of B and (ii) S is *able on reflection alone* to be aware of X in such a way that S justifiedly believes that X is in some way relevant to the appropriateness of holding B.

Although PDSAR manages to avoid requiring for justification the actual possession of an infinite number of increasingly complicated beliefs, it still leads to trouble. For in order to have the justified belief B, S must be able on reflection alone to justifiedly believe that P_1 . And to justifiedly believe that P_1 , S must be able on reflection alone to justifiedly believe that P_2 . Thus, to justifiedly hold B, S must be able on reflection alone to be able on reflection alone to justifiedly believe that P_2 . Given the plausible assumption that *being able on reflection alone to be able on reflection alone* reduces to *being able on reflection alone*,²⁶ we may conclude that for every P_n in the series, S is justified in her belief B only if she is able on reflection alone to justifiedly believe that P_n . But, as was noted above, one needn't go very far in the series before one comes to a proposition that no human is able to grasp let alone justifiedly believe. So although the *potential awareness* option avoids requiring the actual possession of an infinite number of justified beliefs, it is stuck with implausibly requiring the ability to justifiedly hold beliefs of ever-increasing complexity. Like ADSAR, therefore, PDSAR too has skeptical implications that give us a reason to reject it.

Notice that there are two sorts of regress associated with the above two *doxastic* versions of the strong awareness requirement. First, there is what we might call 'the mental state regress'. In the case of both *actual doxastic* and *potential doxastic* understandings of the strong awareness requirement, what is needed is either an infinite number of beliefs or the potential for that many beliefs. But in addition to the mental state regress, there is also a complexity regress. And it is the latter that I've relied on in my objections to the doxastic versions of the strong awareness requirement. It is because these versions of the strong awareness requirement imply that a belief is justified only if one is able to hold justified beliefs of *ever-increasing complexity* that they are so implausible. Notice the following two consequences of this fact. First, we can't prevent the regress from containing an infinite number of beliefs by claiming that earlier beliefs in the series will appear again later on so that we have a circle instead of an infinitely long chain. For given that each belief in the series is more complicated than the previous one, none will make a second appearance.²⁷ Second, it won't be relevant to point out, in response to this vicious regress charge, that we can tacitly hold an infinite number of *dispositional* beliefs. For the issue is the complexity, not the number, of beliefs.

²⁶ Suppose I'm able (given only that I do *some* reflecting) to be able (given only that I do *some more* reflecting) to do A. From that it follows that I'm able *on reflection alone* to do A.

²⁷ This is a point made by Alston (1986: 211).

At this point, those interested in salvaging a strong awareness requirement on justification might be thinking that the problem has to do with requiring *belief*. Perhaps the solution is to require strong awareness—which by definition involves a certain sort of conceiving—without requiring belief. Is that possible? Can one conceive of X as being F without believing that X is F? Paul Moser distinguishes doxastic (or propositional) awareness of X from mere conceptual awareness of X (1989: 186–7). Propositional awareness of X involves a judgment predicating something of X. Mere conceptual awareness of X involves categorizing X according to some classificatory scheme without any judgment being made. Perhaps the internalist can impose a conceptual awareness requirement on justification without imposing a propositional awareness requirement. In our terminology, perhaps the strong awareness required for justification is nondoxastic rather than doxastic.²⁸

The crucial point to recognize here is that concept application (or conceiving of something as being a certain way) can be correct or incorrect; and it can be justified or unjustified in much the same way that believing can be justified or unjustified.²⁹ Therefore, even if the strong awareness required is *nondoxastic*, the internalist who imposes it is stuck with regress problems involving ever-increasing complexity and, therefore, skepticism. To see this, let's consider an actual nondoxastic strong awareness requirement which applies to justification for both concept application and belief:

Actual Nondoxastic Strong Awareness Requirement (ANSAR): S's belief or concept application, Y, is justified only if (i) there is something, X, that contributes to the justification of Y and (ii) S is aware of X in such a way that S justifiedly applies to X the concept of *being in some way relevant to the appropriateness of Y*.

ANSAR says that the justification of both concept application and belief are subject to a strong awareness requirement. There seems to be no good reason to think the justification of *belief* requires strong awareness if one thinks that the justification of *concept application* does not. Why would an externalist account of the justification of concept application be satisfactory if an externalist account of the justification of belief isn't? Likewise, ANSAR says that the conceptual awareness required for the justification of concept application, like the conceptual awareness required for the justification of belief, must itself be *justified* (i.e. that it must involve *justified* concept application). Again, there seems to be no good reason to demand this in the case of belief but not in the case of concept application. If you think insane or irrational concept application is sufficient, why think concept application is even necessary?

²⁸ One response to this sort of move is to deny that there is any nondoxastic conceptual awareness—to insist that all conceptual awareness involves some sort of belief. Given our purposes here, we can ignore this sort of response since taking it seriously only makes things *worse* for the internalist.

²⁹ Cf. Moser (1989: 80–1) and BonJour in BonJour and Sosa (2003: 20–1).

We can see that ANSAR too gives rise to regress problems. According to ANSAR, S's belief B is justified only if:

A₁: S's application to X₁ (a contributor to the justification of B) of the concept *being in some way relevant to the appropriateness of B*

occurs and is justified. And according to ANSAR, concept application A₁ is justified only if:

A₂: S's application to X₂ (a contributor to the justification of A₁) of the concept *being in some way relevant to the appropriateness of A₁*

occurs and is justified. Likewise, ANSAR says that A₂ is justified only if:

A₃: S's application to X₃ (a contributor to the justification of A₂) of the concept *being in some way relevant to the appropriateness of A₂*

occurs and is justified. And so on.

Now consider the concept that is applied in A₃. Spelled out more fully it is:

being in some way relevant to the appropriateness of S's application to X₂ of the concept 'being in some way relevant to the appropriateness of A₁'

which, spelled out even more fully, is:

being in some way relevant to the appropriateness of S's application to X₂ of the concept 'being in some way relevant to the appropriateness of S's application to X₁ of the concept 'being in some way relevant to the appropriateness of B''.

Not an easily digestible concept. And of course it is only the tip of the iceberg when it comes to the complexity we are facing with this regress. Thus, according to ANSAR, justification for a belief B doesn't depend on the subject having an infinite number of *beliefs* in propositions of ever-increasing complexity. But it does depend on the application of an infinite number of *concepts* of ever-increasing complexity.³⁰ And, for reasons similar to those discussed above in connection with the *potential belief* version of the strong awareness requirement, it won't help to require merely the potential for applying an infinite number of concepts of ever-increasing complexity. That way too lies skepticism.

Thus, the familiar sort of regress problem that is associated with requiring actual awareness in the form of a further justified belief can't be avoided if one insists that the awareness required for justified belief is strong awareness. It won't help to require only *potential* strong awareness. Nor will it help to require only *nondoxastic* strong awareness. The problem is that the strong awareness that is being required for the justification of a concept application or a belief involves

³⁰ Notice that even if for every justification-contributor X_n in the series, X_n = X₁, the series still includes an ever-increasing complexity of concepts applied to X₁.

the application, A, of a concept whose content is at least slightly more complex than the content of the concept application or belief for whose justification A is required.³¹ We may conclude, therefore, that requiring strong awareness—whether potential or actual, whether doxastic or nondoxastic—leads to skepticism-inducing regress problems. This establishes premise (III) of my argument.

3.3 Premise (IV): The Subject's Perspective Objection

A number of internalists are familiar with the regress problems associated with requiring strong awareness. And at least some of these—such as Moser (1985, 1989) and Fumerton (1995)—take the regress problems as grounds for opting for the second horn of my dilemma, arguing in effect that the awareness required for justification is weak awareness. But, as I shall now show, this makes them open to the SPO which is standardly used against externalism and to motivate the awareness requirement in the first place. As a result, the main motivation for being an internalist is lost. In order to defend this charge, I will distinguish two kinds of weak awareness and argue with respect to each that requiring it for justification leaves one vulnerable to the SPO.

Recall that weak awareness is awareness that doesn't involve conceiving of the justification-contributor which is the object of awareness as in some way relevant to the truth or justification of the belief in question. Let's say that 'conceptual weak awareness' is weak awareness that involves the application of some concept or other to the object of awareness—i.e. it involves conceiving of the object of awareness in a certain way.³² Thus, although it does not involve conceiving of the object of awareness as in some way relevant to the truth or justification of the relevant belief, it does involve conceiving of it in some way or other. All other weak awareness I will call 'nonconceptual weak awareness'.³³

³¹ For example, if the belief for whose justification the concept application is required is the belief that *p*, then the content of the required concept application will be something like *being in some way relevant to the appropriateness of that belief that p*.

³² In Bergmann (2006), this sort of awareness was called 'conceptual₂ awareness'.

³³ Thus, nonconceptual weak awareness—what in Bergmann (2006) I called plain old 'non-conceptual awareness'—is awareness that doesn't involve the application of any concepts to the object of awareness. Cows and dogs presumably experience pain of some sort. And presumably these animals are *aware* of such experiences. Yet although they are aware of these experiences, it seems likely that they do not apply any concepts to them. Humans too can be nonconceptually aware of experiences they undergo. The difference is that we are also able to be *conceptually* aware of those experiences (by applying concepts to them) whereas dogs and cows presumably aren't able to be conceptually aware of their experiences. The point here isn't to make dogmatic pronouncements on the mental life of dogs and cows. I'm just working with what I take to be common assumptions in order to give the reader some idea of what nonconceptual awareness is. All that we really need here, to grasp what nonconceptual weak awareness is, is an example of a creature that can be aware of a pain experience without applying concepts to it. Choose your own example if you disagree with the account given here of dogs and cows.

Let's begin by considering what merit there might be in imposing a nonconceptual weak awareness requirement on justification. Suppose someone had an externalist view according to which it is necessary and sufficient for the justification of a belief that it be produced by a reliable belief-forming process. And suppose that Jack's belief B is produced by a belief-forming process token of a relevant process type that is, in fact, reliable (call this process token 'RP').³⁴ Now imagine that a proponent of this reliabilist view sees that the SPO counts against her position and takes that to be a good objection. She decides to add to her account of justification the requirement that S is nonconceptually aware of the reliable process token in question—in this case RP. Will that pacify those who endorse the SPO?

It shouldn't. For since the awareness required is nonconceptual, a person can have the required awareness of RP without conceiving of RP in any way—without categorizing it according to any classificatory scheme. But then Jack can be nonconceptually aware of RP without conceiving of RP as relevant at all to the appropriateness of his belief. According to the SPO, if Jack does not conceive of RP as something relevant to the appropriateness of his belief then, even if RP *is* relevant to the appropriateness of his belief, it is an accident from Jack's perspective that his belief is true. Clearly this supposed problem is not solved by requiring Jack to be *nonconceptually* aware of RP.³⁵

Would it help if we added instead the requirement that Jack has a *conceptual* weak awareness of RP? Here again the answer is 'no'. For Jack could satisfy this sort of requirement simply by being aware of RP and applying some concept or other to it (such as the concept *being a physical process token*). And that means that Jack can have a conceptual weak awareness of RP without conceiving of RP as relevant in any way at all to the appropriateness of his belief B. But then, according to the SPO, even if this added requirement were satisfied, it would

³⁴ Every process token is an instance of *some* reliable process type (it is due, in part, to this fact that reliabilism faces the generality problem—see Conee and Feldman (1998) for an explanation). But reliabilism is committed to the view that for each process token, there is a *relevant* type of which it is an instance—a type whose reliability (or unreliability) determines the justification (or lack thereof) of the belief produced by the belief-forming process token in question.

³⁵ Bonjour himself proposes something very much like this objection to requiring only nonconceptual weak awareness when he says: 'If the direct apprehension [or awareness] of the experience involves no claim or assertion regarding its character, so that who thus has such an apprehension is apparently not thereby aware *that* it has such-and-such features, then in what way is his belief *that* he has an experience with those features justified by that apprehension? The basic belief, after all, *is* judgmental: it has the assertive content that something, in this case a sensory experience, has one set of features rather than one of the various others that it might have had. How can a state whose content does not in any way say or indicate that things are one way rather than another nonetheless provide a reason or any sort of basis for thinking that the propositional content of a belief that they are one specific way is true?' (Bonjour and Sosa 2003: 19). Given Bonjour's sympathy with this objection as well as his concern to avoid the regress problems associated with a strong awareness requirement, one might wonder how he can remain a nonskeptical internalist. In Bergmann (forthcoming a) I argue that in the end he can't and that it is only by constantly shifting his position that he avoids facing that conclusion.

still be an accident from Jack's subjective perspective that B is true. For although Jack applies a concept to RP, he doesn't apply the *right sort* of concept to it. He doesn't apply a concept that involves his conceiving of RP as contributing in some way to B's justification (or as indicating that B is likely to be true or some such thing). The only way to guarantee that he *does* apply such a concept to RP is to have B satisfy a strong awareness requirement. Thus, we are forced to concede that by imposing only a conceptual weak awareness requirement, the internalist is vulnerable to the SPO.

We must conclude, therefore, that satisfying a weak awareness requirement, whether nonconceptual or conceptual, is not enough to prevent a view from falling prey to the SPO. And since avoiding the SPO was the main motivation for imposing an awareness requirement on justification, this establishes premise (IV) of the main argument of this chapter—namely, that if the awareness required for justification is weak awareness, then internalism is vulnerable to the SPO, with the result that the main motivation for being an internalist is lost.

To sum up, we have seen that the main motivation for internalism is to avoid being vulnerable to the SPO. But in order to avoid that objection, one must hold views about justification that imply that no belief is justified unless the person holding it is able to justifiedly believe infinitely many other propositions of ever-increasing complexity (or justifiedly apply infinitely many concepts of ever-increasing complexity). The only way to avoid this implication is to require only weak awareness rather than strong awareness. But in doing so, one violates the very intuitions that motivated internalism in the first place.

3.4 Premise (V): Rejecting Internalism

In the previous two subsections, I've defended the following two premises:

- III. If the awareness required by internalism is strong awareness, then internalism has vicious regress problems leading to radical skepticism.
- IV. If the awareness required by internalism is weak awareness, then internalism is vulnerable to the SPO, in which case internalism loses its main motivation for imposing the awareness requirement.

This brings us to premise (V) according to which internalism is to be rejected if it cannot avoid the negative consequences mentioned in the consequents of (III) and (IV):

- V. If internalism either leads to radical skepticism or loses its main motivation for imposing the awareness requirement (i.e. avoiding the SPO), then we should not endorse internalism.

In order to deny premise (V), one needs to argue either that having radical skeptical implications isn't enough to warrant a rejection of internalism or that losing

the main motivation for imposing an awareness requirement isn't enough to warrant a rejection of internalism. Let's consider each of these options in turn.

In Section 3.2 I argued that if justification requires strong awareness, then internalism leads to regress problems that prevent us from having any justified beliefs. I took these implications to be a problem for internalism. But some internalists are known for looking askance at the suggestion that skeptical implications are sufficient to discredit a view (cf. BonJour 1985: 12–13; Fumerton 1995: 42–3). So why can't the internalist just confess that her internalism leads to the noted regress problems and then use *modus ponens* where I use *modus tollens*, accepting the skeptical consequences rather than rejecting her internalism?

It's worth noting that internalists such as Fumerton (1995: 80–1) resist this sort of response even though they think skeptical implications shouldn't automatically discredit a position.³⁶ And I think there is a good reason for this resistance. It is *one thing* to think that justification clearly requires something one naturally expects it to require, such as a good reason, and then to acknowledge that most of our beliefs lack such a reason. An internalist (open to skepticism) who thought this would insist that if we have no good reasons for any of our beliefs, so much the worse for our beliefs; she wouldn't be inclined to think we don't need good reasons for them after all. But it's *another thing* to think that, in order for any belief to be justified, the person holding it must have (or have the ability to form) an infinite number of beliefs of ever-increasing complexity. For in that sort of case, the problem isn't merely that the requirement has skeptical implications. It's also that the requirement just seems excessive, especially when stated so bluntly. I think this is why the internalist, in response to my discussion of strong awareness, will not (and should not) be inclined to accept the radical skeptical results in question.³⁷

It may be helpful to think of these matters in terms of Chisholm's distinction between particularism and methodism (cf. Chisholm 1982: ch. 5). Particularism is the view that we should use, as a starting point in our attempt to identify the criteria for justification, our knowledge of which beliefs are justified and which aren't. Methodism is the view that we should rely on our knowledge of what the criteria for justification are to determine which beliefs are justified and which aren't. The objection to premise (V) that I'm considering here is a methodist objection. It says that we can just tell that the internalist's proposed criterion for

³⁶ Similarly, BonJour argues that, in some ways of formulating the requirements for justification, there is a kind of overintellectualization that is objectionable. What is objectionable about it is that it 'guarantee[s] vicious regresses in all directions, rendering the operation of the intellect inherently futile' (BonJour and Sosa 2003: 65).

³⁷ The internalist might protest that she didn't originally start with the intuition that justification requires an infinite number of beliefs of ever-increasing complexity—she started with something that sounded more plausible. But my claim is that once she sees that her view leads to that bizarre consequence, she should reject her view. The consequence reveals the excessiveness implicit in the original intuition.

justification (namely, that the subject must have a strong awareness of something contributing to the justification of her belief) is correct and that if that criterion leads to the conclusion that none of our beliefs are justified, we should just accept that skeptical conclusion.

Now perhaps the methodist approach is, on occasion, a good one. But there are times when the extreme absurdity of the consequences implied by the criterion we started with demands that we use *modus tollens* with the particularist rather than use *modus ponens* with the methodist. And this is just such a time. For the skepticism implied by the strong awareness requirement is global. Every belief of every believer is unjustified. This is because every believer (with a finite mind) lacks the ability to grasp an infinite number of propositions or concepts of ever-increasing complexity. It turns out, therefore, that if the strong awareness requirement on justification is correct it is literally impossible for a finite mind to have justified beliefs. Surely this implication is a good reason to reject the proposed criterion. It is doubtful *in excelsis* that the sort of justification epistemologists are trying to analyze is *necessarily* unexemplifiable by the beliefs of finite cognitive subjects such as us.

Furthermore, if the internalist were to accept the consequence that none of our beliefs are justified, her position would be self-defeating in an important way. Her extreme confidence in the correctness of the strong awareness requirement on justification has led her to the conclusion that that very confidence was misplaced—misplaced in the sense that it is attached to a belief she is not justified in holding. But once she concedes to her particularist opponent that that confidence is misplaced, it is difficult to see why anyone should approve of her persistence in endorsing the strong awareness requirement. We can make sense of this sort of persistence in the face of absurd consequences only if the one persisting has extreme confidence in the correctness of the requirement she endorses. But it isn't easy to be understanding when we know that the person with this extreme confidence realizes that that confidence is misplaced.

What about the second option for resisting premise (V)? An internalist might say that even though she's lost the main motivation for being an internalist (i.e. to avoid being vulnerable to the SPO), she can remain an internalist by coming up with other good reasons for endorsing a *weak* awareness requirement on justification. But what could those other good reasons be? In light of the arguments given to this point in the chapter, the crucial question that internalists must answer is this:

Crucial Question: When it comes to the justification of a belief, B, why think that weak awareness of something that isn't conceived of as being the least bit relevant to the truth or justification of B is any more helpful than no awareness at all of what contributes to B's justification?

Weak awareness certainly doesn't improve matters from the subject's perspective. That's why those imposing only a weak awareness requirement fall prey to

the SPO. So why is weak awareness better than no awareness at all? Without a good answer to that question, the weak awareness requirement is not adequately motivated.

Several chapters in this book are intended as critical responses to various answers to the Crucial Question noted above. In Chapter 4, I will consider a *positive* motivation for internalism: the view that epistemic justification is a deontological notion requiring epistemic blamelessness. I will argue that, despite initial appearances to the contrary, this does not constitute an adequate motivation for a weak awareness requirement. In Chapters 7 and 8, I will consider two objections to externalism: that it permits epistemically circular beliefs to be justified and that it doesn't provide an adequate response to skepticism. These objections could be viewed as *negative* motivations for internalism. For internalists might think that their position is sufficiently motivated by the thought that imposing a weak awareness requirement on justification enables one to avoid these two objections that externalists can't escape. In response, I will argue that neither of those objections is avoided by imposing a weak awareness requirement on justification. (I will also argue that those objections don't give us a reason to reject externalism.) My conclusion will be that even after a consideration of these possible additional reasons for imposing an awareness requirement, internalism still seems inadequately motivated.

But before examining any of these alternative motivations for internalism, there are two other matters to attend to first. In Chapter 2, I will consider some attempts to escape the negative consequences of the two horns of my dilemma presented in the first chapter. And then, in Chapter 3, I will defend the first premise of my main argument from this chapter, according to which an essential feature of internalism is that it imposes an awareness requirement on justification.

2

No Escape

Internalists are not unfamiliar with the concerns raised by the dilemma in Chapter 1. This is because there is a famous Sellarsian dilemma which has long concerned internalist foundationalists and inspired internalist coherentists.¹ Internalist foundationalists think that beliefs can be justified noninferentially on the basis of an experience. But according to the Sellarsian dilemma, this will work only if the subject is *aware* of the fact that the experience counts as a reason for the noninferentially justified belief. And this raises a problem. For either the required awareness itself will have propositional or conceptual content or it won't. If it *does*, then it too will need justification, and the regress of justification, which supposedly ended in the noninferentially justified, experience-based belief, will continue. If the required awareness does *not* have any propositional or conceptual content, then it cannot itself confer any justification.²

This dilemma, which is intended as an attack on foundationalism, is not exactly the same as the dilemma I used against internalism in the previous chapter.³ But because the two dilemmas are similar, internalists who think they have adequate responses to the Sellarsian dilemma will be tempted to think those responses can be used to deal with the objection to internalism proposed in Chapter 1. The purpose of this chapter is to argue that, whether or not they have adequate responses to the Sellarsian dilemma, these internalists cannot escape the negative consequences highlighted by my dilemma.

My dilemma has two horns: the strong awareness horn and the weak awareness horn. The latter is by far the most popular horn for internalists to take. This is because most internalists who deal with these sorts of problems are convinced that the regress one faces when imposing a strong awareness requirement are vicious. In Sections 2 and 3 I will be looking at two ways of taking the weak awareness horn of the dilemma. First I will consider the work of Richard Fumerton whose eagerness to avoid regress problems leads him to insist that the required awareness is not conceptual.⁴ Then I'll consider an infallibilist way of

¹ The dilemma is traced to Sellars ([1956] 1963: 131–2).

² This is how BonJour describes the Sellarsian dilemma in BonJour and Sosa (2003: 17–21).

³ For one thing, it targets foundationalism (whether externalist or internalist) whereas mine targets only internalism, not foundationalism *per se* or any sort of externalism.

⁴ For reasons similar to Fumerton's, Moser takes this same horn of the dilemma. See Bergmann (2006) for an explanation of why his views cannot escape the objection I propose in Chapter 1.

taking the weak awareness horn of the dilemma—a way inspired by the work of Timothy McGrew. In Section 4, I'll consider the position of Laurence Bonjour who takes the strong awareness horn of the dilemma, but argues that it doesn't lead to the regresses I mention. And in the final section, I'll turn to the views of Evan Fales who also takes the strong awareness horn. Fales's approach is unusual. Unlike most other internalists, he recognizes that his endorsement of the SPO (the Subject's Perspective Objection) pushes him toward the regresses identified in Chapter 1. Yet instead of turning his back on that objection, he argues instead that those regresses aren't vicious.

But before we consider these attempts by internalists familiar with the Sellarsian dilemma to avoid the difficulties highlighted by the dilemma I proposed in Chapter 1, it will be useful to take a more careful look at how my dilemma applies to particular cases. This will give us a better understanding of the challenge these internalists are trying to address in their work. Internalists, such as Bonjour and Fumerton, who are sensitive to the concerns raised by the Sellarsian dilemma often favor introspective beliefs as especially strong candidates for justified non-inferential beliefs in contingent truths. Thus, to see what these internalists are up against in trying to escape from a dilemma of the sort I lay out in Chapter 1, it will be helpful to explain the consequences of my dilemma using an example of an *introspective* belief. Now it turns out that it is actually quite difficult to think clearly about how that dilemma applies to introspective beliefs. So to make matters a little easier, I will begin by looking first at how the dilemma applies in the easier-to-understand case of a *perceptual* belief and then, with that example of perceptual belief in mind, move on to consider an introspective example. After examining these two cases in Section 1, we will have a better grasp of the sort of problem that Fumerton and others are trying to solve. We can then turn to an evaluation of their proposed solutions in Sections 2–5.

1 APPLYING THE DILEMMA TO PARTICULAR CASES

1.1 A Perceptual Belief

Suppose that Jack forms the belief that there is before him a large spherically shaped object (call this belief 'B1'). And suppose that he forms this belief as follows. He walks into a smallish well-lit blue room that is empty except for a large white ball. As a result of fixing his eyes on the ball in the room, he has a visual experience that is of the same type (color-experience-wise) as the experience that a normal human would have in such circumstances (call this visual experience of a large white ball 'WB'). Jack's having WB causes him to have the reliably formed belief B1. Now, is Jack's belief justified? In order to answer that question, there is another important question that will naturally come to the mind of an internalist: does Jack conceive of WB as relevant in any way to the appropriateness of B1?

The internalist faces a dilemma when she considers such a question in connection with a case like Jack's. As an internalist, she will want to insist that B1's being reliably formed does Jack no epistemic good unless B1 satisfies some sort of awareness requirement. But the awareness requirement must be either a strong awareness requirement or not. If it is a strong awareness requirement, then Jack must be aware of some contributor to the justification of B1 and justifiably apply to it a concept such as *being relevant in some way to the appropriateness of B1*. But that way lies the complexity regress and skepticism. So the internalist is forced instead to require that Jack's belief satisfy some weak awareness requirement (requiring either nonconceptual weak awareness or conceptual weak awareness).

Suppose the internalist decides to insist on merely nonconceptual weak awareness of a contributor to the justification of B1. In the case of Jack, this is to require that he has a nonconceptual weak awareness of WB or perhaps a nonconceptual weak awareness of WB's supporting B1. But clearly B1 could satisfy either of those two requirements even if, according to proponents of the SPO, it's the case that *from Jack's subjective perspective* it is an accident that B1 is true. For since Jack is required only to be *nonconceptually* aware of WB, he needn't conceive of WB as evidence for anything at all. The same comment applies to Jack's nonconceptual awareness of WB's supporting B1. It is a little difficult to make sense of such a nonconceptual awareness. I have some idea of what a nonconceptual awareness of a pain would be like. But I'm not sure I have much of an idea of what a nonconceptual awareness of WB's supporting B1 would be like. However, we know this much: one can have such an awareness without conceiving of WB's supporting B1 *as* WB's supporting B1. But if Jack doesn't conceive of WB as evidence for anything and doesn't conceive of WB's supporting the content of B1 *as* WB's supporting that content, it is hard to see how Jack's subjective perspective on B1 has improved in any way.

Switching to conceptual weak awareness won't help Jack either. For Jack's conceptual awareness of WB will do him no good (according to the supporters of the SPO) unless he applies the *right sort* of concept to it—the sort of concept involved in strong awareness. So long as Jack applies only concepts other than ones like *being indicative of B1's truth* or *being a contributor to B1's justification* or *being in some way relevant to the appropriateness of holding B1*, B1's truth will, according to the SPO, be an accident from Jack's perspective.

In short, because it is possible for Jack to have a weak awareness of an experience like WB without applying (or being able to apply) concepts like *being relevant to the justification or truth of B1* to it, it is clear that weak awareness—even if it is conceptual—doesn't enable B1 to avoid the censure of the supporters of the SPO. Therefore, with respect to a perceptual belief like Jack's, the internalist's only choice is to impose a strong awareness requirement leading to a complexity regress (and skepticism) or to impose a weak awareness requirement (of either the nonconceptual or conceptual variety) which leaves her vulnerable to the SPO.

1.2 An Introspective Belief

As we have just seen, the problem with requiring merely weak awareness for the justification of perceptual beliefs like Jack's is this: one could be (conceptually or nonconceptually) weakly aware of an experience like WB without conceiving of it as *being relevant in some way to the appropriateness of holding B1*. This forces the internalist to impose a strong awareness requirement, which leads to a complexity regress and skepticism. Does this same sort of problem arise in connection with introspective beliefs? Yes. But it isn't easy to *see* that it does because it is difficult to think clearly about the type of example needed to demonstrate that it does. Let's take a careful look at an example of this sort and consider its implications for internalists.

Suppose that Jack is being appeared to redly and that he believes that he is being appeared to redly (call this belief 'B2'). And suppose that there is nothing of which Jack is aware to which he applies a concept like *being in some way relevant to the appropriateness of holding B2*. It's true that he is aware of his being appeared to redly. And perhaps he even conceives of that experience in some way (e.g. as a rather uninteresting experience). But, due to severe cognitive malfunction, he doesn't (and isn't able to) think of that experience as something relevant to the appropriateness of holding B2.

This may seem a little hard to swallow. We can see how, in the case of the perceptual belief discussed earlier, Jack might be conceptually aware of WB, the visual experience of a large white ball, without applying to it the concept of *being in some way relevant to the appropriateness of holding B1* (where B1 is his belief that there is a large spherically shaped object in front of him). But is it even possible for Jack to be conceptually aware of his being appeared to redly without applying to it the concept of *being in some way relevant to the appropriateness of holding B2* (where B2 is his belief that he is being appeared to redly)? I don't see why not. It is one thing to have a conceptual weak awareness of one's being appeared to redly. It's another to apply to one's being appeared to redly a concept like *being in some way relevant to the appropriateness of holding B2* (or *being indicative of the truth of B2* or *contributing to the justification of B2*). Granted, the malfunction involved in having such an awareness without applying (or being able to apply) such a concept would be extreme. But that doesn't show it is impossible. Perhaps, in addition to suffering from some sort of cognitive defect, Jack is also in the grip of a philosophical theory which strongly supports his disinclination to apply to his experience of being appeared to redly a concept like *being in some way relevant to the appropriateness of holding B2*. The philosophical theory (which he accepts on the authority of a philosophically astute but malicious prankster) gets Jack *interested* in denying what seems to us to be utterly obvious. And the cognitive malfunction makes it possible for him to do what we are incapable of

doing, namely, genuinely refraining from a concept application that is so obviously correct.⁵

This example provides us with a case of someone who is appeared to redly, who believes that he is appeared to redly, and who is conceptually aware of his being appeared to redly even though, according to proponents of the SPO, it is an accident from his perspective that his belief that he is being appeared to redly is true. For although Jack is aware of his being appeared to redly and although he applies a concept to the object of that awareness, he doesn't (and isn't able to) apply to it a concept like *being in some way relevant to the appropriateness of holding B2*. It's true that the malfunction involved in this example is extreme. But the point is that it is possible. And the proponent of the SPO is committed to requiring (for the justification of B2) that such malfunction not be actual—she's committed to insisting that B2 is unjustified unless it satisfies a strong awareness requirement.

So even in the case of introspective beliefs, the internalist must choose between the extreme skepticism induced by strong awareness requirements and the vulnerability to the SPO that results from imposing only weak awareness requirements. There's no middle ground.

Now that we've seen how my dilemma applies to some particular beliefs, let's consider what some internalist philosophers have to say in response to problems of the sort generated by the dilemma proposed in Chapter 1. I'll begin with Fumerton and then turn to the work of McGrew, Bonjour, and Fales.

2 FUMERTON

In this section, I want to reconsider the introspective belief described in Section 1.2 in light of Fumerton's views on justification. Because of his concern to avoid regress problems, Fumerton tries to steer clear of conceptual awareness requirements of any kind. He says (1995: 75) that a person's belief that *p* is (noninferentially) justified only if she is directly acquainted with the fact that

⁵ What's important here is to distinguish an awareness of being appeared to redly from both (a) the application of a concept to the object of that awareness and from (b) the belief, B2, that one is being so appeared to. Once these three are distinguished, we need only consider some malfunction (due to brain damage or an evil demon) that prevents the subject from applying (and from being able to apply) to the object of the awareness in question a concept such as *being in some way relevant to the appropriateness of holding B2*. Given that the awareness, the concept application and the belief are three distinct things, I don't see why such malfunction is impossible (at the very least, the suggestion that it is impossible would require some defense). And given the possibility of such malfunction, there is no reason to think the following aren't compossible: (i) S's holding the true belief B2 (that she is being appeared to redly), (ii) S's being aware of her being appeared to redly, and (iii) S's being unable to apply to the object of that awareness a concept like *being in some way relevant to the appropriateness of holding B2*.

p , her thought that p , and the relation of correspondence holding between her thought that p and the fact that p . Concerning this direct acquaintance he says the following:

Acquaintance is *not* another intentional state to be construed as a nonrelational property of the mind. Acquaintance is a *sui generis relation* that holds between a self and a thing, property or fact. To be acquainted with a fact is not *by itself* to have any kind of propositional knowledge or justified belief . . . One can be acquainted with a property or fact without even possessing the conceptual resources to *represent* that fact in thought . . . Acquaintance is a relation that other animals probably bear to properties and even facts . . . (1995: 74–5)

Clearly, he has in mind some sort of nonconceptual weak awareness.

In giving the example of Jack's introspective belief, I've already stipulated that he has a nonconceptual weak awareness of the fact that p (i.e. the fact that he is being appeared to redly). We can add that he is also nonconceptually aware of his thought that he is being appeared to redly. Finally, we can add that he is nonconceptually aware of the relation of correspondence holding between his thought that he is being appeared to redly and the fact that he is being appeared to redly. Again, it is a little difficult to make sense of nonconceptual awareness of things like this relation holding (vs. things like being in pain). But for our purposes, what's important is that we know that this awareness or direct acquaintance is nonconceptual (i.e. it doesn't involve the application of any concepts).

Now what happens to Jack's subjective perspective on B2 when we add to our example that Fumerton's conditions on justification are satisfied by B2? Because the direct acquaintance required is nonconceptual, Jack can be directly acquainted with the fact that he is being appeared to redly *without conceiving of it as being in any way relevant to the justification or truth of B2* (this is because nonconceptual weak awareness is the sort of thing that can occur without the application of *any* concepts to the object of awareness). Furthermore, he can be directly acquainted with the relation of correspondence holding between his thought that he is being appeared to redly and the fact that he is being appeared to redly *even if he has no idea* that the relation of correspondence holds between these two items (again, this is because nonconceptual awareness is the sort of thing that can occur without the application of *any* concepts). Thus, Jack's belief that B2 can satisfy Fumerton's requirements even if he conceives of his being appeared to redly as no more relevant to B2 than is the mild pain in his left knee. It is, therefore, exceedingly difficult to see how these direct acquaintances improve things *from Jack's subjective perspective*. If things were epistemically bleak from his perspective *before* he had these nonconceptual acquaintances Fumerton requires, there is no reason to think they will be less bleak afterwards. The only way for things to improve from Jack's perspective is for him to apply to the objects of these direct acquaintances a concept such as *being relevant in some way to the appropriateness of B2*. But to require that is to impose a strong awareness requirement on

justification, something Fumerton refuses to do, knowing full well that it leads to regress problems. I conclude, therefore, that Fumerton's account of the justification of introspective beliefs is vulnerable to the SPO.

At this point one might be inclined to ask: Can it really be that direct acquaintance of the sort Fumerton requires isn't any better than no awareness at all? According to the proponent of the SPO, this *can* be the case. For direct acquaintance of the sort Fumerton requires doesn't guarantee that the person conceives of the object of acquaintance as in any way relevant to the justification of the relevant belief. It is only if Fumerton assumes that such direct acquaintance entails some sort of justified strong awareness that it seems to the proponent of the SPO any better (as far as justification is concerned) than no awareness at all. But Fumerton is rightly adamant about rejecting such a view, knowing that it leads to radical skepticism.

Why is there this tendency among internalists such as Fumerton to reject strong awareness requirements and to think it is possible, while doing so, to adequately motivate a weak awareness requirement on justification? I think the problem is that they mistakenly fall into accepting something like the following line of reasoning:

1. Everyone who is *in fact* aware, in any way at all, of his or her being appeared to redly, justifiably conceives of that experience (or at least is able to justifiably conceive of it) as being relevant to the truth or justification of the belief that he or she is being appeared to redly.
2. Therefore, weak awareness of one's being appeared to redly *entails* at least the potential for a justified strong awareness of one's being appeared to redly. [from 1]
3. The potential for a justified strong awareness of one's being appeared to redly prevents one from being vulnerable to the SPO.
4. Therefore, weak awareness of one's being appeared to redly prevents one from being vulnerable to the SPO. [from 2 and 3]
5. Therefore, a weak awareness requirement on justification is adequately motivated (since it enables one to avoid being vulnerable to the SPO).

Nothing like this is likely to be explicitly rehearsed in the mind of a weak internalist. Nevertheless, it seems to be something like the slip from 1 to 2 in the above line of reasoning that explains their mistake. Premise 1, if true, is true only because no one *in fact* malfunctions in the severe way Jack does in my Section 1.2 example of his introspective belief. But from that we can't conclude that 2 is true.⁶ For that is just to conclude from the fact that no one *actually* malfunctions in that severe way that it is *impossible* to malfunction in that severe way. I think

⁶ Furthermore, if 2 were true, this would give us a reason to think weak awareness is impossible for all of us with finite minds. For to have weak awareness would involve having the ability (required

it is because the kind of malfunction I mentioned is so severe and uncommon (perhaps it never in fact occurs) that weak internalists mistakenly take for granted that requiring weak awareness guarantees the sort of strong awareness that makes one invulnerable to the SPO.

3 INFALLIBILISM

I now want to consider an infallibilist way of taking the weak awareness horn of my dilemma for internalism. An infallibilist is someone who endorses the following view:

Infallibilism: S's belief that *p* is justified only if it is impossible for S to believe *p* when *p* is false.⁷

Let's begin by considering an infallibilist objection (inspired by the work of Timothy McGrew) to my attack on those taking the weak awareness horn of the dilemma.⁸ The objection I have in mind draws attention to the fact that the introspective belief on which I focused in Section 1.2—i.e., Jack's belief that he is being appeared to redly—is a *fallible* belief because it isn't impossible for Jack to believe that he is being appeared to redly when he isn't being so appeared to.⁹ An internalist taking the weak awareness horn who endorsed infallibilism would want me to focus instead on examples of *infallible* introspective beliefs, saying that if I had done so my anti-internalist argument would fail.

An initial difficulty facing anyone endorsing infallibilism about justification is that it seems to force us toward a fairly radical form of skepticism about the external world. What infallible beliefs do we have that could enable us to form

by having justified strong awareness) to apply an infinite number of concepts of ever-increasing complexity. See Section 3.2 of Chapter 1 for further discussion.

⁷ Notice that the claim here is that infallibility is necessary for justification, not that it is sufficient for justification. The suggestion that infallibility (so understood) is sufficient for justification is implausible. All beliefs in necessary truths are relevantly infallible (i.e. it is impossible to believe them when they're false). But not all such beliefs are justified since one can believe a necessary truth for a lousy reason. Notice too that it is possible to be an infallibilist with respect to noninferential justification while denying infallibilism with respect to inferential justification. This seems to be Timothy McGrew's view (1999: 231). I'll be ignoring that wrinkle here.

⁸ See McGrew (1995, 1999). The objection is only *inspired by* the work of McGrew. He isn't responsible for how the objection is formulated.

⁹ We can see that this belief is fallible as follows. Suppose Jack used to be a normal person who regularly formed the belief that he was being appeared to redly when he *was* being appeared to redly and that, in doing so, he picked up the concept of being appeared to redly. Now suppose that, due to brain damage or demon influence, he begins to believe that he is being appeared to redly when he isn't being so appeared to. He doesn't *hallucinate* his being appeared to redly—i.e. it doesn't *appear* to him that he is being appeared to redly when he isn't being so appeared to. Nor does he base his belief that he is being appeared to redly on his acquaintance with how he's being appeared to. What happens is that even when he isn't being appeared to redly, he sometimes succumbs to an overpowering *inclination*—caused by brain damage or demon influence—to believe that he is. Given that being appeared to redly is one thing and that believing that one is being appeared to redly is another, the example seems to be a possible one.

justified beliefs concerning the world around us? Infallible beliefs in necessary truths wouldn't be enough since facts concerning the external world are, presumably, contingent facts. But what infallible beliefs in contingent truths do we have? We've already pointed out that introspective beliefs such as Jack's belief that he is being appeared to redly are fallible. In response to this sort of worry, McGrew points to referential introspective beliefs such as 'I am experiencing *this*' (where the subject introspectively points to some experience she is having). It is impossible, says McGrew, for someone to form a belief like that (when it is false) for this very simple reason: 'If there were nothing for the term *this* to refer to, it would not be possible to form the belief at all' (McGrew 1999: 228) But even granting that we have such infallible beliefs in contingent truths, it is doubtful in the extreme that we can, on the meager basis of such beliefs together with infallible a priori beliefs in necessary truths, arrive infallibly at any conclusions about the external world.

But suppose we set aside such concerns about skeptical implications. There remains the additional and more relevant worry that infallibilism, as defined above, isn't an internalist view. The reason is that it isn't committed to there being an awareness requirement on justification. Infallibilists will concede that although some introspective beliefs are infallible, others are not.¹⁰ But will the infallibilist think that an introspective belief's being infallible rather than fallible can make it justified rather than unjustified even if the believer is completely unaware of its infallibility? It would seem not—at least not if the infallibilist is also an internalist. Internalist infallibilists will think that infallibility of which the believer isn't aware does no more to justify a belief than a high degree of reliability of which the believer isn't aware.

So let's consider an internalist version of infallibilism, one imposing only a weak awareness requirement so as to avoid the vicious regress that results from imposing a strong awareness requirement. Given the importance of the infallibility requirement to infallibilist internalists, the justification-contributor of which they would require awareness would, presumably, be this very infallibility. This suggests that the infallibilist version of the weak awareness requirement would look something like this:

Infallibilist Weak Awareness Requirement (IWAR): S's belief that *p* is justified only if (i) it is impossible for S to believe that *p* if *p* is false and (ii) S has a weak awareness of the fact that it is impossible for S to believe that *p* if *p* is false.¹¹

¹⁰ Even those who think that my belief that I'm being appeared to redly is infallible will grant the fallibility of my introspective belief that my visual image of a speckled hen has 49 speckles rather than 47 speckles. See Sosa's discussion of the speckled hen example in *BonJour* and Sosa (2003: 121 ff.).

¹¹ Some infallibilists might think it is important to require that this weak awareness is infallible too. Let's say that S's weak awareness of X is infallible iff S's being weakly aware of X entails that X exists at the time S is weakly aware of X. I'm happy to think of the weak awareness mentioned in IWAR as being infallible in this sense.

As the inclusion of the second clause makes clear, IWAR is an internalist view whereas infallibilism, as stated above, is not.

Now suppose Jack believes both that he is experiencing *this* (where the 'this' is referring to some pain to which he is introspectively pointing) and that he is being appeared to redly. And suppose furthermore that he isn't aware of the fallibility or infallibility of either of these beliefs. According to IWAR, the latter appeared-to-redly belief isn't justified since it isn't infallible. But the former referential belief fares no better according to IWAR because, even if it *is* an infallible belief, Jack isn't aware of its infallibility. From Jack's subjective perspective, it is an accident that the referential belief he's holding is infallible and an accident that it is true.

What happens when we change the example so that Jack has a weak awareness of the infallibility of his referential belief, thereby satisfying IWAR? Will that satisfy the proponents of the SPO? Certainly not. One thing we know about a weak awareness of a referential belief's infallibility is that one can have such an awareness without conceiving of that infallibility *as* infallibility—without conceiving of it as in any way relevant to the truth or justification of the referential belief in question. So if it was an accident from Jack's perspective before satisfying IWAR that his infallible belief was true, it remains an accident after he has this weak awareness of the belief's infallibility.

Here one might ask: Can it really be that *infallible* awareness of a belief's infallibility is no better than no awareness at all of that infallibility? According to the proponent of the SPO, this *can* be the case. For infallible awareness that is weak awareness doesn't guarantee that the person conceives of the object of awareness as relevant in some way to the justification of the relevant belief. It is only if one requires that the infallible awareness involve some sort of strong awareness that the proponent of the SPO will grant that it seems any better (as far as justification is concerned) than no awareness at all. But of course imposing a strong awareness requirement of any sort is to be rejected because it leads to radical skepticism.

Those tempted to respond to my dilemma by appealing to infallibilism about justification might be implicitly relying on a line of reasoning similar to the line of reasoning expressed in 1–5 at the end of Section 2. The difference is that instead of making the mistake of inferring 2 from 1, they would be making the mistake of inferring from:

- 1*. Everyone who is *infallibly* aware of the infallibility of one of his or her introspective beliefs (such as, for example, the referential belief that one is experiencing *this*), justifiedly conceives of that belief's infallibility (or at least is *able* to justifiedly conceive of it) as being relevant to the truth or justification of that introspective belief.

that:

- 2*. Therefore, infallible awareness of the infallibility of a belief entails at least the potential for a justified strong awareness of that belief's infallibility.

The problem here is the same as it was with 1 and 2: if 1* is true, it is only a contingently true universal generalization, not a necessary truth. 1* could be false if it applied to cognizers who are malfunctioning in a significantly serious way. So 2* doesn't follow from it.

4 BONJOUR

In the previous two sections, we've been considering the problems that arise when taking the weak awareness horn of my dilemma. Let's turn now to a consideration of the problems that arise when one takes the strong awareness horn. I'll begin by looking at the views of Bonjour who wants to take the strong awareness horn while denying that it leads to the regress problems to which I've said it leads.¹²

It will be helpful to focus on the example Bonjour employs to make his case.

Suppose that I am looking directly at a large, multi-colored, abstract painting, under good lighting conditions, with normally functioning vision and no interfering factors, and at a distance at which the image produced by the painting roughly fills my visual field. I come, for whatever reason, to entertain the propositional claim that within my rather complicated visual field there is a dark green, approximately equilateral, triangular shape. I understand the content of this claim, and, perhaps after some scrutiny, I am able to locate an element in the larger visual experience that seems to clearly and unproblematically fit this description. On this basis, I come to have the corresponding belief, and, according to my account, this belief is justified—justified by my awareness of the descriptive fit between the content of the proposition I believe and the relevant aspect of the content of the experience. (Bonjour: forthcoming)

In this example, Bonjour holds the introspective belief that:

P1: There is a dark green triangular shape within my visual field.

Bonjour identifies (at least) two things that contribute to the justification of his belief that P1. First, there is his visual experience of the green triangular shape, which I'll call 'E1'. Second, there is the descriptive fit between the sensory content of this visual experience—that is, the content of E1—and the propositional

¹² In Bergmann (2006), in which I discuss Bonjour (2001), I interpreted him as taking the weak awareness horn of the dilemma. In Bergmann (forthcoming a), in which I discuss Bonjour and Sosa (2003), I interpreted Bonjour as shifting back and forth between taking the strong awareness horn and the weak awareness horn. In Bonjour (forthcoming) he responds to Bergmann (forthcoming a), making it clear that he wants to take the strong awareness horn.

content of P1.¹³ Strictly speaking, however, as the last sentence of the above quotation makes clear, it's not merely the descriptive fit that justifies P1, but the *awareness* of this fit.

But what sort of awareness must this awareness of descriptive fit be? Must it be strong awareness? Or will weak awareness do? Responding to previous criticism by me, BonJour makes his answer clear:

Bergmann argues at some length that weak awareness will not do, a claim with which I entirely agree. As I would put it, weak awareness simply does not yield anything that would count as a reason for thinking that the belief is true, and so leaves the believer in no better position than he would be on an externalist view. (BonJour: forthcoming)

So, in order for his belief that P1 to be justified, BonJour must have a strong awareness of the descriptive fit between the content of P1 and the content of E1. Moreover, in characterizing this required strong awareness (or *seeing of fit* as he calls it), BonJour acknowledges that it is a propositional judgment, one that admits of justification.¹⁴ But, then, it looks like he has started down the path of the vicious regress. For he insists that his belief that P1 isn't justified unless he has a strong awareness of the descriptive fit between the propositional content of P1 and the sensory content of E1. And this strong awareness involves a propositional judgment that:

P2: P1 descriptively fits E1.

As we have seen, however, BonJour admits that the judgment that P2 (the seeing of fit) can be justified and even insists that it is. How then does he think he avoids the regress?

Let's look carefully at what he says. He begins by saying:

Surely, in the example imagined, I do not need some sort of further, independent justification for thinking that the propositional claim [P1] fits my experience [E1] and so is true. On the contrary, my justification [for the judgment that P2] grows out of my awareness

¹³ The sensory content of E1, according to BonJour, is neither propositional nor conceptual. BonJour recognizes (2001: n. 14) that the term 'content' seems a little strange when applied to an experience which is nonconceptual and nonpropositional but he says that he uses it because he knows of 'no better term for what one is conscious of in having sensory or phenomenal states of consciousness'. But he warns us that 'the two sorts of content [propositional and sensory] are importantly different and should not be conflated'.

¹⁴ We can see this acknowledgement in the following passage: 'Since this seeing of fit cannot be *merely* the acceptance of a proposition, it might seem tempting to hold that it is entirely non-propositional in character. But this cannot be quite right either. For one thing, the cognitive act in question certainly cannot be entirely non-*conceptual*, since it essentially involves the conceptual content of the propositional claim [P1]. And it is also extremely natural to formulate it by saying that I see *that* the content of the experience fits the content of the proposition. . . . The upshot [for this *seeing of fit*] is indeed a propositional judgment, but one whose justification is already present in the cognitive act from which it arises' (BonJour: forthcoming).

of the content of the claim [P1] and of the corresponding experience [E1]. To be sure, I must recognize the fit between the two, but this recognition is not a further, cognitively independent judgment which would then require further independent justification, but is instead *cognitively* guided by and based on those experiences themselves. (BonJour: forthcoming)

What exactly is Bonjour saying here? Is he denying that the required recognition of fit is a propositional judgment? No. As I noted above, he makes it clear elsewhere that he thinks it *is* a propositional judgment. Is he denying that this required propositional judgment that P2 must itself be justified? Surely not. Evidently, what he's denying is simply that this propositional judgment must be justified by something *independent* of his awareness of the content of P1 and E1. Later he returns to these same points, specifically attempting to explain how he can avoid a vicious regress while grasping the strong awareness horn of my dilemma:

But I think, for reasons already explained above, that Bergmann is wrong that a requirement of strong awareness must always lead to a vicious regress. On the contrary, in the situation described earlier in which both the content of the claim being justified [P1] and the experiential content [E1] that the claim is about are present in experience in a way that allows the agreement between them to be directly apprehended, it is possible to be aware of the experience [E1] as being relevant to the justification of the belief [that P1] without the need for this judgment of relevance [P2] to be further justified in some independent way. For this reason, no further regress develops. (BonJour: forthcoming)

Here Bonjour admits that his belief that P1 is justified only if he makes a further judgment that P2. Moreover, he also seems to agree that this further judgment itself needs to be justified, though not *by something independent of his awareness of the contents of E1 and P1*. But he gives us no clear statement of what *is* required for the justification of the judgment that P2. And we are left without an explanation of how the regress is avoided. Indeed, it seems clear that no matter what he requires for the justification of the judgment that P2, the regress *cannot* be avoided. Let me explain.

Given the awareness requirement on justification which Bonjour endorses, in order for his judgment that P2 to be justified there must be something, X, of which he is aware that contributes to its justification. (It seems that he thinks that this something, X, just is the content of P1 and E1, of which he is aware. Let's assume that's right.) Now the crucial question arises: must this awareness of X be strong awareness or will weak awareness do? It seems that Bonjour can't say that weak awareness will do since he's already told us that he agrees that weak awareness is insufficient for justification. So he must insist on strong awareness. But to require strong awareness of X is just to require awareness that involves conceiving of X as somehow relevant to the justification of the judgment that P2.

The content of this strong awareness will be something like:

P3: X (the content of P1 and E1) is in some way relevant to the justification of the judgment that P2.

But this just gives us another judgment in the regress, one that is distinct from P1 and P2 and which is itself in need of justification. The regress has not been stopped.

The only thing Bonjour points to in explaining the justification of his judgment that P2 is his awareness of the contents of E1 and P1. But that awareness by itself isn't enough if, as he claims, strong awareness is required: he must also *conceive* of the contents of P1 and E1 as somehow relevant to the justification of the judgment that P2. Maybe Bonjour thinks that—despite the fact that the judgment that P1 is justified only if he makes a further judgment that P2—the judgment that P2 can be justified without a further judgment like P3. But if so, then he is simply taking the weak awareness horn of my dilemma after all. Worse, he is doing so while claiming not to and after he's made it clear that he thinks it's a mistake to take the weak awareness horn. I conclude that Bonjour has failed to show how he can take the strong awareness horn of my dilemma while avoiding the regress problems I highlight.

5 FALES

In this final section, I shall focus on the work of Evan Fales who, like Bonjour, takes the strong awareness horn of my dilemma. His approach differs from Bonjour's insofar as he acknowledges that taking that horn leads to the complexity regress I've identified. His response is to argue that that regress is not a vicious one.

Before critically evaluating what Fales says, I think it is worth noting that his position has a virtue that most other internalist views lack. As I explained in Section 2.2 of Chapter 1, the best and clearest motivation for the internalist's claim that there is an awareness requirement on justification is the intuition supporting the SPO. But if we continue to accept that intuition, we are faced with the regress problems noted. Most other internalists find this consequence troublesome and make their awareness requirement weak rather than strong so that it doesn't lead to these regresses. But because the imposition of only a weak awareness requirement in place of a strong one is not supported by the SPO, these other internalists are abandoning their *raison d'être*. Fales, on the other hand, preserves the motivation for his internalism and chooses instead to argue that the regress isn't vicious.

Let's begin by getting clear on the way in which Fales takes the strong awareness horn of the dilemma for internalism. I'll be focusing on his discussion of simple a priori beliefs such as the belief that $1 + 1 = 2$ or that *if the disjunction* (P or Q)

*is true and P is false, then Q is true.*¹⁵ His view is that a simple a priori belief such as one of these is noninferentially justified only if its content is transparent to the person holding the belief. Here is the clearest account of transparency he gives:

T: A proposition P is transparent to S iff (i) S grasps P's content infallibly and with full clarity—thereby knowing what would count as a truth-maker for it—and (ii) S sees clearly and infallibly that that truth-maker exists.¹⁶

Being a good internalist, Fales thinks that it isn't sufficient, for the noninferential justification of my a priori belief that P (e.g. that $1 + 1 = 2$), that P *is* transparent to me. In addition, he adds the following strong awareness requirement: I must be at least potentially aware *that* P is transparent to me. And that awareness must itself be justified.¹⁷ Thus, for my simple a priori belief that P to be justified, it must be the case not only that P is transparent to me but also that I am able, on reflection alone, to justifiedly believe that:

P*: P is transparent to me.

Likewise, Fales thinks that in order for me to justifiedly believe P*, it must be the case not only that P* *is* transparent to me but also that I am able, on reflection alone, to justifiedly *believe* that:

P**: P* is transparent to me.

For similar reasons, the justification of my belief that P also requires that I am able, on reflection alone, to justifiedly believe that P** is transparent to me, and so on. Thus, for my belief that P to be justified, I must be able to justifiedly believe an infinite number of propositions of ever-increasing complexity. Fales is fully aware of this fact and of the obvious objection that we don't seem capable of even grasping the fifth or sixth proposition in this series, much less of believing such propositions with justification. Nevertheless, he wants to argue that this regress isn't vicious. How does his argument go?

The guiding thought behind his defense of this regress is that, due to the nature of transparency, the transparency to me of one simple proposition entails the transparency to me of an infinite hierarchy of ever more complex propositions:

Transparency, by its very character must (and does) have a kind of self-sufficiency. To be transparent is to be transparently so; were it otherwise, it would not be a case of genuine transparency. (Fales 1996: 163–4)

¹⁵ See Fales (1996: 160–5).

¹⁶ This is a nearly direct quote from an email exchange with Fales in the fall of 2003, from which I'm drawing with his permission. Here is the direct quote: 'Transparency, just to clarify the idea a bit more, is a matter both of grasping the content of a proposition with full clarity—knowing what would count as a truth-maker for it—and seeing clearly that that truth-maker exists. So I find the proposition that $1 + 1 = 2$ transparent because I understand, infallibly, what is being asserted, *and* because I see, infallibly, the fact that $1 + 1$ does indeed equal 2'.

¹⁷ See Fales (1996: 162).

Here Fales is asserting a principle we can call 'Fales's Principle' ('FP' for short):

FP: Necessarily, if some proposition *P* is transparent to *S*, then it is transparent to *S* that *P is transparent to S* (or at least it would be on reflection).¹⁸

If FP were true, then, given that *some* proposition *P* is transparent to *S*, each additional proposition in the hierarchy (*P**, *P***, etc.) would be both true and—because the next one up in the hierarchy is also true—seen to be true.¹⁹ From this Fales concludes that the truth of FP would guarantee that the regress is not vicious. For if the content of my *a priori* belief is transparent to me, it follows that *P**, *P***, and so on are also true and transparent to me. The truth of *P** guarantees that the content of my belief that *P* is transparent to me. And the transparency of *P** (i.e. the truth of *P***) guarantees that I see that *P** is true. Something similar can be said for each of the other propositions in the series.

The first question to ask is why Fales doesn't conclude instead that the truth of FP shows that propositions cannot be transparent to us. After all, if FP were true and a proposition *were* transparent to us, we'd have an ability that it seems clear we lack, namely, the ability to justifiably believe an infinite number of propositions of ever-increasing complexity. But even if we set that reaction aside, we still want to know why we should think FP is true. Its being true is a crucial ingredient in Fales's account of why the regress isn't vicious. What reasons does he give us for accepting it? He mentions that if FP were false, his very demanding internalist commitments would force him to adopt extreme skepticism.²⁰ But he thinks that the fact that holding a belief *B* enables one to avoid skepticism is a very poor reason for holding *B*. So he obviously doesn't think of *that* as a good reason for FP.²¹

He does say things such as the following:

[T]he transparency (and truth-value) of the proposition that *P** is transparent to me [i.e. *P***] is not distinguishable from the transparency of *P* itself, in spite of the greater formal complexity of [*P***].

There is nothing about the content of either *P** or *P*** that involves any information not already contained in the original act of intuiting the content and truth-value of *P*.²²

But those claims are merely asserted, not defended. And they seem very implausible. It's difficult to take seriously his claim that the transparency of *P* (e.g. 1 +

¹⁸ To simplify the presentation in what follows, I'll often drop this parenthetical qualifier.

¹⁹ Take any proposition *P*_i in the series. The next one up in the series says that *P*_i is transparent to the subject. And, by the definition of transparency given above in T, that entails that the subject sees that *P*_i is true. (Or at least it does if seeing what makes *P* true and seeing that that truth-maker exists entails seeing that *P* is true. I have some doubts about that but I won't press them here.)

²⁰ Fales (1996: 161).

²¹ And if it were a good reason to avoid skepticism, then why not simply deny FP or that *P*'s transparency to *S* is necessary for *S*'s belief that *P* to be noninferentially justified? Those denials would also help him to avoid skepticism.

²² Fales (1996: 163). Although it isn't clear from this quotation, he has made it clear in personal communication that he doesn't think *P** and *P*** are the very same proposition.

1 = 2) is indistinguishable from the transparency of, for example, P^{**} . It is one thing to see infallibly and clearly what would make P true and that its truth-maker exists.²³ And it is another thing to see infallibly and clearly what would make it true that *it is transparent to me that it is transparent to me that P* and that its truth-maker exists. Likewise, it is difficult to take seriously his claim that P^{**} or P^{***} involves no information not already contained in intuiting the truth-value and content of P or P^* . For, again, it seems evident that there is less information (about *me*, for example) in a lower-level proposition such as P or *it is transparent to me that P* than there is in the proposition that *it is transparent to me that it is transparent to me that it is transparent to me that P* .

Fales mentions two parallel regress cases that are supposed to help us see that iterating 'it is transparent to me that' doesn't involve any harmful increase in complexity. The first parallel case involves the repeated iteration of 'and P ' added to the claim that P . He points out that the content of $P \ \& \ P \ \& \ P \ \& \ P$ is no different from and no more difficult to comprehend than the content of $P \ \& \ P$.²⁴ But even granting that, we could sensibly insist that the content of *it is transparent to me that it is transparent to me that it is transparent to me that it is transparent to me that P* is both distinct from and more difficult to grasp than the content of *it is transparent to me that P* . The repetition of 'and P ' simply results in our saying the very same thing again and again. But in the iteration of 'it is transparent to me that' we are saying something different and more complicated each time we prefix it to what we've got.

The other parallel case involves iterations of 'it is true that'. Here he points out that if I believe that P is true, then I also would believe, on reflection, that *it is true that P is true* and that *it is true that it is true that P is true* (and so on). He then says: 'Although the distant members of the truth-hierarchy are likewise too complex and numerous for me to entertain them individually and distinctly, I can nevertheless affirm without hesitation that each of them is true, and hence believed by me. That is because I understand the essential triviality of iterating 'is true'' (Fales 1996: 165). But again, even supposing all that is correct, the point doesn't apply to the case at hand.²⁵ For although it is trivial and obvious that P is true entails that *it is true that P is true* (along with every other member of the truth-hierarchy), things are different with the regress we are concerned with. It is definitely *not* trivial and obvious that *P is transparent to me* entails that *it is*

²³ Recall that seeing these two things is what, according to T, is required for P 's transparency.

²⁴ Fales (1996: 163).

²⁵ And in fact, it doesn't seem that all this is correct. He concludes from the fact that he thinks each proposition in a set of propositions is true that he believes each proposition in that set, even though many in the set are too complex to be entertained by him. But that isn't right. I may think that each proposition in the set of propositions God believes is true without believing each of those propositions. It's one thing to believe *of* a proposition, which is too complex for me to grasp, that it is true. It's another to believe *that* proposition. If I can't entertain a proposition, I can't believe it even if I can believe of it that it's true.

transparent to me that P is transparent to me. To say otherwise is just to assert that FP is trivial and obvious. But the truth of FP is the very thing at issue. We are looking for reasons to accept that principle precisely because it *isn't* obviously true.

Thus far, I've been pointing out that Fales doesn't give us any convincing reasons for thinking FP is true and that none of his examples and supporting comments helps us to see intuitively that FP is true. But matters are worse than that, for we also have strong intuitive support for a proposition entailing the *falsity* of FP. The proposition in question is this:

MP: It is metaphysically possible that both the following are true at once: (i) I clearly and infallibly see what would make P (i.e. $1 + 1 = 2$) true and that its truth-maker exists; (ii) due to human limitations or cognitive defects or the machinations of an evil demon, I am unable to grasp P*** (i.e. *it is transparent to me that it is transparent to me that it is transparent to me that P*).

MP seems intuitively very plausible. But its truth entails the falsity of FP. To see this, consider these two abilities:

- (a) the ability to see clearly and infallibly what would make P (i.e. $1 + 1 = 2$) true and that its truth-maker exists
- (b) the ability (on reflection alone) to see clearly and infallibly what would make P*** (i.e. *it is transparent to me that it is transparent to me that it is transparent to me that P*) true and that its truth-maker exists.

Notice that having ability (b) requires having the ability to grasp P***. Thus, MP entails that it is metaphysically possible to exercise ability (a) while lacking ability (b). But FP entails that if I exercise ability (a) I have ability (b). So MP entails the falsity of FP. We therefore have good reason to reject the principle on which Fales's argument (that the complexity regress isn't vicious) depends.

In response to the objection in the previous paragraph, Fales has asked me to try to get the proposition that $1 + 1 = 2$ clearly before my mind and to see its truth and then to ask myself whether, in grasping its truth, I can't help but grasp the truth of the higher-level claim that I am grasping its truth.²⁶ This request strikes me as misguided in two ways. First, the positive answer he expects (that I can't help grasping the truth of that higher-level claim) isn't one he should expect if he asked me instead to consider whether, in grasping the truth of $1 + 1 = 2$, I can't help grasping the truth of the claim that *I'm grasping that I'm grasping that I'm grasping that I'm grasping that its true that $1 + 1 = 2$* . It seems that I certainly can help grasping the truth of *that* when I grasp the truth of $1 + 1 = 2$. And that spells trouble for FP. Second, even if he were right that, upon grasping the truth of $1 + 1 = 2$, I couldn't help grasping all the corresponding higher-level claims, that isn't relevant to the entailment claim at issue. The question

²⁶ In an email exchange in the fall of 2003, from which I'm drawing with his permission.

isn't whether I in fact am *able* to exercise ability (a) without exercising ability (b). Rather, the question is whether it is *metaphysically possible* for me to exercise ability (a) without having ability (b). To determine the answer to that question, it won't be decisive for me to try to exercise the first ability without exercising the latter. For evidence of my *inability* to do A isn't evidence that it is *metaphysically impossible* that I do A. Instead, I must rely on modal intuition to see if it is *possible* for me to exercise ability (a) without having ability (b). And it seems pretty clear that that *is* possible.

Suppose Fales were to assert:

FP*: the fact that makes P* true also makes P** true—i.e. that the truth-maker for *it is transparent to me that P* is also the truth-maker for *it is transparent to me that it is transparent to me that P*.

Would that help? No, for two reasons. First, merely to assert FP* without argument is unconvincing. Some claims carry their evidence on their sleeve. But FP* isn't one of them. If all Fales can say is 'think hard about FP* and you will see that it is true', I'm afraid that won't do the trick. Second, what I said two paragraphs ago against FP counts against FP* as well. There I noted that we have strong intuitive support for the claim that it is possible for me to exercise ability (a) without having ability (b). But if I exercise ability (a), then P is transparent to me (according to T) and so P* is true. And if I lack ability (b), then P*** is not transparent to me, in which case P**** is false. But FP* entails that the fact that makes P* true also makes P**** true. Thus, because we have strong intuitive support for the claim that it is possible for me to exercise ability (a) while lacking ability (b) and this entails the falsity of FP*, we have good reason to reject FP*.

The purpose of this chapter has been to consider whether internalists who think they have adequate responses to the Sellarsian dilemma can somehow avoid the consequences of the dilemma laid out in Chapter 1. I've examined the work of Fumerton, McGrew, Bonjour, and Fales and argued that they can't escape those consequences. The fact that this dilemma is so effective even against those sensitive to the Sellarsian dilemma is further evidence of its strength.

This page intentionally left blank

FOR EDUCATIONAL PURPOSES ONLY

3

Mentalism

The first premise of the main argument of Chapter 1, so far undefended, is:

- I. An essential feature of internalism is that it makes a subject's actual or potential *awareness* of some justification-contributor a necessary condition for the justification of any belief held by that subject.

But this premise has been denied by supporters of a view called 'mentalism'. I'll be exploring two varieties of mentalism in this chapter: one proposed by Earl Conee and Richard Feldman and the other proposed by John Pollock and Joseph Cruz. What these views have in common is their endorsement of something like:

- M. A belief's justification is a function *solely* of (i) which mental states the subject is in and (ii) which mental states of the subject the belief is based on. (i.e. if two possible subjects are exactly alike mentally and in terms of which of their mental states their beliefs are based on, then they are exactly alike justificationaly.)

In addition to endorsing something like M, the four authors mentioned above insist that endorsement of such a position is sufficient for being an internalist and that it doesn't commit one to imposing an awareness requirement on justification. They endorse both mentalism and the following mentalist account of internalism:

- MI. Internalism is equivalent to mentalism.

Consequently, they consider themselves internalists even though, as they realize, they wouldn't count as internalists by the criterion for internalism laid down in premise (I) above. From this we may conclude that they think premise (I) is false.¹

¹ James Pryor might seem at first glance (2001: 108) to be another epistemologist who opposes premise (I) given that he classifies Conee and Feldman's mentalism as an internalist view. But in fact Pryor defines internalism (2001: 104) as a view that endorses what he calls 'Simple Internalism'—what Conee and Feldman (2001: 233) would call 'accessibilism'—according to which justification supervenes on the subject's accessible states. So Pryor would agree with my premise (I) claim that an essential feature of internalism is its requirement for justification that the subject be at least potentially aware of some justification-contributor. The only reason Pryor *seems* opposed to my premise (I) account of internalism, then, is that he mistakenly thinks mentalism,

In this chapter I'll be defending premise (I) against this objection from proponents of MI. But before doing so, I'd like to consider the consequences for my argument if this premise is false. The consequences would not, I think, be serious. My main argument as it stands is an objection to internalism. But I could recast it as an argument against *awareness* internalism, the view that a subject's actual or potential *awareness* of some justification-contributor is a necessary condition for the justification of any belief held by that subject. If I did this (replacing 'internalism' with 'awareness internalism' throughout the argument), the revised premise (I) would be unobjectionable to adherents of MI; and the rest of the argument could be defended in the same way as before. However, the conclusion of the revised argument would be slightly weaker than the original conclusion: instead of arguing that internalism as such should be rejected, I'd be arguing only that awareness internalism should be rejected.² But that weaker conclusion is, by no means, an insignificant one. It is no small victory to show that awareness internalism is mistaken. Indeed, I think it is a significant blow to internalists simply to show that the Subject's Perspective Objection is misguided.³ If everyone were to accept the Chapter 1 conclusion that it is a mistake to impose an awareness requirement on justification, that in itself would be an important advance in the field of epistemology, even if mentalism were left unscathed.

I could, therefore, be content with establishing the weaker conclusion of the revised argument, namely, that awareness internalism should be rejected. But the purpose of this chapter is to do two additional things: argue against mentalism and defend premise (I) against those who endorse MI. The reason I want to argue against mentalism is that my ultimate goal in this book is not merely to object to internalism but also to defend externalism. And as I see it, mentalism is neither an internalist nor an externalist position.⁴ Hence, in addition to arguing against internalism using the main argument from Chapter 1, I also want to argue against the other main nonexternalist position, which is mentalism. The reason I want to defend my original premise (I) is that—despite the fact that I think that the consequences of its falsity are not serious—I think that premise is true. And given that its truth enables me to establish an even stronger conclusion, I'm interested in explaining *why* I think it's true.

as defined by Conee and Feldman, is a version of accessibilism endorsing what he calls 'Simple Internalism'. But if one thinks justification supervenes on one's *non*accessible mental states, one can endorse mentalism while denying accessibilism.

² The original argument objects to awareness internalism and declares internalism to be equivalent to awareness internalism; it is, therefore, an objection to internalism as such. The revised version considered here simply drops the claim that internalism is equivalent to awareness internalism.

³ The Subject's Perspective Objection is introduced in Section 2.2 of Chapter 1. I argue against it in Chapter 1, Section 3.

⁴ See Section 1.2.3 of this chapter where I defend the conclusion that there can be theories of justification that are neither internalist nor externalist.

In defending premise (I), I'll be arguing about how best to understand the technical term 'internalism' that was introduced by specialists in epistemology only thirty years ago. I expect that such an undertaking will strike some readers as being misguided. Why can't epistemologists use such technical terms however they like to use them, stipulating whatever meaning they please? I grant that there is *some* sense behind this sort of worry. Nevertheless, I think the term 'internalism', as it is used in the field of epistemology, has come to be associated with certain views and that it would be a mistake to use the term in a way that ignores these associations. Furthermore, I think that although it might be unreasonable to expect that we can, with any confidence, come up with a complete list of *necessary and sufficient* conditions for being an internalist, it isn't unreasonable to expect that we can identify some very plausible *necessary* conditions for being an internalist. For example, it is entirely reasonable (even if not very helpful) to insist that an essential feature of being an internalist in epistemology is that one reject the view that *a belief is justified if and only if it is formed in a reliable way*. Now it turns out that, as I already noted above, my disagreement with proponents of MI comes down to a disagreement about whether a certain condition—namely, that one imposes an awareness requirement on justification—is *necessary* for being an internalist. Given that we do seem to be able to identify some necessary conditions for being an internalist, we needn't assume from the start that, because 'internalism' is a technical term, it is impossible to settle such a disagreement in a reasonable way.⁵

This chapter will be divided into two main sections: the first section will focus on the mentalism of Conee and Feldman (CF-mentalism) and the second will be devoted to the mentalism of Pollock and Cruz (PC-mentalism). In each section I will ask the following three questions of the version of mentalism under discussion:

1. What does this version of mentalism say?
2. Is it equivalent to internalism?
3. Is it true?

My responses to question (2)—one response for each section—will together constitute my defense of premise (I) from the first chapter against the objection from proponents of MI. And my responses to question (3)—again, one for each section—will together constitute my case against mentalism.

⁵ Feldman (2004: 150) is likewise not entirely pessimistic about making at least some progress in settling the question of how best to understand internalism. Here's what he says: 'It is also worth emphasizing at the outset of this discussion that the words 'internalism' and 'externalism' have not been used uniformly and any decision about their usage here is to some extent stipulated. Still, I think that the origins of the debate provide some basis for rejecting the more restrictive accounts of what internalism is'.

1 CONEE AND FELDMAN

1.1 What is CF-Mentalism?

To my knowledge, Conee and Feldman (2001: 233) were the first to use the term ‘mentalism’ in the way I’m using it in this chapter. However, their definition of ‘mentalism’ differs slightly from the definition I gave above. I said above that mentalism is the view that:

- M. A belief’s justification is a function *solely* of (i) which mental states the subject is in and (ii) which mental states of the subject the belief is based on. (I.e. if two possible subjects are exactly alike mentally and in terms of which of their mental states their beliefs are based on, then they are exactly alike justificationally.)

Conee and Feldman, on the other hand, offer the following account of mentalism: ‘If any two possible individuals are exactly alike mentally, then they are alike justificationally, e.g. the same beliefs are justified for them to the same extent’ (2001: 234). The two definitions might seem different because M makes justification a function of two things—what the believer is like mentally and what her beliefs are based on—whereas Conee and Feldman say it is a function only of the first.⁶ But this apparent difference is not really a difference at all. For M is speaking of doxastic justification while Conee and Feldman are speaking of propositional justification.⁷ And Conee and Feldman think a belief has doxastic justification (what they call ‘well-foundedness’) just in case it has propositional justification and it is properly based.⁸ Thus they would agree that *doxastic* justification is a function of both what the believer is like mentally and what her beliefs are based on (and nothing else). We can, if we like, think of M as providing a mentalist account of *doxastic* justification corresponding to Conee and Feldman’s mentalist account of *propositional* justification. Since the sort of justification that I’ll be focusing on in this book is doxastic justification, it is M in which we are interested.

The answer to our first question, then, is that CF-Mentalism is just M where the justification mentioned in M is doxastic justification.

⁶ But isn’t what a belief is based on *included* in what the believer is like mentally? Well, perhaps Conee and Feldman would say it is. If so, then the concern I raise isn’t a concern at all. The reason I consider it as a possible concern is that Conee and Feldman say that facts about causal connections aren’t mental factors (2001: 234). And given that facts about basing are widely considered to be, at least in part, facts about causal connections, I thought some might be inclined to think facts about basing aren’t mental factors.

⁷ See my discussion (in Section 1 of Chapter 1) of the distinction between doxastic and propositional justification.

⁸ See Feldman and Conee (1985: 24) and Feldman (2004: 147–8; 2005: 274–5).

1.2 Is CF-Mentalism Equivalent to Internalism?

I said in the introduction to this chapter that Conee and Feldman equate internalism with mentalism. I also said that they deny that mentalism imposes an awareness requirement on justification. Here are their words:

We find two distinct but closely related characterizations of internalism in [the literature]. One characterization uses a notion of access. What we shall call ‘accessibilism’ holds that the epistemic justification of a person’s belief is determined by things to which the person has some special sort of access. . . . [Another characterization] suggest[s] that internalism is the view that a person’s beliefs are justified only by things that are internal to the person’s mental life. We shall call this version of internalism ‘mentalism.’ A mentalist theory may assert that justification is determined entirely by occurrent mental factors, or by dispositional ones as well. As long as the things that are said to contribute to justification are in the person’s mind, the view qualifies as a version of mentalism.

We think it likely that philosophers have not separated mentalism from accessibilism because they have tacitly assumed that the extensions of the two do not differ in any significant way. They have assumed that the special kind of access on which many internalist theories rely can reach only mental items and perhaps all mental items, or at least all that might be counted as playing a role in justification.

We think that simplicity and clarity are best served by understanding internalism as mentalism. (2001: 233)⁹

Feldman and Conee here distinguish what they say are ‘two distinct’ characterizations of internalism. According to the first, internalism is equivalent to accessibilism (the view that justification is determined by things relevantly accessible to the subject). According to the second, internalism is equivalent to mentalism (the view that justification is determined by the subject’s mental states).¹⁰ Conee and Feldman opt for the mentalist account of internalism (i.e. MI) over the

⁹ In this passage Feldman and Conee say that it is likely that the reason philosophers haven’t separated mentalism from accessibilism is that they have assumed something like the thesis that necessarily all and only mental items are accessible items. But even if it were true that necessarily all and only mental items are accessible items, mentalism and accessibilism would still be distinct views. For a person could still sensibly endorse one view and reject the other so long as she could sensibly reject the thesis that necessarily all and only mental items are accessible items. The most you could conclude concerning these views, from the premise that necessarily all and only mental items are accessible items, is that these two distinct views (mentalism and accessibilism) have the same implications about which beliefs are justified and which aren’t.

¹⁰ In another paper, authored alone, Feldman (2004: 150) gives the impression that he wants to remain agnostic about whether these are two distinct accounts of internalism. But the consideration he mentions there lends no support to such agnosticism. He points out that he’s unsure about whether all mental items are accessible. But even if all mental items were accessible, that would show at best only that (perhaps unbeknownst to their adherents) accessibilism and mentalism have the same implications about which beliefs are justified. As I noted in the previous footnote, it wouldn’t show that accessibilism and mentalism are the *same view*. Those who sensibly denied that all mental items are accessible could sensibly affirm mentalism while rejecting accessibilism. Nor would it show that the mentalist account of internalism is equivalent to the accessibilist account of it. Given that a view could count as a mentalist view without counting as an accessibilist view, it could qualify as an internalist view according to a mentalist account of internalism while failing

accessibilist account. Thus, although they insist that to be an internalist one must be a mentalist, they don't demand that, to be an internalist, one must impose an awareness requirement on justification.¹¹

The question that concerns us here is whether Conee and Feldman are right that MI is to be preferred to the 'awareness requirement' account of internalism given in premise (I) from Chapter 1. Before explaining why I think they are *mistaken* about this, I'd like to consider their reasons for endorsing MI. In order to understand their reasons, it will be helpful to lay out a framework for comparing theories of justification. We can then use that framework to see which theories would be classified as internalist theories by those who equate internalism with mentalism. Likewise, we can use that framework to see how that classification differs from the classification implied by my 'awareness requirement' understanding of internalism. Ultimately, this will put us in a better position to understand and evaluate the reasons Conee and Feldman offer in support of MI.

1.2.1 A Framework for Comparing Theories of Justification

The framework I have in mind takes for granted that the following assumption is accepted by all theories of epistemic justification worthy of our attention, whether internalist or externalist:

to count as an internalist view according to an accessibilist account of internalism. So *if* indeed Feldman is proposing (in Feldman 2004: 150) a reason to give up his earlier claim with Conee (just quoted in the text) that these two accounts of internalism are distinct, he seems to have made a mistake.

¹¹ Feldman does however seem attracted to awareness requirements. In Feldman (2003: 142 ff.) he worries about the skeptical challenge according to which people's perceptual evidence doesn't provide a better reason to believe ordinary external-world propositions than to believe rival skeptical hypotheses (such as the evil demon hypothesis or the brain-in-a-vat hypothesis). He considers with approval (2003: 148–51) the proposal that our perceptual evidence *is* better explained by the hypothesis that our sense experience is caused by physical objects stimulating our sense organs (call this 'the standard hypothesis') than it is by the rival skeptical hypotheses. But then he wonders whether this is enough for the justification of our perceptual beliefs in ordinary external-world propositions. He asks (2003: 150) whether 'it is enough for the explanation in fact to be the best explanation, or does the person have to realize that it is the best explanation?' His suggested solution (2003: 150–1) is to require, for the justification of ordinary external-world propositions, not only that the standard hypothesis *be* the best explanation of our perceptual evidence but that the subject is *justified in believing* the higher-level proposition that the standard hypothesis is the best explanation. It is clear that he has in mind here the requirement that the subject has *propositional* justification for that higher-level proposition, not that the subject believe that higher-level proposition with *doxastic* justification. But he concludes (2003: 156, n. 25) by saying that, although this required propositional justification for a higher-level proposition is enough for the *propositional* justification of the ordinary external-world propositions, it isn't enough for *doxastically* justified belief in those ordinary external-world propositions. He doesn't say what would be sufficient. But once we see that the standard hypothesis's *being* the best explanation isn't enough and that *having propositional justification* for the claim that the standard hypothesis is the best explanation also isn't enough, it really seems that Feldman thinks what is required is an *awareness* of the standard hypothesis's being the best explanation. This is why I say that he seems inclined to impose an awareness requirement on justification.

Assumption: A belief is justified only if it is held in the right way. This requires, at the very least, that the belief is nondeviantly caused by an appropriate input to the belief-forming process.¹²

Lest this assumption seem too externalist in character, let me point out again that the justification in question is *doxastic* justification. Even internalists agree that a belief isn't justified in *that* sense unless it is appropriately based—something that has to do with how the belief is formed or held.¹³

With Assumption before us, we can state the two questions to be used in placing theories of justification into the framework I have in mind:

- (A) What sorts of things can be appropriate inputs to our belief-forming processes?
- (B) Must the appropriateness of such an input (for leading to the belief to which it leads) be due to an essential property of that input?

Here are some possible answers to question (A) ranging from the more internal to the less internal:

- only states of which the subject can, on reflection alone, be strongly aware can count as appropriate inputs.
- only states of which the subject can, on reflection alone, be aware (strongly or weakly) can count as appropriate inputs.
- only mental states of the subject can count as appropriate inputs.

¹² Three features of Assumption require further comment: (i) The belief-forming process I have in mind is the process token that is responsible for the belief's *now* being held by the subject. This may be different from the one responsible for the belief's *first* being held by the subject. (ii) To say an input is an appropriate input to the belief-forming process is to say that it is appropriate for such an input to lead to the belief to which it in fact leads. (iii) The point of the nondeviance requirement is to emphasize that not just any sort of causing will do. To get an idea of what is required for nondeviance, consider the following example of deviant causation given by Plantinga (1993b: 69, n. 8): 'Suddenly seeing Sylvia, I form the belief that I see her; as a result I become rattled and drop my cup of tea, scalding my leg. I then form the belief that my leg hurts; but though the former belief is a (part) cause of the latter, it is not the case that I accept the latter on the evidential basis of the former'.

The belief about seeing Sylvia causes the belief about leg pain but not in a nondeviant way. It is no easy matter to say what exactly is required for nondeviance. Fortunately, I can ignore that difficulty here.

¹³ The view that the basing relation holds between a belief and its basis (or ground) *only if* that basis *causes* the belief is the dominant view, though it isn't universally held. I realize, therefore, that some epistemologists will think Assumption is false. Nevertheless, because (i) Assumption is widely held and, in my judgment, extremely plausible, (ii) it makes possible the very useful framework for comparison that I will be introducing here, and (iii) responding carefully to those who reject Assumption would take us too far afield, I will simply ignore the objections of the nay-sayers. Those interested in the question of whether there is a causal element to the basing relation should consult Korcz (2000) and Mittag (2002), being careful to distinguish our question of whether nondeviant causation is *necessary* for basing from the distinct question of whether nondeviant causation is *sufficient* for basing. For a detailed survey of the literature on the topic, see Korcz (1997).

- only states of the subject (mental or nonmental) can count as appropriate inputs.
- even things completely external to the subject can count as appropriate inputs.

In response to question (B), the answer will be either ‘yes’ or—if the appropriateness of an input (for leading to the belief to which it leads) can be due, even in part, to a contingent property of the input—‘no’.

To illustrate how these two questions can be used to compare theories of justification, let’s consider how a mentalist and a reliabilist would answer them. In response to question (A), mentalists say that only mental states of the subject can count as appropriate inputs. They think of these input states as belief grounds—i.e. as things on which beliefs are *based*. And they think of the nondeviant causation mentioned in Assumption as the nondeviant causation that is required for the basing relation to hold between the mental states which are inputs and the beliefs based on them. In short, the mentalist agrees with Assumption because she thinks *a belief is justified only if it is based on an appropriate mental state*. For the mentalist, then, question (B) asks whether the appropriateness of such a mental state must be due to an essential property of it.¹⁴ And the mentalist answer to this question is ‘yes’. For if it were a contingent feature, then two possible individuals could be in the same mental states and be basing their beliefs on the same things and yet differ justificationaly.¹⁵

Let’s consider next how reliabilists would answer the two questions. In response to question (A), many reliabilists would allow that an appropriate input could be something external to the person, so long as that thing nondeviantly caused the belief in a reliable way. For example, many reliabilists would allow that (if God exists) it is possible for people to know things via divine revelation. Such revelation could result from God *directly* (and reliably) causing true beliefs to be formed in the believers, without using states internal to the believers as intermediate causes of those beliefs. The input that is nondeviantly causing these beliefs is some state or activity in God, which we can plausibly think of as being external to the believer.¹⁶ In answer to question (B), reliabilists will say that the

¹⁴ When I speak of the appropriateness of an input, I have in mind its appropriateness *for leading to the belief to which it leads via the belief-forming process to which it is an input*. I won’t spell that out each time I talk about the appropriateness of an input but it should be understood as implicitly assumed.

¹⁵ Notice that, according to mentalism, what’s essential to the mental state which is the basis is that it’s appropriate (for leading, in the right way, to the belief to which it leads—see the previous footnote), not that it in fact *is* the basis for a belief with that content.

¹⁶ Of course in ordinary run-of-the-mill cases of perceptual belief, there will typically be some other mental state of the subject in the causal history of the belief. But the point is just that many reliabilists would not *require* the input state to be a mental state. The fact that reliabilists would allow for the possibility of knowledge in the way described in the divine revelation example shows this.

appropriateness of the input depends on a contingent property of it, namely, it's being the front end of a *reliable* belief-forming process.¹⁷

Thus, the mentalist and the reliabilist can agree on Assumption while answering the two questions above very differently. We can think of Assumption along with those two questions as providing a framework for comparing theories of justification. To place such theories within the framework, we need only consider how the theories being considered answer those two questions. In addition to helping us compare *theories of justification*, we can also use the framework to help us compare *accounts of internalism* by considering which theories in the framework are classified by those accounts as internalist theories. Consider first MI. According to it, every theory of justification that answers question (A) by requiring appropriate inputs to be mental states and answers question (B) by insisting that the appropriateness of an input *must* be due to its essential properties is an internalist theory. Consider next an accessibilist account of internalism. According to it, internalist theories are those that answer question (A) by requiring appropriate inputs to be *accessible* states and which give a positive answer to question (B), saying that such states count as appropriate or not in virtue of their *essential* properties (maybe because it is assumed that only *those* properties are relevantly accessible—i.e. accessible by reflection alone). Finally, consider the 'awareness requirement' account of internalism proposed in premise (I) from Chapter 1. According to it, what matters for being an internalist theory is whether *some* justification-contributor—either the input itself or the fact that it is appropriate—must be (at least potentially) an object of the subject's awareness. If we suppose that the appropriateness of the input won't be accessible to the subject on reflection alone if the input itself isn't, then this account of internalism would say that internalist theories are those that answer question (A) by requiring appropriate inputs to be accessible states.¹⁸

¹⁷ This particular example of divine revelation *might* be viewed as one in which the input is *essentially* reliable (given God's essential goodness and omniscience). Nevertheless, it is reliability, not essential reliability, that the reliabilist requires. And in most cases such reliability is contingent.

It's worth noting here that, by requiring that the belief-forming process token leading from the input to the belief is reliable, it seems one is requiring the reliability of the *relevant* type of which that process token is an instance. And as Conee and Feldman (1998) have pointed out, it's no easy matter to identify what the relevant type is (this is the Generality Problem for reliabilism). But for my purposes here, nothing hangs on this since nothing here hangs on the truth of reliabilism. And in fact, my own account of justification defended in Chapter 5 is not a reliabilist account of justification (I consider that to be one of its virtues). So I don't need to solve the Generality Problem, though I'll register my sympathy with William Alston's attempt to solve this problem (1995, 2005a: ch. 6), despite the attacks on his solution by Conee and Feldman (1998).

¹⁸ Since the 'awareness requirement' account gives only a necessary condition for being an internalist (rather than a set of conditions that are together necessary and sufficient), it is silent about whether a theory must give an affirmative answer to question (B) if it is to count as an internalist theory. Moreover, it differs from an accessibilist account of internalism in that although it says that to be an internalist one must impose an (actual or potential) awareness requirement on justification, it doesn't insist that to be an internalist one must say that justification is a function *solely* of one's accessible states.

1.2.2 Conee and Feldman's Reasons for Equating Internalism with CF-Mentalism

Now that we have before us this framework for comparing theories of justification and accounts of internalism, we can consider the reasons Conee and Feldman offer for preferring MI to my 'awareness requirement' account of internalism from premise (I).

In their co-authored work, they don't have much to say in support of MI. As I noted above, they say that 'simplicity and clarity are best served by understanding internalism as mentalism' (2001: 233). And apparently in defense of this they say: 'One advantage of our way of understanding the distinction between internalism and externalism in epistemology is that it closely parallels the counterpart distinctions in the philosophy of mind and ethics' (2001: 234). Later they acknowledge that this is only a 'modest asset' (2001: 235). But I think even that is overstating the case. If the goal is to capture accurately the way in which the term 'internalism' has been used in epistemology, the alleged virtue doesn't strike me as even a modest asset. For we have no reason to assume that the term 'internalism' has been used in epistemology in ways that closely parallel its use in philosophy of mind and ethics. The only other defense Conee and Feldman offer is the claim that MI, like the accessibilist account of internalism, seems to fit with what *some* epistemologists have said about internalism. But as they admit, that doesn't tip the scales in favor of MI over the accessibilist account of internalism, since both accounts have this same virtue.

As far as I can tell, neither Feldman nor Conee seems to say anything else in defense of the claim that to be an internalist one must give the mentalist answer to question (A)—the question asking what sorts of things count as appropriate inputs to our belief-forming processes. However, in a paper authored alone, Feldman defends MI by arguing that to be an internalist one must give the mentalist answer to question (B)—the question asking whether the appropriateness of an input state must be due to one of its *essential* properties. Here's what he says in discussing Alston's reaction to mentalism:

Alston seemed to find [mentalism] to be a peculiar kind of theory . . . [I]f my feeling of warmth contributes to my being justified in believing that I am warm, then there are [thinks Alston] two justifiers present: the feeling and the fact that the feeling justifies the belief. Alston took internalists to be committed to the view that justified believers have access to all justifiers. He rejected the requirement that one have access to the linking fact [i.e. the fact that the feeling justifies the belief]. But he also took the linking fact to be a contingent fact. This is the truly externalist feature of his theory. (Feldman 2004: 155)

The justifiers that Feldman calls 'linking facts' are what I have been thinking of as facts concerning the appropriateness of the input on which the belief is based. In effect, what Feldman is here emphasizing is that it is Alston's answer to question (B) that is the 'truly externalist feature' of Alston's theory. Presumably

the point here isn't that one's answer to question (A) is irrelevant. (Surely Feldman wouldn't count as internalist a view that said the inputs were completely external to the believer though their appropriateness as belief grounds was due to an essential feature of them.) Rather Feldman's point is that Alston's agreement with the mentalist answer to question (A) isn't sufficient to make him an internalist: he must also agree with the mentalist answer to question (B). Why does Feldman think this? He thinks it because a contingent-reliability requirement like Alston's 'imposes something closely resembling the contingent natural relations that the paradigm externalist theories—the causal theory and reliabilism—imposed' (Feldman 2004: 155).

We can summarize the above discussion, then, as follows. On the one hand, Conee and Feldman think that MI is to be preferred to my account of internalism in terms of awareness requirements. On the other hand, they don't give us any good reason for thinking internalism requires the 'mental state' answer to question (A) rather than the 'accessible state' answer to it. In the end, *Feldman's* main defense of MI comes down to this: internalism requires an affirmative answer to question (B), since a negative answer results in too close a resemblance to certain paradigm externalist theories.

But even if he is right that, to be an internalist, one must require for justification that the appropriateness of a belief input be due to its *essential* properties, this doesn't help him in his defense of MI. For his claim that, to be an internalist, one must impose that 'essential property' requirement on input appropriateness doesn't count against my claim that, to be an internalist, one must impose an *awareness* requirement on justification. For it may be that *both* are required for being an internalist. So his claim doesn't count against my 'awareness requirement' account of internalism—in which case it doesn't favor a mentalist account of internalism over my own. I conclude, therefore, that Conee and Feldman haven't given us anything compelling in support of their claim that CF-mentalism is equivalent to internalism.

1.2.3 Why CF-Mentalism Isn't Equivalent to Internalism

I've pointed out that Conee and Feldman haven't given us any good reason for equating internalism with CF-mentalism. But do we have any good reason for rejecting the claim that the two are equivalent? I think we do. *My* main reason for objecting to Conee and Feldman's mentalist account of internalism is that it seems clearly to misclassify certain possible theories of justification. Consider for example the following theory about justification:

T1: The justification of our beliefs is determined by those of our mental states that are of kind K. It is highly uncommon for a person to be *aware* (or even potentially aware) of mental states of kind K. But it isn't at all uncommon for a person to *have* mental states of kind K that justify her beliefs. As a result, most of

our justified beliefs are justified in virtue of our being in mental states we aren't aware of (or even potentially aware of). Thus, most of our justified beliefs are justified despite the fact that we aren't aware or even potentially aware of *anything at all* contributing to their justification.

It seems clear that, given assumptions that have become entrenched in the field of epistemology over the last few decades, T1 is *not* an internalist theory. For according to T1, a person's beliefs can be justified even if she has no idea whether any justification-contributors are present. And yet if internalism is defined as CF-Mentalism, T1 must be classified as an internalist theory. I take that result to count decisively against Conee and Feldman's mentalist account of internalism.

What if T1 is a necessarily false view due to the fact that mental states are essentially things of which their subjects are aware, or potentially aware? Does that prevent it from being a counterexample to MI? Not at all. Presumably all false theories of justification are, due to the fact that they are noncontingent, *necessarily* false. What's at issue is how to *classify* all theories of justification, not which theories of justification are *true* or possibly true. The fact that MI would classify T1 as an internalist view shows that it is mistaken.

I've offered T1 as a clear example of a noninternalist theory that MI wrongly classifies as an internalist theory. A possible response might be to point out that since T1 makes justification depend on the subject's mental states, it isn't an externalist theory. But that is beside the point unless all nonexternalist theories of justification are internalist theories. And they aren't. It isn't difficult to identify theories of justification that fit into neither category. CF-mentalism itself strikes me as a perfectly good example of such a theory and, hence, a legitimate counterexample to the claim that all nonexternalist theories are internalist. For, as I've just argued, T1 is a version of CF-mentalism but is not an internalist theory. And I am sympathetic with Conee and Feldman's claim that—both in restricting belief-forming process inputs to *mental states* and in saying their appropriateness as inputs must be due to their *essential* features—mentalism is imposing justification requirements that externalism (as epistemologists have thought of it over the years) does not impose.

Lest the use of CF-mentalism as an example of a noninternalist, nonexternalist theory seem question-begging in this context, here's another example of a theory that seems neither internalist nor externalist:

T2: Each of the necessary conditions of justification is internal to normal humans, in the sense that a normal human can tell on reflection alone whether a belief of hers satisfies it. However, it isn't required for justification that the subject be aware or potentially aware of *anything* contributing to the justification of her beliefs. Thus, an abnormal human could satisfy all of the necessary conditions of justification (and, therefore, have justified beliefs) without being aware or potentially aware that she has satisfied any of them. Likewise, perfectly normal nonhuman extraterrestrial cognizers could satisfy all of the necessary

conditions of justification without being aware or potentially aware that they have satisfied any of them. What matters for justification is that these necessary conditions are satisfied, not that the subject has any idea that they are satisfied. It just so happens that normal human cognizers *are* aware, or at least potentially aware, that they are satisfied.

Given that T2 permits justification when the subject has no idea whether any necessary conditions of justification are satisfied, the view seems (in light of how epistemologists have been thinking of internalism) to be a noninternalist view. And given that it stipulates that each of the necessary conditions of justification is internal to normal adult humans, it appears (in light of how epistemologists have been thinking of externalism) to be a nonexternalist theory as well. We should conclude, therefore, that not all theories of justification are either internalist or externalist. Some are neither. Thus, the fact that T1 isn't an externalist theory doesn't prevent it from being a clear case of a noninternalist theory that a CF-mentalistic account of internalism wrongly classifies as an internalist theory.

I conclude, therefore, that with T1 we have a decisive counterexample to MI which shows that it misclassifies a noninternalist theory of justification as an internalist theory. Given that the 'awareness requirement' account in premise (I) from Chapter 1 doesn't misclassify T1, it is to be preferred as an account of internalism. With this defense of the conclusion that CF-mentalism is not internalism behind us, we can now turn to considering whether CF-mentalism is true.

1.3 Is CF-Mentalism True?

I will be arguing here that CF-mentalism is mistaken. The first order of business, in Section 1.3.1, will be to expose and set aside a misguided but perhaps influential reason for thinking CF-mentalism is plausible. Then, in Sections 1.3.2–1.3.3, I will consider a variety of other reasons for endorsing CF-mentalism, explaining why those reasons fail. Finally, in Section 1.3.4, I will give two reasons for thinking that CF-mentalism is false.

1.3.1 *The Subject's Perspective Objection and CF-Mentalism*

There is, I believe, a temptation to find CF-mentalism plausible due to the fact that it seems to be supported by the very intuitions that support the Subject's Perspective Objection. That objection, you will recall, went as follows:

The Subject's Perspective Objection (SPO): If the person holding a belief isn't aware of what her belief has going for it, then she isn't aware of how its status is any different from a stray hunch or an arbitrary conviction. From that we may conclude that from her perspective it is an accident that her belief is true. And that implies that it isn't a justified belief.

How can this objection to externalism be used to motivate CF-mentalism? Suppose you found the SPO persuasive and that, as a result, you thought justification required actual or potential awareness of some justification-contributors. One way to flesh out such a view is to think of justification as something requiring the subject to be aware both of her belief *grounds* (i.e. that on which her beliefs are based) and of the *adequacy* of those grounds for justification. Thinking those two things could very easily lead you to endorse the following two characteristic theses of CF-mentalism: (a) the items on which a belief must be based if it is to count as justified are *mental states* and (b) it's a thing's *essential* properties that make it an appropriate basis for a justified belief.

To see how you might be led to endorse (a), note that if you thought that the subject had to be aware of her belief grounds, you might find it natural to conclude that these grounds must be *mental states* of the subject. For you might think the subject's mental states are the only candidate belief grounds of which we have the capacity to be directly aware.

To see how you might find (b) tempting, consider the fact that if you thought the subject had to be aware of the adequacy of her belief grounds (if her beliefs were to be justified), you might conclude, if you also assumed that our beliefs often *are* justified, that the adequacy of those grounds had to be due to their *essential* properties. For the contingent facts about belief grounds that seem relevant to their adequacy for justification—facts such as whether the belief ground is a reliable indicator of the truth of the belief based on it—don't seem to be facts that we can be aware of on reflection alone. This might suggest that it is only if the adequacy of a belief ground was due to its essential properties that one could be guaranteed to be aware of it.

Thus, by endorsing a certain sort of awareness requirement on justification—one requiring awareness of both the ground of the belief and the adequacy of that ground—one could be led to endorse both (a) and (b) above, which amounts to an endorsement of CF-mentalism. This is how the intuitions supporting the SPO—and, thereby, prompting the imposition of an awareness requirement on justification—can lead one to think CF-mentalism is plausible. But the problem is that proponents of CF-mentalism should *reject* the intuitions supporting the SPO. The CF-mentalists' endorsement of (a) and (b) must, therefore, be cut loose from this motivation for it.

My reasons for thinking that proponents of CF-mentalism should reject the intuitions supporting the SPO are, by now, familiar. It isn't that I think there is any *special* problem that results from combining CF-mentalism with support for the SPO. Rather, my concern is that CF-mentalists will face the same problem described in Section 3 of Chapter 1: to avoid the SPO, you must impose awareness requirements on justification; but if you don't require *strong* awareness, you won't avoid that objection; and if you do require strong awareness, you are faced with vicious regress problems; hence, you should reject the intuitions that lead to the SPO. In Chapter 1 I argued that you shouldn't use the SPO in

support of internalism. For the very same reasons, you shouldn't use it in support of mentalism.

I suspect that much of the initial attractiveness of CF-mentalism is derived (via the sort of reasoning laid out earlier in this subsection) from the *prima facie* appeal of the intuitions supporting the SPO. But once we recognize the bankruptcy of those intuitions, we must train ourselves not to be taken in by them. Those who endorse CF-mentalism must be careful, therefore, not to let that endorsement be motivated by SPO intuitions. But without that motivation, what is left to motivate CF-mentalism? In the next two subsections, I will consider some answers to that question.

1.3.2 Conee and Feldman's Defense of CF-Mentalism

I find two main reasons in the writings of Conee and Feldman for endorsing CF-mentalism. One is that they think CF-mentalism offers the best explanation of our intuitive judgments about justification differences in a collection of pairs of examples they've selected. The other reason, offered by Feldman in an article of his own, consists of a defense of the thesis that the appropriateness of a belief ground (or other input to a belief-forming process) must be due to its *essential* properties. Below I'll respond to both objections, though in each case my response will depend on points I develop at greater length in Chapter 5. (See the end of the Preface where I explain why the material from Chapter 5, on which I depend in giving the arguments in this chapter, comes after this chapter rather than before it.)

Conee and Feldman give us six pairs of examples in each of which either 'someone has a justified belief in a proposition while someone else's belief in that proposition is not justified, or one person's belief is better justified than the other's' (2001: 236). Their conclusion is that the justificational differences we recognize in these pairs are clearly due to mental differences. They then say that 'it is reasonable to generalize from these examples to the conclusion that' justificational differences supervene on mental differences (2001: 238).

My objection to this conclusion is that although CF-mentalism is consistent with our intuitive judgments about the examples Conee and Feldman give us, there is an externalist account that is also consistent with those judgments. Moreover, this externalist account of justification is superior to CF-mentalism insofar as it does a better job of explaining the truth of our intuitive judgments about justification in connection with *other* examples, while at the same time satisfactorily handling the examples Conee and Feldman mention. The superior externalist account I have in mind is the proper function account of justification to be presented and defended in Chapter 5. In that chapter, I will argue for, among other things, the following three conclusions:

- The pairs of examples Conee and Feldman use in support of their mentalist view are handled easily by the proper function account of justification.

- There are many examples which are better handled by the proper function account of justification than by their mentalist view.
- Unlike reliabilism, the proper function account of justification handles the intuition that a belief can be justified even though formed in an unreliable way.

In short, my response to Conee and Feldman's defense of CF-mentalism on the basis of the pairs of examples they mention is that the points just highlighted, if true, completely undermine that defense.

The other reason given in support of CF-mentalism is proposed by Feldman. He defends CF-mentalism by defending his affirmative answer to question (B)—i.e. his claim that the appropriateness of an input to a belief-forming process must be due to an essential, not a contingent, feature of it. Note, however, that even if this claim were true, it wouldn't be enough to establish the truth of CF-mentalism because CF-mentalism also insists that the input must be a mental state. Nevertheless, if Feldman were right, this would show an advantage that CF-mentalism has over externalism, since externalists don't insist that the appropriateness of an input must be due to an essential feature of it.

Here is what Feldman says in defense of his position:

Alston holds that what justifies a belief must have some significant connection with the truth of that belief [such as entailing that the belief is likely to be true]. As a result, he must reject the strong supervenience thesis that, on my view, is at the heart of internalism. Demon world examples make the point vivid. A person in the normal world can have a mental twin in a demon world. As I see it, internalism implies, quite correctly, that the person and the twin are justified in believing the same propositions. Alston's insistence on his version of the truth connection requires him to reject this plausible assumption. (2004: 155)

Feldman is here pointing out that by insisting that the appropriateness of an input is due to an essential feature of that input rather than a contingent feature of it (such as being a reliable indicator of the belief's truth), we can preserve our intuition that the beliefs of my mental twin in a demon world are as justified as mine are. He takes this to support his view that CF-mentalism is to be preferred to externalism, because CF-mentalism takes the appropriateness of an input to be due to *essential* features of it.

The problem with this argument is that it shows only how CF-mentalism preserves the following two closely related points: that justification doesn't supervene on reliability and that you can be justificationaly the same as your radically deceived, demon-world mental twin. But the externalist proper function account of justification I give in Chapter 5 also preserves these same two points.¹⁹ And it does so while emphasizing and defending the claim that the appropriateness of

¹⁹ See Section 2.2.1 in Chapter 5 for details.

an input is a *contingent* feature of it, depending not on whether it is a reliable indicator but on considerations having to do with proper function. Hence, if the argument of Chapter 5 is successful, this second reason for preferring CF-mentalism to externalism also fails.

1.3.3 Other Defenses of CF-Mentalism

One can't use the intuitions supporting the SPO to motivate CF-mentalism. And the reasons proposed by Conee and Feldman aren't sufficient to show that CF-mentalism is preferable to externalism. What other reasons are there?

One reason that might be offered for preferring CF-mentalism to externalism is that the externalist but not the CF-mentalist rejects the very plausible *basing requirement* on justification. Externalists deny that justification requires the input to the belief-forming process to be a mental state. But according to this objection to externalism, a belief's ground (i.e. that on which the belief is based) must be a mental state of the subject. Nothing else can count as a belief ground. Externalists are, therefore, forced to deny the basing requirement. For, given this 'mental state' account of what it is to be a belief ground, externalists think a belief can be justified by an input despite the fact that it isn't *based* on that input (since, according to externalists, the input needn't be a mental state). These considerations give rise to the objection that, because the basing requirement on justification is extremely plausible, the externalist rejection of it shows that externalism is inferior to views like CF-mentalism which endorse it.

I could respond by quibbling about the suggestion that only mental states count as belief grounds. But I'm willing to grant that some of the things externalists allow as inputs to belief-forming processes resulting in justified beliefs do not count as belief grounds. So I'm willing to grant that externalists don't insist on a basing requirement on justification. Furthermore, I agree that there is something extremely plausible about the basing requirement on justification: merely having a good reason for a belief doesn't make that belief justified if I don't actually base that belief on that reason. How, then, can I defend externalism in light of these admissions? I can defend it by drawing attention to the fact that externalists endorse Assumption (from Section 1.2.1 of this chapter), according to which a belief is justified only if it is held in the right way, which involves its being nondeviantly caused by an appropriate input to the belief-forming process. In cases where the input is a mental state of the subject, this amounts to a basing requirement on justification. The only difference between Assumption and the basing requirement on justification is that Assumption allows for the input to the belief-forming process to be something other than a mental state of the subject. Yet even in such cases Assumption requires for justification that the input be appropriate and that it nondeviantly cause the belief.

Thus, Assumption has all the virtues of the basing requirement without its problems. It has the virtue of requiring more, for a belief's justification, than

the mere existence of something that would be an appropriate input to a belief-forming process resulting in that belief; for it also requires that the belief be held *because of* that input (i.e. that it be nondeviantly caused by it). This is the very thing that makes the basing requirement so plausible: it requires more than the mere existence of the reason; it requires, in addition, that the belief be held *because of* that reason (i.e. that it be nondeviantly caused by it). But unlike the basing requirement, Assumption doesn't require that the input be a mental state. And this is exactly the result we are after. For, as I will point out in Section 1.3.4, it seems that the input to the belief-forming processes causing our justified beliefs needn't include any mental states.²⁰

Another reason that might be offered in support of CF-mentalism is the view that our beliefs are directly causally sensitive to our mental states (whether they are accessible or not) but not to inaccessible nonmental facts inside or outside our skin. If true, this would be relevant to the topic under discussion so long as the *nondeviance* of the causal connection mentioned in Assumption—between the belief and the input to the belief-forming process—requires our beliefs to be directly causally sensitive to these inputs. For then CF-mentalism would be demanding something which Assumption also demands, namely, that the input is something that could *nondeviantly* cause the belief.

But even if we grant that our beliefs are directly causally sensitive to our mental states but not to inaccessible nonmental facts, this doesn't help the mentalist's case. For it isn't a *necessary* truth that the beliefs of all cognizers are directly causally sensitive to their inaccessible mental states but not to inaccessible nonmental facts inside or outside the skin. At best that is a contingent truth. What is required for justification, according to the proposal under consideration, isn't that an input is a mental state but rather that it is something to which the belief is directly causally sensitive.²¹

1.3.4 Reasons to Reject CF-Mentalism

We've seen in Section 1.3.1 that the intuitions supporting the SPO can't be relied on by CF-mentalists to motivate their view. But once CF-mentalists cut their

²⁰ If, as I argue here, a belief can be (doxastically) justified without satisfying a basing requirement, then it seems that, contrary to what some philosophers think, doxastic justification for S's belief that *p* doesn't require that *p* be propositionally justified for S. Propositional justification is justification a proposition *p* has for a person S just in case S is in possession of evidence E which is such that if S were to hold a belief that *p* on the basis of E, that belief would be doxastically justified. Since such evidence can be lacking in cases where a belief is doxastically justified without satisfying a basing requirement, propositional justification can also be lacking (in which case it isn't required for doxastic justification).

²¹ In Section 2.3.1 of this chapter, I discuss a view that is very much like the view identified here—i.e. the view according to which justification requires that the inputs to the belief-forming process be something to which the belief is directly causally sensitive. That view is closely connected with PC-mentalism (the version of mentalism endorsed by Pollock and Cruz) which, I argue, is false when applied to ordinary justification.

ties to those intuitions, it is difficult to see why they impose their favored constraints on justification. For as I've argued in Section 1.3.2, the reasons offered by Conee and Feldman don't adequately motivate these constraints. Neither do the two reasons discussed in Section 1.3.3. In light of these failures to adequately motivate CF-mentalism, we have good reason to refrain from endorsing it.

But not only are we lacking good reasons for endorsing CF-mentalism, we also have at least two reasons to think it is false. CF-mentalism can be viewed as having two components, consisting of its answers to the following two questions:

- (A) What sorts of things can be appropriate inputs to our belief-forming processes?
- (B) Must the appropriateness of such an input (for leading to the belief to which it leads) be due to an essential property of that input?

Thus, one can object to CF-mentalism either by rejecting its answer to (A) or by rejecting its answer to (B). Let's consider both in turn, beginning with its answer to (B).

The CF-mentalist answer to question (B) is 'yes'. But, as I already suggested near the end of Section 1.3.2, that is a mistake. I won't take the time to argue for this here, but in Chapter 5 I will defend the following conclusion:

- The appropriateness of an input (for leading to the belief to which it leads) is due to a contingent feature, not an essential feature, of that input.²²

The truth of that conclusion shows that the CF-mentalist's answer to question (B) is mistaken. And that is sufficient for showing that CF-mentalism itself is mistaken.

Consider next the CF-mentalist answer to question (A). As I've already indicated, I'm relying on the following widely accepted assumption:

Assumption: A belief is justified only if it is held in the right way. This requires, at the very least, that the belief is nondeviantly caused by an appropriate input to the belief-forming process.

Question (A) asks what sorts of things can count as appropriate inputs to the belief-forming process. The CF-mentalist says that only mental states can play that role. Is she right about that? Near the end of Section 1.2.1 of this chapter, I mentioned that many externalists will be inclined to think it is possible for God, if he exists, to reveal things to us (thereby giving us justified beliefs in the truths so revealed) by directly causing beliefs in us, without the causal intermediation of other mental states. My rejection of the mentalist answer to (A) is based on this very line of thought. It just seems, intuitively, that such divinely caused beliefs

²² As I noted earlier, I explain near the end of the Preface why I save this argument until Chapter 5 rather than giving it here.

could be justified. Hence, it seems intuitively that the mentalist answer to question (A) is mistaken.

That divine revelation example is a convenient one mainly because it is easy to make clear that there are no mental states that are mediating causes of the belief: God simply causes the belief in us directly. But we can think of other examples making the same point. Imagine alien cognizers who form the belief that there is water nearby via a belief-forming process that bypasses their other mental states. Suppose, for example, that water in the environment of these aliens causes in them the belief that there is water nearby, without using any other mental states as intermediate causes of those beliefs. And suppose, furthermore, that these beliefs are not only reliably formed but also formed in accord with what counts as proper function for these cognizers (because these cognizers were in some sense designed by God or evolution to form these beliefs in this way). Once again, these beliefs seem, intuitively, to be justified beliefs even though they are caused directly by events that are entirely external to the believer.

I realize the considerations in the previous two paragraphs won't convince diehard mentalists who will simply deny the intuitions on which the points I make rely. However, I find the intuitions persuasive and I hope other readers will too. In any case, even without this objection to the mentalist's answer to question (A), we still have the objection (noted above and discussed at greater length in Chapter 5) to their answer to question (B). And we still have the fact (defended in sections 1.3.1–1.3.3 of this chapter) that there seems to be no adequate motivation for endorsing the mentalist constraints on justification. Given that mentalist constraints are unmotivated, it's best not to endorse them.

1.4 Conclusion on Conee and Feldman

In this first part of this chapter, I've answered the following three questions as they apply to CF-mentalism:

1. What does this version of mentalism say?
2. Is it equivalent to internalism?
3. Is it true?

The answer to (1) is just that CF-mentalism (with respect to *doxastic* justification) is basically identical to M. The answer to (2) is that CF-mentalism is *not* equivalent to internalism. Conee and Feldman give us no good reason for preferring MI to the 'awareness requirement' account I gave in premise (I) from Chapter 1. Furthermore, we have the following excellent reason for preferring my account of internalism to their mentalist account of it: theirs misclassifies as 'internalist' a theory of justification (i.e. T1) that my account correctly says is not internalist.

As for question (3), I argued, first, that although it is tempting to let one's support for CF-mentalism be motivated by the intuitions that support the SPO, this temptation should be resisted. This leaves us with the task of explaining why

we should endorse the constraints on justification imposed by CF-mentalism. Neither the reasons offered by Conee and Feldman nor the others I considered were up to that task. Furthermore, we have the two objections to CF-mentalism given in section 1.3.4. On the basis of these considerations, I concluded that CF-mentalism should be rejected.

2 POLLOCK AND CRUZ

Let's turn now from CF-mentalism to PC-mentalism, the version of mentalism proposed by Pollock and Cruz.

2.1 What is PC-Mentalism?

I said in the introduction that Pollock and Cruz endorsed something like:

M. A belief's justification is a function *solely* of (i) which mental states the subject is in and (ii) which mental states of the subject the belief is based on. (i.e. if two possible subjects are exactly alike mentally and in terms of which of their mental states their beliefs are based on, then they are exactly alike justificationaly.)

Here is their own description of internalism, a view they endorse: 'Internalist theories make the justifiability of a belief a function of the internal states of the believer, in the sense that if we vary anything but his internal states the justifiability of the belief does not vary' (Pollock and Cruz 1999: 137). As for internal states, they define them as follows:

We propose to define internal states as those states that are directly accessible to the cognitive mechanisms that direct our epistemic cognition. The sense in which they are *directly* accessible is that access to them does not require us first to have beliefs about them [i.e. they can be accessed without our first having to make a judgment about whether we are in those states]. (1999: 133)

And, in considering which states are directly accessible (and therefore internal), they say this: 'We have characterized internalist theories in terms of direct accessibility, but we have not said anything in a general way about which properties and relations are directly accessible. It seems clear that directly accessible properties *must be* in some sense 'psychological', (1999: 133; emphasis added). What seems to be suggested here is that a belief's justifiability is a function solely of the psychological states of the believing subject (since it is a function of internal states, which are directly accessible states, which *must be* psychological states). If we take psychological states to be mental states, then PC-mentalism appears to be the view that justifiability is a function solely of which mental states the subject is in.

This view of justification appears to differ from M in just the same way that Conee and Feldman's view seemed at first to differ from M: it makes a belief's

justifiability a function only of which mental states the subject is in and not at all of what that belief is based on. But this difference is merely apparent for basically the same reason mentioned in discussing Conee and Feldman's account of justification. Pollock and Cruz (1999: 79) draw a distinction between a *justified* belief and a *justifiable* belief. A justifiable belief is one that has what I call 'propositional justification'. It is one for which the subject has some adequate reason or evidence and it counts as justifiable whether or not it is based on that reason or evidence. A justifiable belief becomes a justified belief—i.e. one that has what I call 'doxastic justification'—only if it is *based on* the reason or evidence that makes it justifiable.²³ Notice that what Pollock and Cruz in fact say in the first of the three passages quoted above is that *justifiability* (not justifiedness) is a function of the subject's internal states and that we can vary anything but the internal states without affecting the belief's *justifiability*.²⁴ By making it clear that a justified belief is simply an appropriately based justifiable belief, they are thereby saying that a belief's justification is a function of only these two things: which internal (or mental) states the subject is in—this determines the belief's justifiability—and what that belief is based on.

At first (or at least second) glance, therefore, it looks like PC-mentalism is equivalent to M. That is why I said that they endorse 'something like' M. There are two problems, however, that prevent us from saying that PC-mentalism is 'exactly like' M. The first has to do with Pollock and Cruz's identification of internal states with psychological or mental states. The second has to do with the fact that the justification with which Pollock and Cruz are concerned is *procedural* justification. Let's consider these two problems in turn.

2.1.1 Internal States

Pollock and Cruz identify internal states (i.e. states that are directly accessible to cognition) with psychological or mental states. But this identification is doubtful. To see why, we need to see first that their notion of direct accessibility to cognition is very different from the sort of *special epistemic accessibility to the subject* that is commonly emphasized by internalists. The accessibility to cognition that Pollock and Cruz have in mind is not *epistemic*: 'The sense in which they [i.e. the circumstances in response to which we do things, such as form beliefs] must be directly accessible is that our cognitive system must be able to access them without our first having to make a *judgment* about whether we are

²³ See Section 1 of Chapter 1 for a discussion of the difference between doxastic and propositional justification.

²⁴ Unfortunately, Pollock and Cruz aren't consistently careful about this. Sometimes they seem to switch back and forth between 'justifiedness' and 'justifiability' as if they mean the same thing. Thus, for example, in Pollock and Cruz (1999: 133) they say 'internalist theories make justifiedness a function exclusively of the believer's internal states'. See also Pollock and Cruz (1999: 24–5) where they seem to use the two terms interchangeably.

in circumstances of that type. We must have non-epistemic access' (1999: 132). Nor is it accessibility to the *subject*:

[C]ognition is an internal process that accesses internal states directly. Cognition works by noting that we have certain beliefs [or other internal states] and using that to trigger the formation of further beliefs. However, it is *cognition* that must note that we have certain beliefs—we do not have to note it ourselves. The sense in which cognition notes it is metaphorical—it is the same as the sense in which a computer program accessing a database might be described as noting that some particular item is contained in it. (1999: 136)

The idea seems to be that, in order for something X to count as directly accessible to our cognition, the mechanisms by which beliefs are formed in us must be responsive to X—they must be able to detect X in the way a computer can detect items in a database. I myself might have no idea whether X is present to be detected. But so long as the mechanisms by which my beliefs are formed are responsive to the presence of X, X counts as directly accessible to my cognition (or, loosely speaking, to me). States that are directly accessible in that sense are what Pollock and Cruz think of as *internal*. Clearly, this sense of the term 'internal' is quite different from the 'epistemic access to the subject' sense of 'internal'.

Now why think that only psychological or mental states are, in that sense, internal states? Consider again the possibility that, if God exists, he could directly cause beliefs in me. If I were designed to regularly form beliefs in that way, then certain states or activities of God would be directly accessible to my cognition and, hence, be states internal to me. Likewise, if a species of cognizers had evolved so that their beliefs were formed directly in response to remotely transmitted radio signals (without any mediation by other mental states), then those signals would be directly accessible to their cognition and, therefore, states internal to them. I don't know what Pollock and Cruz would say about these cases. I don't know whether they'd want to deny that those radio signals or states of God would count as internal states of those cognizers. And if they would want to deny this, I'm not sure what they would say about direct accessibility to explain that denial. If they want to say that, by definition, only mental states counted as directly accessible, then they should have just defined internal states as mental states to begin with and skipped the discussion of direct accessibility. Short of that, I don't see how they could guarantee that all states that are, in their sense, directly accessible are mental states.

It seems, therefore, that the official account of PC-mentalism runs as follows:

M_{PC} A belief's justification is a function *solely* of (i) which internal states (i.e. states directly accessible to the subject's cognition) the subject is in and (ii) which internal states of the subject the belief is based on.

Since it isn't at all obvious that, necessarily, only mental states are directly accessible to our cognition, it isn't at all obvious that PC-mentalism (defined as M_{PC})

entails or is equivalent to M. Given that PC-mentalism defines justification in terms of internal states and not mental states, one might wonder why I say it is a form of *mentalism*. I do so because its form is so similar to M and because Pollock and Cruz say that internal states must be mental (or at least psychological) states.

2.1.2 *Procedural Justification*

Even if we waive the concern just discussed and take for granted that, necessarily, only psychological or mental states are directly accessible to our cognition, there remains a more significant problem with the suggestion that PC-mentalism entails or is equivalent to M. That problem is that, although Pollock and Cruz agree that a belief's justification depends on its being properly based, they don't have in mind the sort of justification I'm talking about in this book—namely, ordinary justification, which I described in Section 1 of Chapter 1. They may endorse the words used in stating M, but only if M is interpreted so that the term 'justification' that appears in it is understood as *procedural* justification. Thus, in order to understand this way in which PC-mentalism differs from M (applied to ordinary justification), we need to be clear about what procedural justification is and how it differs from ordinary justification.

Pollock and Cruz make the following a matter of definition: a belief is procedurally justified only if it is formed in accord with correct epistemic norms that are usable by our automatic cognitive processing systems to guide our belief-formation.²⁵ Furthermore, they make it a matter of definition that the *correctness* of our epistemic norms cannot be contingent on features of those norms that aren't usable (detectable) by us in selecting which norms to employ.²⁶ As a matter of definition, therefore, they say that the only factors on which procedural justification depends are factors usable by our cognition. On the basis of this definition of procedural justification, they argue that procedural justification is a function of internal states (i.e. states that are directly accessible to our cognition) together with facts about what our beliefs are based on. For, as they argue,

²⁵ Initially they say (1999: 14) that, as a matter of definition, a belief is procedurally justified only if it is formed in accord with norms which *we* could use if we were deciding what to believe. But later (1999: 132–6) it becomes clear that it is our cognition, not us, that must be able to use the norms in 'deciding' what to believe.

²⁶ In Pollock and Cruz (1999: 140–1) they say that some externalists tell us: 'that if old patterns of reasoning [i.e. old norms] are unreliable and new patterns [norms] are reliable, then regardless of whether we *know* these facts about reliability, we should not reason in accordance with the old patterns [norms] and we should reason in accordance with the new patterns [norms]. What could the point of this claim be? It cannot be taken as a recommendation about how to reason, because it is not a recommendation anyone could follow. We can only alter our reasoning [i.e. the norms we employ] in response to facts about reliability if we are apprised of those facts'. They go on to make it clear that they are taking this fact—i.e. that this alleged good-making feature of our norms (reliability) isn't usable (because not detectable) by our cognition—to entail, by definition, that this externalist position fails as an account of *procedural* justification.

it is only states that are directly accessible to our cognition that are usable by our cognition (Pollock and Cruz 1999: 132–5).

It is important to notice here that there are at least two ways to arrive at a conclusion about what a kind of justification depends on. One can arrive at such a conclusion by *analyzing* the sort of justification in question in the way philosophers often do—for example, by using the particularist methodology of considering possible examples where the justification in question is clearly present and other possible examples where it is clearly absent and then formulating analyses that correctly account for our judgments about these cases. Or one can make it a matter of *stipulative* definition. So, for example, there are at least two ways that you could arrive at the conclusion that justification is reliability-entailing. You could say that what clear cases of justified belief have in common is that they are formed in reliable ways whereas clear cases of unjustified belief are formed in unreliable ways. Or you could just note that the sort of justification you have in mind is *reliabilist* justification which, by stipulative definition, is reliability-entailing. As already noted above, Pollock and Cruz's insistence that justification depends only on factors usable by cognition is defended as a matter of definition: it is *procedural* justification they are concerned with and, by definition, that sort of justification depends on such factors.²⁷

Now consider ordinary justification, which is the sort of justification I'm focusing on in this book. Whether it depends only on factors usable by cognition is *not* a matter of definition. My own view is that it doesn't depend solely on such factors but, again, this isn't a matter of definition. It would be perfectly sensible for me to entertain arguments for the conclusion that ordinary justification *does* depend only on factors usable by cognition. But that wouldn't be a sensible thing for me to do if it were a matter of definition that ordinary justification didn't depend on such factors.

We may conclude, therefore, that whether ordinary justification is equivalent to procedural justification is a *substantive* (rather than a trivial or definitional) issue. For as we've just seen, what's true by definition of procedural justification isn't true by definition of ordinary justification. And Pollock and Cruz give us no good reason for thinking procedural justification is equivalent to ordinary justification. Moreover, as I will explain in Section 2.3.1 of this chapter, we have good reason to think procedural justification is *not* equivalent to ordinary justification.

This gives us another reason to reject the identification of PC-mentalism—understood as M_{PC} applied to procedural justification—with M. M claims that *ordinary* justification is to be understood in terms of *mental* facts and facts about basing whereas PC-mentalism defines *procedural* justification in terms of *internal* facts (i.e. facts directly accessible to cognition) and facts about basing. M is a *substantive* and controversial thesis about the nature of justification. PC-mentalism,

²⁷ This sort of point comes out a number of times in their discussion in Pollock and Cruz (1999: 132–42).

on the other hand, is a stipulative claim about how the term ‘procedural justification’ is to be used.

2.2 Is PC-Mentalism Equivalent to Internalism?

PC-mentalism, as described in the previous paragraph, is what Pollock and Cruz call ‘internalism’. So they think that the answer to the question comprising the title of this subsection is ‘yes’. In defense of their characterization of internalism as PC-mentalism, Pollock and Cruz say only this: ‘This definition makes the internalist/externalist distinction precise in a way that agrees at least approximately with the way it has generally been used, although it is impossible to make it agree with everything everyone has said about it because philosophers have drawn the distinction in different ways’ (1999: 133). Clearly they aren’t interested in defending their definition vigorously. Nevertheless, I think it is worth mentioning that, contrary to what they suggest here, their definition of internalism as PC-mentalism does not seem to agree very well with the way the term ‘internalism’ has been used in the literature. For given their account of internalism, the following theory about procedural justification would count as an internalist view:

T3: The (procedural) justification of our beliefs is determined by internal states of kind K. It is highly uncommon for a person to be *aware* (or even potentially aware) of internal states of kind K. But it isn’t at all uncommon for there to be internal states of kind K that justify her beliefs. As a result, most of our justified beliefs are justified in virtue of internal states we aren’t aware of (or even potentially aware of). Thus, most of our justified beliefs are justified despite the fact that we aren’t aware or even potentially aware of *anything at all* contributing to their justification.

To classify such a view as an internalist view does *not* seem to fit at all with the way the term ‘internalism’ has been used in the literature. This casts serious doubt on their claim that their definition of internalism agrees approximately with the way the term has been used. At the very least we can say this: the ‘awareness requirement’ account of internalism from premise (I) of Chapter 1 is a considerable improvement over the PC-mentalist account of internalism, even when the sort of justification in question is procedural justification.

No doubt Pollock and Cruz would think it a waste of time to argue about how the word ‘internalism’ should be used. What I expect they would consider more worthy of discussion is their claim that PC-mentalism is true. For the reasons given at the beginning of this chapter, I disagree with them insofar as I think there is some value in discussing the way in which the word ‘internalism’ should be used. Nevertheless, I agree that it is also important to consider whether PC-mentalism is true. So let’s turn to that question now.

2.3 Is PC-Mentalism True?

The short answer to the title question of this section is ‘yes’. But it is a disappointing ‘yes’ because it is neither surprising nor interesting to learn of the truth of a position that is true by definition. The title question is whether procedural justification (which by definition depends on factors usable by our cognition) is a function of our internal states and what our beliefs are based on. And the answer is ‘yes’ given the trivial truth that all factors *usable* by our cognition are, in Pollock and Cruz’s sense, directly *accessible* to our cognition (and are, therefore, internal states).

Thus, although we have good reason to endorse PC-mentalism (i.e. M_{PC}), this is only because it is true by definition. And it isn’t interesting in the least to find out that a claim of that sort is true. What *would* be interesting is to learn that a claim parallel to PC-mentalism is true with respect to ordinary justification. Or, what comes to the same thing, what would be interesting is to find out that procedural justification is equivalent to ordinary justification. But Pollock and Cruz don’t defend that claim and they can’t defend it simply by stipulating that PC-mentalism is true. In fact, as I will now argue, the ordinary justification parallel of M_{PC} seems to be mistaken. In order to see just how damaging that conclusion is for PC-mentalists, I will turn, after defending that conclusion, to a comparison of PC-mentalism with a reliabilist position that is true by definition.

2.3.1 PC-Mentalism Applied to Ordinary Justification

Is PC-mentalism applied to ordinary justification true? Here the answer isn’t settled by definition. Let’s first get clear on what PC-mentalism about ordinary justification requires that is not required by externalists. Once again, we will take for granted that both sides make the following assumption about ordinary justification:

Assumption: A belief is justified only if it is held in the right way. This requires, at the very least, that the belief is nondeviantly caused by an appropriate input to the belief-forming process.²⁸

Given this, it will be helpful, if we want to identify ways in which PC-mentalism about ordinary justification requires more than externalism requires, to compare their respective answers to questions (A) and (B) introduced earlier in this chapter in Section 1.2.1:

- (A) What sorts of things can be appropriate inputs to our belief-forming processes?

²⁸ As I noted in Section 1.2.1 of this chapter where this Assumption was first introduced, the internalist’s endorsement of Assumption is made evident by their insistence that justification is subject to a basing requirement which has a causal element to it.

- (B) Must the appropriateness of such an input (for leading to the belief to which it leads) be due to an essential property of that input?

PC-mentalism answers question (A) by saying that only things that are directly accessible (in a non-epistemic sense) to our cognition can be appropriate inputs to our belief-forming processes. In Section 2.1, we noted that this required only that our cognitive processors (not we ourselves) were able to detect the input in the way a computer detects items in a database. Notice, however, that Assumption itself (which externalists accept) imposes a *similar* requirement: it says that the inputs must nondeviantly cause the belief. In each case, the main point seems to be that our belief-producing mechanisms must be causally sensitive (in some appropriate way) to their inputs. Perhaps the only difference is that PC-mentalists specify that the nondeviance of the causal connection involves some sort of directness whereas Assumption doesn't specify this. But even that might not constitute a difference between PC-mentalists and externalists. Pollock and Cruz say that the 'directness' qualification is meant to convey only that one needn't first form a belief about the input before it is used by our cognition.²⁹ And in that sense, externalists also required the causal connection to be direct. If there is a difference between how PC-mentalists answer question (A) and how externalist proponents of Assumption answer (A), it isn't much of a difference.

The similarity between the PC-mentalists and the externalist answers to question (A) suggests that if PC-mentalism about ordinary justification requires more than externalism requires, it must be because it demands more in answer to question (B). And this is exactly what we find. For PC-mentalists insist that what makes an input appropriate must be its essential properties whereas externalists don't require this.³⁰ Now that we have identified a constraint on justification imposed by PC-mentalism and not externalism, the question arises whether this constraint is adequately motivated. Does ordinary justification require that the appropriateness of the inputs to a belief-forming process be due to their essential properties or not? Here I will again direct the reader to Chapter 5 where I defend the conclusion that the appropriateness of an input is due to a contingent feature, not an essential feature, of that input.³¹ If my argument there is successful, then we may conclude that ordinary justification does *not* require what

²⁹ In Pollock and Cruz (1999: 133) they say: 'The sense in which they are *directly* accessible is that access to them does not require us first to have beliefs about them.'

³⁰ Pollock and Cruz say (1999: 137): '[T]he only properties of and relations between internal states to which internalist norms can appeal are those that cannot be varied without varying the internal states themselves. In other words, it must be necessarily true that if we are in those states then they have those properties and stand in those relations to one another. In short, they are the 'logical' properties and 'logical' relations between internal states'. In (1999: 25) they call this view 'cognitive essentialism'.

³¹ See Section 1.2 of Chapter 5.

PC-mentalism requires of procedural justification, namely, that the appropriateness of the inputs be due to their essential properties.³²

Thus, although we know *by definition* that PC-mentalism about procedural justification is true, we discover *by argument* (here and in Chapter 5) that PC-mentalism about ordinary justification is mistaken. The argument summarized in the previous paragraph gives us a good reason, therefore, for thinking that procedural justification is *not* equivalent to ordinary justification. It also suggests that establishing that the truth of PC-mentalism—if one does so only by arguing that it is true by definition of procedural justification—results in a hollow victory for PC-mentalists. To see why, let's consider this result in light of an example concerning a version of reliabilism that is true by definition.

2.3.2 Reliabilism that is True by Definition

The example I have in mind is an imaginary dispute between some reliabilists about justification and their opponents. The reliabilists in my example begin by insisting that justification requires reliability. Their opponents are willing to consider that suggestion and evaluate it. Instead of ruling it out as, by definition, impossible, they appeal to cases that seem, intuitively, to be cases of justified belief even though the belief in question isn't reliably formed. They appeal to the very plausible intuition that my mental twin who is the victim of an evil demon could still have justified beliefs just as I do, even though his beliefs aren't reliably formed. In response, the reliabilists in my example say that their concern is with *reliabilist* justification which, by definition, is reliability-entailing (i.e. all beliefs that have it are, by definition, reliably formed). They conclude by making the following point:

The sense of epistemic justification with which we are concerned is reliability-entailing, and if it is acknowledged that opposition to reliabilism bears only upon another sense of justification then our main point has been conceded.

³² For those who are beginning to worry about keeping track of all the conclusions from Chapter 5 that I depend on here in Chapter 3, I'll list all four of them together in this note: (i) The pairs of examples Conee and Feldman use in support of their mentalist view are handled easily by the proper function account of justification. (ii) There are many examples which are better handled by the proper function account of justification than by their mentalist view. (iii) Unlike reliabilism, the proper function account of justification handles the intuition that a belief can be justified even though formed in an unreliable way. (iv) The appropriateness of an input (for leading to the belief to which it leads) is due to a contingent feature, not an essential feature, of that input.

The first three points are mentioned early in Section 1.3.2; the last point is mentioned in the paragraph to which this note is attached and in Section 1.3.4. There is one other place in this chapter where I appeal to conclusions established in Chapter 5, namely, the end of Section 1.3.2. The points mentioned there are basically the last two of the four points mentioned above in this note. In Chapter 5, I will make a special point of noting when each of the four claims above is defended.

Now what is likely to be our reaction as observers of this sort of interchange? My reaction is that the reliabilist's 'victory' is without much significance. When focusing on a sort of justification where the dispute couldn't be settled by definition—where instead it had to be settled by some other method such as the particularist approach to settling such disputes—the opponents of reliabilism seemed to have the advantage (because of their appeal to the example of the unreliable but seemingly justified beliefs of demon victims). The reliabilist gained her point in reply only by declaring that she was concerned solely with a sort of justification about which her views were true by definition. We can't object to her claim that she is right concerning that sort of justification. But we can register our serious doubts about how worthwhile it is to be right about that while being mistaken when the focus is a sort of justification about which neither side is right by definition.

The lesson to be gleaned and which we can apply to our discussion of PC-mentalism is obvious. PC-mentalists are like the reliabilists above. They win their dispute only when they stipulate that the sort of justification they're interested in is one about which their view is, by definition, correct. Furthermore, Pollock and Cruz conclude their case in support of PC-mentalism in just the way the reliabilists above concluded their case in support of reliabilism. Here's what Pollock and Cruz say: 'The sense of epistemic justification with which we are concerned is the reason-guiding or procedural sense, and if it is acknowledged that norm externalism bears only upon another sense of justification then our main point has been conceded' (1999: 141). This response is just as unsatisfactory as the reliabilist response in the example above. For when it comes to focusing on a sort of justification about which the dispute can't be settled by definition, the PC-mentalist position turns out to be false. I conclude, therefore, that even if PC-mentalists are, by definition, right about procedural justification, that 'victory' is without much significance because that version of mentalism isn't an interesting one. And when we turn to a sort of justification that can't be settled by definition, PC-mentalism seems to be mistaken.

Pollock and Cruz might wish to respond by repeating a point they make a number of times in their book, namely, that procedural justification is the focus of traditional epistemology (1999: 14, 135, and 143). They might go on to claim that this makes their victory by definition more significant than the reliabilist victory in my example. But here it is important to distinguish two claims. One is the claim that the focus of traditional epistemology is the concept of procedural justification which, by definition, depends only on factors usable by our cognition. The other is the claim that justification has traditionally been understood by epistemologists as depending only on factors usable by our cognition. Both claims focus on the thesis that justification depends only on factors usable by our cognition. But the first says that traditional epistemology has treated this as a matter of definition while the second says only that epistemologists have

traditionally taken it to be true. The second claim is controversial.³³ But the first claim, which is the one Pollock and Cruz need for their response under discussion here, is worse: it is highly implausible. If it had been a matter of *definition* that justification depends only on factors usable by cognition, then externalists would have had that pointed out to them whenever they expressed their views. It would have been silly for internalists to engage externalists in a debate about a matter of definition. But if, on the other hand, it was merely a deeply entrenched and unquestioned assumption that justification depends only on factors usable by cognition, then the fact that internalists have *argued* with externalists about this (instead of simply reminding them of their definition of ‘justification’) is much easier to understand.

2.4 Conclusion on Pollock and Cruz

In this second part of this chapter, I’ve addressed the following three questions as they apply to PC-mentalism:

1. What does this version of mentalism say?
2. Is it equivalent to internalism?
3. Is it true?

In response to (1), we have seen that although PC-mentalism appears, initially, to be quite similar to M, in fact it is better stated as the following view about *procedural* justification:

M_{PC} A belief’s justification is a function *solely* of (i) which internal states (i.e. states directly accessible to the subject’s cognition) the subject is in and (ii) which internal states of the subject the belief is based on.

It differs from M in its focus on procedural justification rather than ordinary justification (the sort of justification I’m focusing on in this book), in its claim that the states on which that justification supervened were *internal* states (rather than mental states), and in its status as a stipulative definition of procedural justification rather than a substantive thesis about ordinary justification. In response to (2), I argued that PC-mentalism is *not* equivalent to internalism since it wrongly classifies as ‘internalist’ a theory of justification (i.e. T3) that is clearly not an internalist theory.

Finally, in response to question (3), we’ve seen that although PC-mentalism is true by definition when applied to procedural justification, it isn’t an *interesting* version of mentalism; indeed, it’s no more interesting than a reliabilist theory that is true by definition. And if we turn to an interesting version of PC-mentalism—one that pertains to ordinary justification—it turns out to be mistaken.

³³ See Frederick Schmitt (1992: 7–10 and chs. 1–3) for an objection to the charge that externalism constitutes a recent departure from traditional epistemological concerns and claims.

This page intentionally left blank

FOR EDUCATIONAL PURPOSES ONLY

4

Deontology

Deontology is the view that epistemic justification is to be understood in terms of concepts such as duty, obligation, and blame. Thus, one deontological account of epistemic justification says that a belief is justified if and only if one has not violated one's epistemic duties (or done what one epistemically ought not to have done) in forming the belief. Another deontological account says that a belief is epistemically justified if and only if one is epistemically blameless in holding the belief.

One prominent way of defending internalism is to point out that deontology of some sort is true and that it entails internalism. The purpose of this chapter is to evaluate that proposed defense of internalism. More specifically, it is to determine whether a deontological defense of internalism can be of help in responding to the argument against internalism presented in Chapter 1. The conclusion there was that once we give up on the main motivation for internalism—i.e. that it enables us to avoid the Subject's Perspective Objection (SPO) which initially seemed so effective against externalism—we are left without a good reason for endorsing internalism and should, therefore, reject it. The proposal to be considered here is that deontology provides the internalist with what she needs given that she can't rely on the SPO, namely, a good reason for imposing an awareness requirement on justification. In effect, the proposal to be evaluated in this chapter is an objection to premise (V) from my Chapter 1 argument against internalism. That premise said:

- V. If internalism either leads to radical skepticism or loses its main motivation for imposing the awareness requirement (i.e. avoiding the SPO), then we should not endorse internalism.

In response to (V), the deontologist view I'll be criticizing here says: 'even if we lose what you call 'the *main* motivation' for internalism, internalism remains adequately motivated by deontological considerations, so we have no reason to give it up'.

There are two sorts of question that can arise in connection with the deontological defense of internalism. There are those having to do with whether deontology is true. And there are those concerned with whether deontology entails internalism. Much of the literature on deontology has to do with questions in the first category. Consider for example the widely discussed problem of doxastic

voluntarism, which is concerned with whether we have voluntary control over our beliefs. That discussion has to do with whether deontologism is true. For it seems that if we lack voluntary control over our beliefs then we cannot have any duties or obligations with respect to them, nor could we be blamed for holding them, in which case deontologism is false.¹ In this chapter I will be ignoring that problem. Indeed, I will be ignoring altogether questions of the first sort. My focus will not be on whether deontologism is true but on whether, if true, deontologism gives us a good reason for internalism.

I'll begin in Section 1 by considering whether justification understood in terms of epistemic *duty* entails internalism. I will look at both an objective and a subjective account of epistemic duties and argue that in neither case do we get a successful argument from deontologism to internalism. In Section 2 I will examine a prominent argument proposed by Carl Ginet and criticized by William Alston—an argument that appeals to the 'ought implies can' principle. My contention there will be that it fails (as does a similar argument that Alston accepts) for reasons neither Alston nor Ginet recognize. In the third section I will focus on what I think is the most promising argument from deontologism to internalism—what I call 'the Blamelessness Argument'. I'll consider a number of ways of developing that argument and conclude that none of them succeeds. There I will also offer an explanation for why it is so tempting, in light of that argument, to make the mistake of thinking that deontologism *does* entail internalism. Finally, in Section 4, I will consider the arguments of Matthias Steup who has been more explicit than most contemporary internalists in proposing deontological arguments for internalism. I conclude that his arguments fail, largely because they depend on inferences of the sort I criticize in Sections 1 and 3.

1 SUBJECTIVE AND OBJECTIVE EPISTEMIC DUTY

According to some versions of deontologism, a belief is justified if and only if one has not violated one's epistemic duty in forming the belief. But philosophers often distinguish two sorts of duty: objective duty and subjective duty.² Your objective duty is what you in fact ought to do. Your subjective duty is what you *think* you ought to do. Or perhaps it is what you nonculpably and sanely think you ought to do (so that if your belief that you have a certain duty D is relevantly due to insanity or to some previous culpability on your part, then D won't count as your subjective duty). For our purposes, it won't matter which of these two understandings of subjective duty we employ. My main arguments in connection with subjective duty—which I consider in Section 1.2 below—will apply to either one.

¹ See Alston (1988), Audi (2001), Feldman (2000, 2001), Ginet (2001), and Steup (2000).

² See Alston (1985: 86–8), Donagan (1977: 52–7, 112 f.), Goldman (1986: 73), Kvanvig (1984: 72–4), Plantinga (1993b: 16–17), Pollock (1979: 109–10), and Unger (1986: 163–4).

It isn't uncommon to think that you can be nonculpably unaware of what (at least some of) your objective duties are and, therefore, blameless if you fail to fulfill them.³ But it is also common to think that there are duties of another kind which are such that you are always blameworthy if you violate them. And it seems that the distinction between subjective and objective duties was designed especially for discussing both (a) the view that some duties are such that you are always blameworthy if you violate them while others are such that you can blamelessly violate them and (b) the view that the reason you aren't always blameworthy for violating the latter duties is that you can be nonculpably unaware of the fact that they are your duties.⁴ The idea is that if the duty is a subjective duty, then you think you've got it, in which case you can't reasonably escape blame by pleading nonculpable ignorance as an excuse. But you *can* plead nonculpable ignorance for objective duties and, hence, be blameless in violating them.

In this section I want to consider two versions of deontologism: one that relies on a 'subjective duty' understanding of deontologism and another that relies on an 'objective duty' understanding of it. We can think of these two understandings of deontologism as follows:

D_{subj} : A belief is justified if and only if one has not violated one's *subjective* epistemic duties in holding it.

D_{obj} : A belief is justified if and only if one has not violated one's *objective* epistemic duties in holding it.

It's easy to see why someone might think that D_{subj} implies internalism. That conclusion seems to be based on what I'll call 'The Argument from Subjective Duty':

The Argument from Subjective Duty: Suppose that D_{subj} is true. Then justification is equivalent to epistemic blamelessness, in which case you are blameworthy for your unjustified beliefs. But you aren't blameworthy for something unless you can tell that you ought not to have done it. So you can tell whether or not your beliefs are justified. And this implies that there is an awareness requirement on justification.

Clearly, this argument is relevant to whether D_{subj} entails internalism. But just as clearly, it isn't by itself enough to show that D_{obj} entails internalism. For the connection that seems to hold between blameworthiness and the violation of subjective duty does not seem to hold between blameworthiness and the violation of objective duty. This gives the strong impression that if the argument from deontologism to internalism involves thinking of justification in terms of

³ See Donagan (1977: 52–7, 112 f.) and Plantinga (1993b: 16–17).

⁴ Richard Feldman (1988b: 407) says: 'A large part of the motivation for introducing subjective justification (or obligation) is that it often seems inappropriate to blame a person for failing to do what is in fact best'.

epistemic *duties*, then we should understand deontology in accord with D_{subj} , not D_{obj} . But is this impression correct? In the remainder of this section I'll consider first whether we can build on the Argument from Subjective Duty to show that deontology understood as D_{obj} entails internalism. Then I'll consider whether the Argument from Subjective Duty successfully shows that D_{subj} entails internalism.

1.1 Deontology and Objective Duty

Despite the fact that it initially seems hopeless to argue from D_{obj} to internalism, Alvin Plantinga has identified a potential connection between deontology, understood as D_{obj} , and internalism. Here's the line of reasoning Plantinga lays out:

1. A person can always see (i.e. cannot make a nonculpable mistake about) what subjective epistemic duty requires.
2. In a large and important set of cases, objective epistemic duty for properly functioning humans coincides with their subjective epistemic duty.
3. Therefore, in a large and important set of cases, a properly functioning human can simply see (i.e. cannot make a nonculpable mistake about) what objective epistemic duty requires.⁵

The starting assumption here is that, for reasons of the sort mentioned in the Argument from Subjective Duty above, we can always tell what subjective duty requires. That gives us premise 1. And since subjective duty and objective duty coincide for properly functioning humans in a large and important set of cases, we may conclude that in such cases, a properly functioning human can always tell what objective duty requires. Hence, just as our special epistemic access to what *subjective* duty requires suggests that there is an awareness requirement on justification understood in accord with D_{subj} , so also our (just established) special epistemic access to what *objective* duty requires suggests that there is an awareness requirement on justification understood in accord with D_{obj} .

When viewed as a deontological defense of *internalism*, the argument of the previous paragraph is, as Plantinga would say, about as imperforate as an afghan knit by an elephant.⁶ For one thing, the conclusion isn't that justification understood in accord with D_{obj} is subject to awareness requirements. It is acknowledged that the access to what objective duty requires is at best guaranteed only for a large and important class of cases and only if the person holding the belief is functioning properly. So the access in question clearly

⁵ See Plantinga (1993b: 15–25). Premise 2 in this argument corresponds to what Plantinga calls 'M2' (the second internalist motif) and the conclusion, 3, corresponds to what he calls 'C1' (the first corollary of that motif).

⁶ See Plantinga (1974: 17).

isn't *required* for justification itself (not even if proper function itself is also required for justification). Moreover, as Anthony Brueckner points out (1996: 532), according to many views concerning what objective duty requires, premise 2 in the above argument is false and Plantinga does nothing to show those views are mistaken. How shall we account for Plantinga's uncharacteristically allying himself—or at least seeming to ally himself—with such a weak argument?

The answer to that question is rather simple. He *isn't* allying himself with such an argument. Plantinga is trying to explain why Descartes and Locke were inclined to think not only that D_{subj} leads to internalism about subjective justification (in the way the Argument from Subjective Duty suggests) but also that D_{obj} leads to internalism about a more objective sort of justification too. He takes Descartes and Locke to endorse premise 2 above. And he takes them to be concerned only with properly functioning humans in the large and important class of cases mentioned. As far as I can tell, he doesn't himself endorse the argument or claim that if its premises were correct, internalism would be true. His purpose is only to explain why Descartes and Locke were inclined to connect deontologism with internalism.⁷

It seems, therefore, that we should stick with our original impression that, among those views which say justification is a matter of epistemic dutifulness, it is only deontologism understood as D_{subj} that has a chance of giving us a reason for internalism. We can at least see, via the Argument from Subjective Duty, why people *think* that deontologism so understood entails internalism. But nothing like the Argument from Subjective Duty seems available in support of the conclusion that D_{obj} entails internalism. Thus, if someone wants to defend internalism using a duty-based sort of deontologism, deontologism will have to be thought of as D_{subj} .

1.2 Deontologism and Subjective Duty

Consider again the Argument from Subjective Duty:

The Argument from Subjective Duty: Suppose that D_{subj} is true (i.e. that a belief is justified if and only if one has not violated one's *subjective* epistemic duties in holding it). Then justification is equivalent to epistemic blamelessness, in which case you are blameworthy for your unjustified beliefs. But you aren't blame-worthy for something unless you can tell that you ought not to have done it. So you can tell whether or not your beliefs are justified. And this implies that there is an awareness requirement on justification.

⁷ One might wonder whether it is very charitable of Plantinga to interpret Descartes and Locke in the way he does. And one might also raise the sorts of questions Brueckner raises (1996: 531–2) concerning how this bad argument is supposed to be relevant to contemporary defenses of internalism. I won't take the time here to pursue these questions.

In particular, consider the second sentence. That sentence, insofar as it says that D_{subj} entails that justification is equivalent to epistemic blamelessness, seems to take for granted that one is always blameworthy for failing to do one's subjective epistemic duty. Given the account of subjective duty we're working with, this amounts to saying:

A: Necessarily, one is blameworthy for failing to believe in accord with the epistemic duties one thinks one has.

But A seems implausible—or at least it seems implausible if we grant that, due to our nonculpable ignorance, we could be blameless in violating some of our duties. For suppose I think that I ought always to believe a proposition if and only if it is supported by all the available evidence. Given how difficult it is to consider a proposition in light of all the available evidence and how difficult it is to correctly determine whether or not it is supported by that evidence, it seems that I could easily be nonculpably unaware of whether I've formed my beliefs in accord with that duty I think I have. But then it seems I can be blameless in violating epistemic duties I think I have. In fact, so long as it is possible for there to be some property F such that (i) I (nonculpably and sanely) think I have the duty to believe a proposition if and only if it has F and, at the same time, (ii) there is some proposition I'm considering such that I am nonculpably unaware of whether it has F, then A is false. But that does seem possible, as the example I just gave shows. Hence, the Argument from Subjective Duty seems mistaken because of the inference in its second sentence.⁸

The argument from deontology understood as D_{obj} tried to link objective dutifulness with subjective dutifulness, subjective dutifulness with blamelessness, and then blamelessness with awareness requirements. In Section 1.1 I pointed out the weakness of the alleged link between objective dutifulness and subjective dutifulness. In this section I've pointed out the weakness of the alleged link between subjective dutifulness and blamelessness (i.e. violation of subjective duty doesn't entail blameworthiness). But even if we grant that these two links aren't sufficiently strong to be of use in an argument from deontology to internalism, deontologists needn't despair. For at most this shows that *duty-based* versions of deontology don't entail internalism. Nothing follows about whether versions of deontology that focus instead on blamelessness entail internalism. Thus, I will set aside the Argument from Subjective Duty and turn, in Section 3, to a way

⁸ One might be inclined to respond as follows to this objection to the Argument from Subjective Duty: 'That objection works only if you focus on believed duties to obey general maxims or policies. But suppose you focus only on believed epistemic duties to refrain from a particular belief. For example, suppose you revised the Argument from Subjective Duty so that it took for granted not A but A*: *Necessarily, one is blameworthy for believing p if one thinks one has a duty not to believe p*. Then that argument wouldn't be subject to the sort of objection proposed above in the text.' In response, I'll simply note that I deal with something like A* below in Section 3 (see EBW and my discussion of it).

of developing the latter part of it which I'll call 'the Blamelessness Argument'. But before leaving the related notions of duty and obligation behind, I want to consider, in the next section, an argument by Carl Ginet that differs from the Argument from Subjective Duty by relying on the *ought implies can* principle.

2 GINET'S ARGUMENT

2.1 Ginet and Alston

Ginet's well-known argument for internalism begins with a deontological conception of justification and runs as follows:

Assuming that *S* has the concept of justification for being confident that *p*, *S* *ought* always to possess or lack confidence that *p* according to whether or not he has such justification. At least he ought always to withhold confidence unless he has justification. This is simply what is meant by having or lacking *justification*. But if this is what *S* ought to do in any possible circumstance, then it is what *S* *can* do in any possible circumstance. That is, assuming that he has the relevant concepts, *S* can always tell whether or not he has justification for being confident that *p*. But this would not be so unless the difference between having such justification and not having it were always directly recognizable to *S*. (Ginet 1975: 36)

Here is Alston's formulation of this argument:

Ginet's Argument

1. *S* ought to withhold belief that *p* if he lacks justification for *p*.
2. What *S* ought to do *S* can do.
3. Therefore, *S* can withhold belief wherever *S* lacks justification.
4. *S* has this capacity only if *S* can tell, with respect to any proposed belief, whether or not *S* has justification for it.
5. *S* can always tell this only if justification is always directly recognizable.
6. Therefore justification is always directly recognizable. (Alston 1986: 217)

Alston's restatement seems to me to be fair and accurate.

The first thing to consider is what is meant by 'can do' in premise 2. Suppose there is a locked safe and that you know the combination and that I don't know the combination (nor do I have any other way of breaking into the safe or finding out what the combination is). Now consider these two accounts of 'S can do A':

C1: *S* *can*₁ *do A* iff (a) *S* has the capacity to perform the activities which would bring it about that *S* has done *A* and (b) *S* is able to do *A* at will.

C2: *S* *can*₂ *do A* iff (a) *S* has the capacity to perform the activities which would bring it about that *S* has done *A* even if (\sim b) *S* is unable to do *A* at will.

You *can*₁ open the safe whereas I cannot₁ open the safe. Nevertheless, I *can*₂ open the safe given my ability to use my hands to turn the dial left and right. The latter

sense of 'can do' would be the relevant one if my employers were considering whether to train me on a new job and one asked the other 'Do you think he can do that job?'. They both know I cannot₁ do that job now but they wonder whether I can₂ do it. If I cannot₂ do it, then there is no point in trying to train me. However, it is pretty clear that premise 2 is employing 'can₁' and not 'can₂'. Of course, proponents of 2 will think 2 is true on either of these two uses of 'can' since can₁ typically implies can₂. But if 'can₂' is what is being used in 2, then 4, which refers back to that capacity, will be completely implausible.

The can₁ interpretation of 'can do' in 2 suggests a particular interpretation of 'ought' in 2. As I noted earlier in this chapter, it is not unusual for philosophers to think that it is possible for us to have duties and obligations that we aren't justly blamed for failing to fulfill. This happens in cases where we are nonculpably ignorant of those duties or obligations. It also happens in cases where we are nonculpably ignorant of which actions of ours will count as fulfilling certain objective duties or obligations that we know we have. In these latter cases, although we objectively ought to fulfill the duties or obligations in question, we cannot₁ do so even if we can₂ do so. The reason we cannot₁ fulfill them is that, due to the fact that we are nonculpably ignorant of which actions of ours will count as fulfilling them, we aren't able to fulfill them at will. Clearly, the 'ought' in 2 doesn't refer to this sort of obligation since obligation of that sort doesn't imply can₁. Instead, the 'ought' in 2 refers to a kind of obligation which is such that I can always tell which of my actions will count as fulfilling it. Otherwise premise 4 wouldn't be the least bit plausible. In light of this, let's define a sense of 'ought' as follows:

O*: *S ought* to do A* iff that obligation to do A is such that a person can always tell which of his or her actions will count as fulfilling it.

If premise 4 is to be convincing, the 'ought' in premises 2 and 1 should be understood as ought*.

One final preliminary remark is required before we begin evaluating Ginet's argument. It seems that the justification mentioned in the argument is propositional justification. For the idea behind 1 seems to be that S should first consider whether he has justification for a proposition before making a decision about whether to believe it. It is, therefore, the sort of justification a proposition *p* can have for a person even if the person doesn't believe *p*. Putting this together with the conclusions of the previous paragraph, we can think of premise 1 as saying:

1* S ought* to refrain from believing *p* if he lacks propositional justification for *p*.⁹

⁹ I've replaced 'withhold belief' with 'refrain from believing'. I do this in order to better capture Ginet's and Alston's clear intention to allow that both believing that *p* is false and withholding judgment about the truth-value of *p* count as ways of withholding belief in *p*.

The rest of the argument, adjusted to fit with 1*, goes like this:

- 2* What S ought* to do S can₁ do.
- 3* Therefore, S can₁ refrain from belief wherever S lacks propositional justification.
- 4* S can₁ do this only if S can tell, with respect to any proposed belief, whether or not S has propositional justification for it.¹⁰
- 5* S can always tell this only if propositional justification is always directly recognizable.
- 6* Therefore propositional justification is always directly recognizable.

Ginet and other internalists typically think that S's belief that *p* is doxastically justified only if *p* is propositionally justified for S.¹¹ Thus, they could use 6* to argue that access to at least one factor contributing to doxastic justification—namely, the subject's possession of propositional justification for its content—is required for doxastic justification. This is why the argument above is relevant to internalism about doxastic justification.¹²

Let's turn now to Alston's analysis of Ginet's argument. According to Alston, the argument goes wrong at 5*.¹³ For, says Alston, although we must be able to tell whether or not we have done what we ought* to do, we needn't be able to tell this on reflection alone. It may take investigation. He mentions as an example that it may take some investigation to determine whether he would be morally justified in resigning his professorship as late as April 12 in order to accept a position elsewhere for the following fall (Alston 1986: 218). Since such investigation involves more than reflection alone, the facts about what one ought* to do might not be *directly* recognizable. Alston goes on to argue that such facts needn't be recognizable in any of the special senses internalists typically have in mind. He concludes that what is required is only that these facts about what we ought* to do be knowable in some way or other (Alston 1986: 220). It seems, therefore, that Alston accepts an argument like Ginet's if we replaced 5* and 6* with:

¹⁰ The 'can' in 'can tell' isn't sensibly read as 'can₂' so I won't bother specifying that it is to be read as 'can₁'.

¹¹ That proposal makes sense if (a) there is a basing requirement on doxastic justification requiring the subject's belief to be based on an adequate ground and (b) the subject's having that adequate ground is sufficient (even without basing) for the belief's being propositionally justified for the subject. However, if one denies that there is a basing requirement on justification, then it is no longer clear that doxastic justification requires propositional justification (see n. 20 in Chapter 3 for more on this). For an explanation of why I deny the basing requirement (despite my interest in accommodating the sensible motivations for it), see Chapter 3, Section 1.3.3.

¹² See Chapter 1, Section 1, for a discussion of the difference between propositional and doxastic justification.

¹³ Alston, of course, speaks of premise 5, not 5*, but we can take his remarks as applying to 5* as well.

5** S can always tell this only if propositional justification is always *knowable in some way or other*.

6** Therefore propositional justification is always *knowable in some way or other*.

But Alston would insist that this conclusion isn't enough to get you internalism.

One can imagine a line of response to Alston's objection which pointed out that, although what the subject ought* to do *later* (i.e. after investigation) isn't something directly recognizable to him now, what he ought* to do *now* (i.e. investigate the matter more fully before deciding whether to resign his professorship at such a late date) *is* directly recognizable to him now. However, I won't pursue this line of response to see whether, in the end, it will help the proponent of a Ginet-style argument. For even if Alston's objection fails, there is another objection that applies to both Ginet's argument and to Alston's 'improved' argument—the argument that replaces 5* and 6* with 5** and 6**.

2.2 Another Objection to Ginet's Argument

The other objection I have in mind is to premise 4*:

4* S can₁ do this only if S can tell, with respect to any proposed belief, whether or not S has propositional justification for it.

It will be helpful, in evaluating 4*, if we make its antecedent more explicit by saying what 'this' refers to. A quick look at premise 3* reveals that it refers to following this policy:

P: to refrain from belief wherever S lacks propositional justification.

But P can be interpreted in either of these two ways:

P1: to refrain from believing a proposition *if* S lacks propositional justification for that proposition.

P2: to refrain from believing a proposition *if and only if* S lacks propositional justification for that proposition.

This gives us two ways to restate 4* with its antecedent made more explicit:

4*_{P1} S can₁ follow policy P1 only if S can tell, with respect to any proposed belief, whether or not S has propositional justification for it.

4*_{P2} S can₁ follow policy P2 only if S can tell, with respect to any proposed belief, whether or not S has propositional justification for it.

So which is it? Premise 1* gives us the answer. What S ought* to do is follow the policy of refraining from the belief that *p* *if* (not *if and only if*) one lacks propositional justification for *p*. So 4* should be understood as 4*_{P1}, not 4*_{P2}.

But 4*_{P1} is false. To see why, it will be helpful if we state it more fully as follows:

4^*_{P1} S can₁ follow the policy of *refraining from believing a proposition if it lacks propositional justification for S* only if S can tell, with respect to any proposed belief, whether or not it is propositionally justified for S.

By stating it this way, we can see more easily that 4^*_{P1} gets whatever plausibility it has from the fact that it is an instance of the following general principle:

GP S can₁ follow the policy of *refraining from doing A to an X if that X lacks F* only if S can tell, with respect to any X, whether or not it is F.

But GP is clearly false. We can see this by considering this consequence of it:

C S can₁ follow the policy of *refraining from shooting a living thing moving in the bushes if it is a human* only if S can tell, with respect to any living thing moving in the bushes, whether or not it is a human.¹⁴

Sometimes hunters can tell what sort of living thing is moving in the bushes: they can tell that it's a bear or that it's a deer or that it's a human. But sometimes they can't tell what sort of living thing it is. And yet despite the existence of cases of the latter sort, hunters still can₁ follow the policy of *refraining from shooting a living thing moving in the bushes if it is a human*. For they can₁ follow the policy of not shooting at something moving in the bushes unless they are sure it isn't human. Following that latter policy enables them to follow the former policy. This shows that C and, therefore, GP are false.

Remarks parallel to those made in connection with the hunter example apply to 4^*_{P1} . Suppose that there are some propositions for which S can tell she *has* propositional justification, that there are others for which S can tell she *lacks* propositional justification, and that there are yet others for which S cannot tell whether or not she has propositional justification. Does the existence of those latter cases imply that S cannot₁ follow the policy of *refraining from believing a proposition if she lacks propositional justification for it*? No. For S can₁ simply make it her policy never to believe a proposition unless she is sure she has propositional justification for it.¹⁵ This shows us that 4^*_{P1} is false for the same reason that C is false. From this we may conclude that Ginet's argument fails, as does Alston's modified version of it which relies on 5^{**} and 6^{**} .

At this point, a defender of Ginet's argument might point out that the above objection to 4^*_{P1} doesn't apply to 4^*_{P2} . It's true that—contrary to 4^*_{P1} —you can₁ follow policy P1 even if there are some propositions which are such that

¹⁴ To see that C is a consequence of GP, notice that C could be restated more awkwardly as follows: S can₁ follow the policy of *refraining from shooting a living thing moving in the bushes if it lacks the property of being nonhuman* only if S can tell, with respect to any living thing moving in the bushes, whether or not it is nonhuman.

¹⁵ Obviously, I'm assuming for the sake of argument here that some version of doxastic voluntarism is true. As I noted earlier, whether it is or not is relevant to the truth of deontology. But in this chapter, I'm ignoring that question and considering only what follows *if* deontology (and the doxastic voluntarism it requires) is true.

you can't tell whether you have propositional justification for them. But 4^*_{P2} is right in saying that you can₁ follow policy P2 only if you can tell, for each proposition you consider, whether you have propositional justification for it. For to follow P2, you must *refrain* from believing every proposition for which you lack propositional justification while also *believing* every proposition for which you have propositional justification. So 4^*_{P2} doesn't suffer from the problem afflicting 4^*_{P1} . Indeed, I suspect that Ginet and Alston both interpreted 4^* as 4^*_{P2} , given that they both took 4^* to be so plausible. Why then can't Ginet's argument be saved simply by interpreting 4^* as 4^*_{P2} ?

The problem is that if 4^* is interpreted as 4^*_{P2} , 1^* has to be replaced with:

1^{**} S ought* to refrain from believing p if *and only if* he lacks propositional justification for p .

But both Alston and Ginet recognize that 1^{**} is implausible. In the passage quoted earlier from Ginet, he said: 'Assuming that S has the concept of justification for being confident that p , S ought always to possess or lack confidence that p according to whether or not he has such justification. At least he ought always to withhold confidence unless he has justification' (Ginet 1975: 36).

In that last sentence, Ginet is moving from the stronger 1^{**} to the weaker 1^* . Alston notes this and says the following:

Ginet recognizes that we are intellectually obligated to refrain from believing that p in the absence of justification, but he wisely holds back from claiming that we are obligated to believe that p wherever we have a justification. The presence of justification gives me a *right* to believe, but I am not obliged to exercise that right; I have a choice as to whether or not to do so. . . . Ginet's recognition of this point is evinced by his modifying ' S ought always to possess or lack confidence that p according to whether or not he has such justification' to 'At least he ought always to withhold confidence unless he has justification'. (Alston 1986: 217, n. 39)

As Alston sees it, then, 1^* is to be preferred to 1^{**} because, in general, when it comes to *believing*, we have permissions, not obligations, but when it comes to *refraining* from belief, we have obligations, not mere permissions.

We can think of the reasoning behind the rejection of 1^{**} as follows. There are some propositions which are such that S is nonculpably unaware of whether or not he has propositional justification for holding them.¹⁶ This shows that, contrary to 1^{**} , it is false that S ought* to follow *both* of the following policies (which together comprise P2):

P1: to refrain from believing a proposition *if* S lacks propositional justification for that proposition.

¹⁶ This claim is not likely to meet with resistance. For it isn't very plausible to insist that, for every person X , every proposition is such that X is either aware or culpably unaware of whether or not she or he has propositional justification for it.

P3: to believe a proposition *if* S has propositional justification for that proposition.

If it was true that S ought* to follow both P1 and P3, then he *couldn't* be nonculpably unaware of whether he had followed them both. For given our earlier definition of 'ought*':

O*: *S ought* to do A* iff that obligation to do A is such that a person can always tell which of his or her actions will count as fulfilling it,

we can infer, from the fact that S ought* to follow both P1 and P3, that S can't be nonculpably unaware of whether he had done so. But, as I just noted above, there are some propositions which are such that S *is* nonculpably unaware of whether he has propositional justification for them. From this it follows that he *can* be nonculpably unaware of whether he has followed *both* policies.

Now it's true that S could have an obligation* to follow one of these policies and not the other.¹⁷ For S can₁ follow P1 by refraining from believing a proposition whenever he isn't sure he has propositional justification for it. Likewise, S can₁ follow P3 by believing a proposition whenever he isn't sure he *lacks* propositional justification for it. So, assuming S has an obligation* to follow at least one of these policies, which is it? Does S have an obligation* to *believe* a proposition when he isn't sure he lacks justification for it? Or does S have an obligation* to *refrain* from believing a proposition when he isn't sure he has justification for it? The most natural answer—the one that will be most appealing to internalists given their desire not to be guilty of the externalist's 'sin' of making justification too easy to come by—is that S has an obligation* to *refrain* from believing a proposition when he isn't sure he has propositional justification for it. This, I think, explains why the first premise in Ginet's argument is 1*.

We can summarize my objection to Ginet's argument, then, as follows. For the reasons just noted, the first premise is 1*, not 1**. This forces Ginet to interpret 4* as 4*_{P1} rather than as 4*_{P2}. But 4*_{P1} is false, as I showed using the example of the hunters. Thus, Ginet's argument fails to show that deontologism entails internalism.

3 THE BLAMELESSNESS ARGUMENT

In Section 1 I presented some objections to the claim that duty-based versions of deontologism entail internalism. And in Section 2 I objected to Ginet's 'ought implies can' argument for the conclusion that a version of deontologism understood in terms of obligation entails internalism. In this section, I want to leave the related concepts of duty and obligation behind and focus on a version of deontologism according to which epistemic justification should be understood

¹⁷ S has an obligation* to do A iff S ought* to do A.

in terms of blamelessness. I want to develop the Blamelessness Argument, which, as I noted at the end of Section 1, is basically the latter part of the Argument from Subjective Duty. We can state the Blamelessness Argument, roughly, as follows:

The Blamelessness Argument: Justification is equivalent to epistemic blamelessness. Therefore, you are blameworthy for your unjustified beliefs. But you aren't blameworthy for something unless you can tell that you ought not to have done it. So you can tell whether or not your beliefs are justified. And this implies that there is an awareness requirement on justification.

As will become clear below, my dissatisfaction with this argument has to do with the third and fourth sentences in it.¹⁸ After briefly explaining how I'll be thinking of blameworthiness and blamelessness and how justification thought of as blamelessness is relevant to the concerns of this book, my task in this section will be twofold. First, I will consider several ways of more carefully formulating the Blamelessness Argument. I will conclude that, despite the fact that the Blamelessness Argument seems to be the most promising defense available for the conclusion that some sort of deontology entails internalism, it fails. Second, I will explain how focusing on *blameworthiness*, as the original version of the Blamelessness Argument does at points, contributes to the illusory appeal of the idea that justification thought of as epistemic blamelessness is subject to awareness requirements.

3.1 Blameworthiness, Blamelessness, and Justification as Blamelessness

Let's begin by considering what seems like a respectable account of generic blameworthiness:

BW: S's doing A at *t* is something for which she is blameworthy if and only if *either* (i) S believes at *t* that she ought not to do A *or* (ii) S's failure to believe at *t* that she ought not to do A is relevantly due to some other doing or failure of hers for which she is blameworthy.¹⁹

This account is motivated largely by the assumption that a person is not blameworthy for doing A if she is nonculpably unaware that she ought not to do A. I take that assumption to be relatively uncontroversial, especially in this context since it seems to be the sort of thing to which deontological defenders of internalism will want to appeal.

Two questions arise in connection with this account of blameworthiness. First, is BW correct in saying that it is sufficient for S's blameworthiness that she

¹⁸ See nn. 31 and 32 below.

¹⁹ Clause (ii) will require some work if one is to say exactly what counts as being 'relevantly' due to some other blameworthy doing or failure. Fortunately, for our purposes, we can leave that work undone here.

believes that she ought not to have done the thing she did? Take for example Huck Finn. Was he blameworthy for helping the slave, Jim, to escape when he believed that doing so was immoral and that he'd be sent to hell for it? Shouldn't we say instead that he was doing something praiseworthy, not blameworthy?

One could answer this first question by defending BW as follows:

If Huck Finn believed only that he was going against what his society (and their god) claimed was right even though what he was doing seemed to him to be the only decent thing to do, then he wasn't really going against his conscience. However, if he seriously thought that what he was doing was an immoral thing to do and yet he did it anyway, then he *is* blameworthy because he is violating his conscience. (Of course if, instead, he had turned Jim in because his conscience told him to do so, you might think he had something *else* wrong with him morally, such as having a confused sense of right and wrong.)

That's a defensive response. A concessive response would concede the point behind the question and say that although satisfying clause (i) of BW isn't sufficient for blameworthiness (in which case the 'if' part of BW's 'if and only if' claim is mistaken), this problem could be solved by changing BW from an 'if and only if' principle to an 'only if' principle.²⁰ My own response to the question is just this: for reasons to be explained below,²¹ we can see that if we weaken BW by making it an 'only if' principle rather than an 'if and only if' principle, that only makes matters worse for supporters of the Blamelessness Argument. Hence, if the concerns raised by the question at the end of the previous paragraph are genuine, they cause trouble not for me but for those who want to use the Blamelessness Argument to show that deontology entails internalism. Thus, because I want to give the Blamelessness Argument a careful hearing, I'll just ignore the concerns raised by the first question.

The second question connected with BW has to do with the fact that the term 'blameworthy' appears in the second clause of what is supposed to be an explanation of what makes an act blameworthy. Isn't that in some sense circular? No. To see why, notice that an act's blameworthiness can be original or derived. The blameworthiness of an act is derived if that blameworthiness is relevantly due to some other blameworthy act of the agent. It is original if it isn't. In other words, an act's blameworthiness is original if it is blameworthy because clause (i) above is satisfied; it is derived if it is blameworthy because clause (ii) is satisfied. Now if I am blameworthy for something, not all of the blameworthy behavior in my life is derivatively blameworthy. Instead, the derived blameworthiness of my behavior must always be traced back ultimately (perhaps via the derived blameworthiness of other acts of mine) to behavior that is *originally* blameworthy. Thus, the

²⁰ Notice that if we changed the 'if and only if' in BW to 'only if', it would be asserting nothing more than the extremely plausible claim that a person isn't blameworthy for doing A if she is nonculpably unaware that she ought not to do A.

²¹ See n. 25.

account above avoids circularity. Clause (ii) of BW functions as a recursive clause while clause (i) functions as a base clause.

To make use of this notion of blameworthiness in arguing for an awareness requirement on justification, we should first remind ourselves that the concept we're speaking of isn't *generic* blameworthiness but *epistemic* blameworthiness. Looking to BW as our example, we can define epistemic blameworthiness as follows:

EBW: S's believing *p* at *t* is something for which she is *epistemically* blameworthy if and only if *either* (i) S believes at *t* that she ought not to believe *p* *or* (ii) S's failure to believe at *t* that she ought not to believe *p* is relevantly due to some other doing or failure of hers for which she is epistemically blameworthy.²²

Next, we can spell out the corresponding notion of epistemic blamelessness which is just the absence of epistemic blameworthiness:

EBL: S's believing *p* at *t* is something for which she is epistemically *blameless* if and only if (i) S doesn't believe at *t* that she ought not to believe *p* and (ii) S's failure to believe at *t* that she ought not to believe *p* is not relevantly due to some other doing or failure of hers for which she is epistemically blameworthy.

I'll start by assuming that when the Blamelessness Argument above says, in its first sentence, that justification is equivalent to epistemic blamelessness, it is speaking of epistemic blamelessness understood in accord with EBL.

Even though we have yet to consider a careful formulation of the Blamelessness Argument, there is one question that immediately comes to mind. We can now see how the Blamelessness Argument can be used to argue from a version of deontologism to internalism about justification if our focus is justification understood as epistemic blamelessness. But I have already made clear in Section 1 of Chapter 1 that I'm interested in a *different* sort of justification—what I there called 'ordinary justification', which I said is more objective than mere epistemic blamelessness. So why bother with this particular argument from deontologism to internalism?

There are two reasons. First, I think that the argument from deontologism to internalism fails even if the internalist is using it to defend internalism about justification understood as epistemic blamelessness. And I don't want internalists to think my objection to internalism is successful only if I focus on more objective sorts of justification and ignore justification thought of as epistemic blamelessness. Second, although it is a given that ordinary justification (the sort in which I'm interested) is not *merely* epistemic blamelessness, it isn't a given that it doesn't

²² According to this definition, it is sufficient for being epistemically blameworthy in holding B that you think you *ought* not to hold it. But what if you think both that you *epistemically* ought to hold B and that you *morally* ought *not* to hold it? Are you still epistemically blameworthy for holding B? See Bergmann (2000a: 93) for some discussion related to these questions. To simplify matters, I will ignore these questions here.

require epistemic blamelessness and more besides. True, epistemic blamelessness isn't *sufficient* for ordinary justification. But it might be argued that it is *necessary* for ordinary justification. And if that were correct, then an argument from the necessity for justification of epistemic blamelessness to the conclusion that internalism about such justification is true, would be of use in defending internalism about ordinary justification. For if blamelessness is required (even though insufficient) for ordinary justification and it is the sort of thing that is, of necessity, something accessible to the subject, it follows that ordinary justification requires access to at least one of the things contributing to it. For this reason, the Blamelessness Argument, used to argue from deontology to internalism about justification thought of as epistemic blamelessness, can be viewed as relevant to whether internalism about *ordinary* justification is true.

3.2 Some Versions of the Blamelessness Argument

Our question, then, is whether internalism is entailed by justification's requiring epistemic blamelessness, where the latter is understood in accord with EBL. A natural first reaction is that the answer is 'no'. For a person can't always tell on reflection alone whether her failure to think she's violating a duty is due to some *other* failure (perhaps in the past) for which she is blameworthy. In other words, derived epistemic blamelessness isn't always accessible to a person. In light of EBL, this implies that a person can't tell on reflection alone whether she is epistemically blameless (since that requires the absence of both original and derived epistemic blameworthiness). And if justification is a matter of epistemic blamelessness, this suggests that a person *can't* tell on reflection alone whether her beliefs are justified. The problem, in a nutshell, is that because current blameworthiness can depend on past blameworthiness and past blameworthiness isn't always accessible, I am not guaranteed to be able to tell on reflection alone whether or not I am currently blameworthy.²³

But perhaps we can still use this version of deontology (according to which epistemic blamelessness is required for justification) to argue for internalism. Consider the following more careful formulation of the Blamelessness Argument:

The Blamelessness Argument: Version 2

1. S's belief, at *t* that *p*, is justified only if S is epistemically blameless in believing *p* at *t*.²⁴

²³ Cf. Greco (1990: 256–7) and Alston (1986: 208).

²⁴ I want this argument to be applicable to both ordinary justification *and* justification thought of as epistemic blamelessness. It is for this reason that I've formulated this first premise as an 'only if' rather than an 'if and only if' claim. It is only the 'only if' claim that has a chance of being true of ordinary justification (see the last paragraph of Section 3.1 of this chapter for discussion of this point). And the 'if and only if' claim isn't required for this argument to succeed since the 'only if' part does all the work.

2. S is epistemically blameless in believing p at t only if S doesn't believe at t that she ought not to believe p . [from EBL]
3. Therefore, S's belief, at t that p , is justified only if S doesn't believe at t that she ought not to believe p .

By focusing in premise 2 on only a *necessary* condition of epistemic blamelessness, we avoid mentioning the 'past blamelessness' requirement, the satisfaction of which isn't guaranteed to be accessible on reflection alone. We thereby sidestep the worry that surfaced in the previous paragraph.²⁵

If the conclusion of Version 2 of the Blamelessness Argument implied that there is an awareness requirement on justification, then we would have an argument from deontology to internalism. But it doesn't. For the conclusion isn't that justification requires awareness of a justification-contributor. Instead, it's that justification requires the *absence* of awareness of a justification-defeater.

We could alter our accounts of epistemic blameworthiness and blamelessness in order to get an argument with a conclusion that requires not merely the absence of awareness of a justification-defeater but, instead, the presence of awareness of a justification-contributor. This could be done by defining a *stronger* version of epistemic blamelessness:

EBL_S: S's believing p at t is something for which she is epistemically blameless if and only if S believes at t that she is *permitted* to believe p .

According to EBL_S, it isn't enough for epistemic blamelessness that S satisfies the negative requirement of (nonculpably) *refraining* from believing that she ought not to hold the belief in question. In addition, S must also satisfy the positive requirement of *believing* that she is permitted to hold the belief in question. It is because it imposes a positive requirement, not merely a negative one, that EBL_S is stronger than EBL. With this account of epistemic blamelessness at our disposal, we can replace Version 2 of the Blamelessness Argument with the following:

The Blamelessness Argument: Version 3

1. S's belief, at t that p , is justified only if S is epistemically blameless in believing p at t .
- 2*. S is epistemically blameless in believing p at t only if S *believes* at t that she is permitted to believe p . [from EBL_S]
- 3*. Therefore, S's belief, at t that p , is justified only if S *believes* at t that she is permitted to believe p .

²⁵ I can now explain why the Blamelessness Argument would be in trouble if we had simply conceded the point made in the 'Huck Finn' objection I mentioned above to BW near the beginning of Section 3.1. That concession would have led us to change both BW and EBW from 'if and only if' claims to 'only if' claims. And that would have forced us to change EBL from an 'if and only if' claim to an 'if' claim. The result would be that we couldn't derive premise 2 above from EBL and, hence, that the Blamelessness Argument wouldn't go through.

This time the conclusion is that there is an awareness requirement on justification since, for her belief *B* to be justified, the subject must have the further *belief* that she is permitted to hold *B*.

But this revised argument is only as plausible as the account of epistemic blamelessness on which it depends. As I noted above, those who prefer *EBL_S* over *EBL* think that *refraining* from thinking you ought not to believe *B* isn't enough for being epistemically blameless in holding *B*—in addition, you must *believe* that holding *B* is permissible. But surely such opponents of *EBL* won't think that just *any* such belief about *B*'s being permissible, no matter how insane or irrational or blameworthy it is, is enough for being blameless in holding *B*. The very intuitions that push one to require, for epistemic blamelessness in believing *B*, a further belief that *B* is permissible also demand that only a *justified* belief about *B*'s permissibility will be sufficient for *B*'s epistemic blamelessness. This suggests that opponents of *EBL* will think of justification in terms of a version of epistemic blamelessness that is even stronger than *EBL_S*, a *super strong* version which we can define as follows:

EBL_{SS}: *S*'s believing *p* at *t* is something for which she is epistemically blameless if and only if *S* *justifiedly* believes at *t* that she is permitted to believe *p*.

This changes the Blamelessness Argument as follows:

The Blamelessness Argument: Version 4

1. *S*'s belief, at *t* that *p*, is justified only if *S* is epistemically blameless in believing *p* at *t*.
- 2** *S* is epistemically blameless in believing *p* at *t* only if *S* *justifiedly* believes at *t* that she is permitted to believe *p*. [from *EBL_{SS}*]
- 3** Therefore, *S*'s belief, at *t* that *p*, is justified only if *S* *justifiedly* believes at *t* that she is permitted to believe *p*.

But although this conclusion gives us an awareness requirement on justification, the obvious problem is that it gives us a *strong* awareness requirement since it says that justification for one belief requires a further justified belief.²⁶ And, as I argued in Chapter 1, this leads to vicious regress problems since the required additional belief is at least slightly more complicated than the original belief for whose justification the additional justified belief is required. Moreover, for the reasons discussed in Section 3.2 of Chapter 1, it won't help to require only the potential for such an additional justified belief.²⁷ Thus, the strong version of epistemic blamelessness (*EBL_S*) results in a deontological account of justification

²⁶ See Chapter 1, Section 2.1, for an account of the difference between strong and weak awareness.

²⁷ One might try replacing *EBL_{SS}* with *EBL_{SS}**: *S*'s believing *p* at *t* is something for which she is epistemically blameless if and only if *S* epistemically blamelessly believes at *t* that she is permitted to believe *p*. But then epistemically blameless belief that *p* requires epistemically blameless belief in

that is too permissive to satisfy opponents of EBL; and the *super* strong version of epistemic blamelessness (EBL_{SS}) results in a deontological account of justification that is too demanding to be plausible. The lesson seems to be that we must stick with the original account of blamelessness (EBL) which requires only the absence of a belief about being obligated not to hold the belief in question.

But perhaps there is another way to save the Blamelessness Argument. Perhaps we can somehow derive an awareness requirement from the original 'absence of awareness' requirement. Consider the result of extending version 2 of the Blamelessness Argument as follows:

The Blamelessness Argument: Version 5

1. S's belief, at *t* that *p*, is justified only if S is epistemically blameless in believing *p* at *t*.
2. S is epistemically blameless in believing *p* at *t* only if S doesn't believe at *t* that she ought not to believe *p*. [from EBL]
3. Therefore, S's belief, at *t* that *p*, is justified only if S doesn't believe at *t* that she ought not to believe *p*.
4. Necessarily, if S doesn't believe at *t* that she ought not to believe *p*, then S is aware or potentially aware of the fact that she doesn't believe at *t* that she ought not to believe *p*.
5. Therefore, S's belief, at *t* that *p*, is justified *only if* S is aware or potentially aware of the fact that she doesn't believe at *t* that she ought not to believe *p*.

The idea here is that the alleged necessary truth stated in 4 enables us to derive a requirement for one sort of awareness from the requirement that another sort of awareness is absent. Unfortunately, 4 seems false. For even if we allow that the claim 4 says is necessary is contingently true for all actual persons, there's no reason to think it is necessarily true for all possible persons. Why couldn't some believer be defective in such a way that she is incapable (due to brain damage or the interference of an evil demon) of being aware of the fact that she doesn't believe that she ought not to believe *p*? Given that *not* believing you ought not to believe *p* is one thing and that being aware of that absence of belief is another, it's hard to see why you couldn't be rendered incapable of detecting such an absence.²⁸

something slightly more complex than *p*. And this will lead to a vicious regress exactly parallel to the one discussed in Section 3.2 of Chapter 1. Moreover, the prospects in this case for avoiding the viciousness of the regress are just as slim as they are in the case of the regress discussed in Chapter 1.

²⁸ One possible response to this objection to premise 4 is to argue as follows: "The proponent of the argument from deontology to internalism can avoid this objection to premise 4 by imposing a proper function condition on justification. For then we could change the argument by adding premise 3.5: *S's belief, at t that p, is justified only if S's cognitive faculties are functioning properly in forming B*. And we could change 4 to 4*: *Necessarily, if (a) S doesn't believe at t that she ought not*

But, some might say, the falsity of the claim that 4 says is necessary is a remote possibility: in most, perhaps all, *actual* cases, a person can tell on reflection alone when she doesn't believe that she ought not to hold some belief. Let's suppose that's right. How does it help? The question still arises: is B justified in the scenario (acknowledged to be possible) in which the person holding B both (a) doesn't believe she ought not to hold B and (b) is incapable of being aware of the fact that she doesn't believe she ought not to hold B? If the correct answer to that question is 'yes', then 5 above is false and this attempt to establish an awareness requirement on justification has failed. Perhaps, as a last resort, the internalist could simply *claim* that the correct answer is 'no', thereby insisting on 5 without inferring it from the conjunction of 3 and 4. The internalist could grant the point made in the previous paragraph—that you *could* fail to believe something (i.e. that you ought not to hold belief B) without being able to be aware of the fact that you don't believe it. But she could then insist that not believing you ought not to hold B isn't enough for the justification of B. In addition, you must be *aware* (or potentially aware) of the fact that you don't believe that you ought not to hold B. Unfortunately, if this is the approach taken by the internalist, then she is no longer *arguing* for the awareness requirement using the original Blamelessness Argument together with premise 4 above. Instead, she is merely *asserting* that there is such a requirement since she is claiming, without an argument from deontologism, that 5 is true. My point here is *not* that the merely asserted claim (i.e. 5) is false (though, for the record, I think it is). Rather, my point is that if 5 is merely asserted, then we haven't been given an argument from deontologism to internalism, which is our quarry in this chapter.

Thus, the prospects for the argument from deontologism to internalism don't look promising. It won't do to argue from deontologism understood in accord

to believe p and (b) S's cognitive faculties are functioning properly in forming B, then S is aware or potentially aware of the fact that she doesn't believe at t that she ought not to believe p. And 4, together with 3 and 3.5, will yield the desired conclusion, 5.'*

There are three problems with this suggestion. The first is that we no longer have an argument from deontologism to internalism. Instead, we have (at best) an argument from *deontologism and proper functionalism* to internalism. The second problem is that 4* seems false. For the actual or potential awareness it says is required (for the proper function of the cognitive faculties producing B and, given 3.5, for justification) will be either strong awareness or weak awareness. In Chapter 1 I explained why both requirements are implausible. The strong awareness requirement is implausible because of the regress problems it engenders. And the weak awareness requirement is implausible because this sort of awareness seems to offer nothing of significance beyond what is provided already by the *object* of the awareness. (See Chapter 6, Section 3.2 for a discussion of some related points.) The third difficulty is that this suggestion won't be very appealing to internalists because it is entirely contrary to the spirit of internalism to impose a proper function condition without also requiring, for justification, that the subject be aware (or potentially aware) of her belief's satisfaction of the proper function condition. (Anyone who thinks that justification requires the satisfaction of a proper function or reliability condition without also requiring awareness of that satisfaction is a clear example of a noninternalist.) But then the proponent of the deontologism-cum-proper-function defense of internalism will have to explain how deontologism motivates that additional awareness requirement. None of the above versions of the Blamelessness Argument seem to do that and it isn't easy to see how it could be done.

with D_{obj} . And things aren't any more hopeful when we turn to D_{subj} . Moreover, Ginet's argument from deontologism to internalism doesn't work. Nor do the various versions of the Blamelessness Argument discussed above. It's beginning to look, therefore, as if there is no good way to argue from deontologism to internalism.

3.3 Why Deontologism Seems Relevant to Internalism

I've presented objections to what I think of as the most promising arguments from deontologism to internalism. But although those arguments seem to fail, there remains the sense that there is some connection between justification (understood as epistemic blamelessness) and awareness requirements. If we can identify what that connection is and see why it doesn't enable us to argue from deontologism to internalism, then we can be more confident in concluding that deontologism can't be used to defend internalism.

What we noted above is that there is no awareness requirement on blamelessness (epistemic or otherwise). Instead, there is an 'absence of awareness' or 'absence of higher-level belief' requirement on B's blamelessness: you must not believe you are violating a duty in holding B. For this reason, we can't appeal to the necessity for justification of epistemic blamelessness in arguing for an awareness requirement on justification. At best, that would enable us to establish that there is an 'absence of awareness' requirement on justification.

But suppose we focus, as the original rough statement of the Blamelessness Argument does, on epistemic *blameworthiness*. Recall that we are thinking of that notion as follows:

EBW: S's believing p at t is something for which she is *epistemically* blameworthy if and only if *either* (i) S believes at t that she ought not to believe p *or* (ii) S's failure to believe at t that she ought not to believe p is relevantly due to some other doing or failure of hers for which she is epistemically blameworthy.

Earlier we distinguished between original and derived blameworthiness. A belief is originally epistemically blameworthy if clause (i) is true of it; it is epistemically blameworthy in a derived way if clause (ii) is true of it. Thus, S's belief, at t that p , is originally epistemically blameworthy if and only if S believes at t that she ought not to believe p .²⁹ What this shows is that there is an awareness (or higher-level belief) requirement on original epistemic blameworthiness: original epistemic blameworthiness for a belief B requires you to have the higher-level belief that you ought not to hold B.³⁰

²⁹ It is perhaps worth noting that the 'Huck Finn' objection to BW needn't trouble us here. The reasoning laid out here goes through even if we change EBW from an 'if and only if' claim to an 'only if' claim. Likewise, we could change the sentence to which this note is attached from an 'if and only if' to an 'only if' claim without affecting the point I'm making here.

³⁰ No regress results from this requirement because the belief that you ought not to hold B needn't be justified in order for it to result in your being epistemically blameworthy for holding B.

Thus, there *is* a connection between, on the one hand, deontological considerations having to do with epistemic blameworthiness and, on the other hand, awareness requirements. But the connection is due to awareness requirements on original epistemic blameworthiness, not on derived epistemic blameworthiness or on epistemic blamelessness or on justification. One can't infer from the fact that there is an awareness requirement on original epistemic blameworthiness that there is such a requirement on derived epistemic blameworthiness.³¹ For this reason, one can't infer that there are such requirements on epistemic blamelessness (since that requires the absence of both original *and* derived epistemic blameworthiness). And if one can't infer that there are such requirements on epistemic blamelessness, then there seems to be no hope of inferring that there are such requirements on a deontological version of justification. One can't even infer, from the fact that there are awareness requirements on original epistemic blameworthiness, that there are such requirements on the *absence* of original epistemic blameworthiness—a state we might call 'original epistemic blamelessness'.³² All we can say about original epistemic blamelessness in holding B is that it requires the *absence* of awareness of a certain higher-level belief—namely, that one ought not to hold B. But that isn't an awareness requirement. And, in light of the failure of the argument relying on premise 4 near the end of Section 3.2, there doesn't seem to be any good way of deriving an awareness requirement from the requirement that another awareness is absent. In short, the only awareness requirement we can get from deontology is an awareness requirement on original epistemic blameworthiness; and we can't get from there to an awareness requirement on justification. These considerations explain both why it is tempting to agree with the argument from deontology to internalism as well as why that argument is mistaken.

4 STEUP'S ARGUMENTS

In the first three sections, I objected to several attempts to show that deontology yields internalism. In this final section, I will use some of the conclusions established in Sections 1 and 3 in my evaluation of some arguments proposed by Matthias Steup. Steup's work contains some of the most explicit support of the deontological defense of internalism to be found in recent epistemology. I noted

³¹ By recognizing the fallaciousness of this inference, one can see what is wrong with the third sentence of the original (rough) statement of the Blamelessness Argument, which said: *But you aren't blameworthy for something unless you can tell that you ought not to have done it.* For if there is no awareness requirement on derived blameworthiness then there is no reason to endorse that third sentence, even if there is an awareness requirement on original blameworthiness.

³² The attractiveness of this mistaken inference is what tempts people to endorse the fourth sentence in the Blamelessness Argument. That fourth sentence asserts that you can infer from the third sentence (according to which you're blameworthy only if you can tell you've done what you ought not to have done) that you can tell *whether or not* your beliefs are justified or blameless.

above that deontological defenses of internalism have two main components: the claim that deontology about justification is true and the claim that deontology entails internalism. In defense of the first of these claims—deontology itself—Steup has offered impressive responses to two of the main objections to it: (a) the charge that epistemic justification understood deontologically can't adequately account for the plausible view that a belief's being *epistemically* justified must make the truth of that belief probable and (b) the worry that deontology is implausible insofar as it depends on the truth of doxastic voluntarism.³³ But our concern in this chapter has been with the second claim, according to which deontology entails internalism. For this conclusion, Steup has identified two arguments which he has discussed in a number of places: the Direct Argument from deontology to internalism and the Indirect (Evidentialist) Argument from deontology to internalism.

4.1 The Direct Argument

Let's begin with Steup's statement of the Direct Argument:

The Direct Argument from Deontology to Internalism

1. Justification is a matter of epistemic duty fulfillment.
2. Therefore, what determines justification is identical to what determines epistemic duty.
3. What determines epistemic duty is directly recognizable.
4. Therefore, what determines justification is directly recognizable.
5. If what determines justification is directly recognizable, then justification itself is directly recognizable.
6. Therefore, justification is directly recognizable.³⁴

The troublesome premise on which I wish to focus is 3, according to which what determines epistemic dutifulness is directly recognizable. Premise 3 won't seem the least bit plausible if the epistemic duties in question are thought of as objective duties that one can violate blamelessly if one is nonculpably ignorant of them. And as I argued in Section 1.2, for similar reasons 3 won't seem plausible if the duties in question are thought of as subjective duties (since one can be nonculpably ignorant of whether one has fulfilled duties one thinks one has). What we need to do is focus on duties which one cannot violate blamelessly. But if we do that, we may as well focus on justification understood in terms of epistemic blamelessness rather than in terms of epistemic duty.

³³ See Steup (1996: 79–84, 1997: 381–3) for his response to the first objection and Steup (2000) for his response to the second objection. (In saying these responses are impressive, I'm not thereby endorsing them.)

³⁴ This argument is taken verbatim from Steup (2001b: section 5).

We could revise Steup's Direct Argument so that it focused on epistemic blamelessness instead of epistemic duty by changing the first three premises as follows:

- 1*. Justification is a matter of epistemic blamelessness.
- 2*. Therefore, what determines justification is identical to what determines epistemic blamelessness.
- 3*. What determines epistemic blamelessness is directly recognizable.

But why think 3* is true? I've shown in Section 3 that awareness requirements on original epistemic blameworthiness don't imply such requirements on epistemic blamelessness. And yet apart from appealing to the awareness requirements on original epistemic blamelessness, it's difficult to see why one would think that epistemic blamelessness (or what determines it) is directly recognizable. I conclude, therefore, that Steup's first argument stumbles at premise 3*: in light of my arguments from Section 3, that premise is without adequate motivation.

4.2 The Indirect Argument

The Indirect (Evidentialist) Argument from deontology to internalism has two parts: a subargument from deontology to evidentialism and another subargument from evidentialism to internalism. Steup gives two quite different versions of the Indirect Argument and two quite different versions of each of the subarguments that compose it. I'll call these two versions the 'Stanford version' and the 'Oxford version' because the first version of the Indirect Argument with its accompanying subarguments appears in the online *Stanford Encyclopedia of Philosophy* and the second appears in an Oxford University Press anthology edited by Steup.³⁵

4.2.1 The Stanford Version

As I just noted, both the Stanford version and the Oxford version of the Indirect Argument are composed of two subarguments. After laying out the Stanford version of the subargument from evidentialism to internalism (which I won't give here), Steup turns to the other subargument—the one from deontology to evidentialism—which begins with the deontologist's assumption that justification is a matter of not violating one's epistemic duties. He tells us that these duties are ones that stipulate what one ought to do in the pursuit of truth. He then asks this question: 'in the pursuit of truth, exactly what is it that one ought to do?' Here is his answer:

Evidentialists would say: it is to believe what, and only what, one has evidence for. Now if that is one's epistemic duty, then those who take justification to be deontological can

³⁵ See Steup (2001a, 2001c).

employ the argument considered above (which proceeds from evidentialism to J-internalism) to derive the conclusion that deontological justification is internal justification. So the combination of deontology about justification with evidentialism allows for a pretty straightforward derivation of J-internalism. (2001c: section 5)

(J-internalism is just internalism about justification.) Notice what has happened here—or rather, what hasn't happened. Steup hasn't really given us an argument from *deontologism* to internalism. Instead, he simply tells us (in the opening claim of the passage just quoted above) that evidentialists will interpret deontologism in such a way that evidentialism is true; and they can then use the other subargument to get from evidentialism to internalism. Steup himself admits as much when he says that he's given us an argument for internalism from the *combination* of deontologism and evidentialism, rather than from deontologism itself. In fact, it is more like an argument from evidentialism alone to internalism. The deontologism seems like dead weight in the argument. Why even bother noting that evidentialists will interpret *deontologism* in such a way that evidentialism is true and that evidentialism can then be used to get internalism? Why not simply note that evidentialists will accept *evidentialism itself* and that once you get evidentialism, you can combine it with the argument from evidentialism to internalism to get internalism?

Thus, the Stanford version of the Indirect Argument from Deontologism to Internalism isn't an argument from *deontologism* to internalism. Instead, it's an argument from evidentialism to internalism.³⁶ Let's turn now to the Oxford version of the Indirect Argument to see if it fares any better.

4.2.2 The Oxford Version

The Oxford version of the Indirect Argument begins with the following subargument from deontology to evidentialism:

But can evidentialism (as I have construed it here) be derived from deontology? It seems to me it can. The derivation is a complex one and an adequate account and defense of it would go beyond the confines of this essay. I will therefore merely outline how it would go. It involves three arguments. The first is an ingredient of Bonjour's well-known counterexamples to reliabilism. If some item *x* is to make it epistemically responsible of a person *S* to believe that *p*, then *x* must be in *S*'s cognitive possession. The second argument would have to establish that there is one and only one way in which *x* can come into *S*'s cognitive possession: *S* must be in an evidential state that constitutes evidence for *x* and thus enables *x* to justify *S* in believing that *p*. The job of the third argument would

³⁶ And although I won't take the time to explore this here, the Stanford version of the argument from evidentialism to internalism seems to rely on a premise that is in as much need of defense as internalism itself, namely, the premise that justification is determined by one's evidence (where one's evidence is thought of as one's accessible or directly recognizable mental states). See the argument from evidentialism to internalism in Steup (2001c: section 4).

be to show that, if this condition is met, then x is either an evidential state itself or a mere fifth wheel: an item that plays no role as a justifier for S 's belief that p since all the needed justificatory work is being done by S 's evidence for x . (Steup 2001a: 137)

As he says, this is just an outline of this subargument—one based on three further supporting arguments which together are supposed to yield the conclusion that deontologism entails evidentialism. Let's try to get clear on the structure of this subargument.

The conclusion of the first of the three additional supporting arguments connects epistemic responsibility and cognitive possession:

ER/CP: What makes one epistemically responsible is within one's cognitive possession.

The conclusion of the second of the three further supporting arguments connects cognitive possession with evidence one has:

CP/EV: Something is within one's cognitive possession if and only if one *has* evidence for it.

At first glance, CP/EV seems implausible. Or at least it seems implausible if we combine it with another claim Steup seems to accept, namely, that a person can have evidence itself in her cognitive possession.³⁷ For the combination of that claim with CP/EV entails that if you have any bit of evidence in your cognitive possession, you also *have* (presumably within your cognitive possession) evidence for that evidence. This seems to entail that, so long as you have one bit of evidence in your cognitive possession, you have in your cognitive possession an infinite hierarchy of ever more complex bits of evidence.

But perhaps this problem is easily solved. Consider the result of replacing CP/EV, the conclusion of the second of the three supporting arguments Steup relies on, with

CP/EV*: Something is within one's cognitive possession if and only if it *is* evidence one has.

Not only does CP/EV* avoid the regress problems facing CP/EV, it also seems to be just the conclusion Steup needs to plug into the Oxford version of the subargument from deontologism to evidentialism because with it, the need for his additional third supporting argument disappears. For by starting with deontologism and adding to it ER/CP and CP/EV* (together with a definition of evidentialism), we get the following:

³⁷ In Steup (2001c: n. 25) he talks of evidence one possesses; presumably he is thinking here of evidence in one's cognitive possession. And in the passage I quote in the previous paragraph in the text above, he allows that x itself might be an evidential state (and x is clearly something he thinks might be within one's cognitive possession).

Oxford Version of the Subargument from Deontologism to Evidentialism

1. Justification is a matter of being epistemically responsible.
2. ER/CP: What makes one epistemically responsible is within one's cognitive possession.
3. Therefore, what determines justification is within one's cognitive possession.
4. CP/EV*: Something is within one's cognitive possession if and only if it is evidence one has.
5. Therefore, what makes one justified is evidence one has.
6. Evidentialism is the view that what makes one justified is evidence one has.
7. Therefore, evidentialism is true.

Given that the main premises of this argument—namely, ER/CP and CP/EV*—are the conclusions of the first two supporting arguments Steup mentions, it seems to capture the essence of the argument he is giving in the extended quotation two paragraphs back.

We don't need to bother with the Oxford version of the subargument from evidentialism to internalism.³⁸ For the Oxford version of the subargument from deontologism to evidentialism seems to suffer from a familiar problem. According to its second premise (i.e. ER/CP), what makes one epistemically responsible is within one's cognitive possession. It seems to me that this premise gets all its plausibility from reasoning of the following sort:

A Defense of ER/CP

1. S isn't epistemically blameworthy if she can't tell she's done anything wrong.
2. Therefore, S isn't epistemically irresponsible unless she can tell she is.
3. Therefore, S isn't epistemically responsible unless she can tell she is.
4. Therefore, what makes one epistemically responsible is within one's cognitive possession.

But there are two familiar problems with this line of reasoning. First, its third premise doesn't follow from its second premise. What follows instead is something like:

- 3*. Therefore, S isn't epistemically responsible unless it's false that she can tell she is epistemically irresponsible.

An awareness requirement on epistemic irresponsibility yields an 'absence of awareness' requirement, not an awareness requirement, on epistemic responsibility (just as an awareness requirement on epistemic blameworthiness yields an 'absence of awareness' requirement, not an awareness requirement, on epistemic blamelessness). And, for reasons of the sort discussed in Section 3, there

³⁸ That argument is very short: Steup simply *defines* internalism as evidentialism. He rejects an account of internalism in terms of accessibility and replaces it with an account in terms of evidential states. See Steup (2001a: 135–7).

doesn't seem to be much hope of deriving an awareness requirement from an absence of awareness requirement. The second problem is that the first premise in the Defense of ER/CP is mistaken. For as I pointed out in Section 3, a person can be currently epistemically blameworthy due to past blameworthiness to which she currently lacks any special access. Current epistemic blameworthiness (and epistemic irresponsibility) can be *derived*, in which case there's no guarantee that the subject is now aware of it (even if she once was). A defender of ER/CP could insist that it is plausible independently of the Defense of ER/CP which I just criticized. But once you recognize that premise 1 of that Defense is false and that its third premise doesn't follow from its second premise, ER/CP seems to lose whatever initial plausibility it seemed to have.

Thus, the Oxford version of the Indirect Argument from deontology to internalism fails too. For its component subargument from deontology to evidentialism relies on the premise that what determines one's being epistemically responsible is within one's cognitive possession. And that claim is no more adequately motivated than the claim, undermined in Section 3, that epistemic blamelessness is subject to awareness requirements.

Our goal in this chapter has been to consider whether deontology could replace the main motivation for internalism that was rejected in Chapter 1. I looked at many different arguments from deontology to internalism. The variety in the arguments was due in part to the fact that they didn't all begin with the same version of deontology. Some versions of deontology I considered conceived of justification in terms of epistemic duty, others conceived of it in terms of epistemic obligation, and still others conceived of it in terms of epistemic blamelessness or epistemic responsibility. None of the arguments from deontology to internalism succeeded. Moreover, we can see why the idea that deontology entails internalism is so appealing despite the fact that it's mistaken: it's appealing because there *is* an awareness requirement on a certain kind of epistemic blameworthiness and it's tempting (though mistaken) to conclude from that that there is also an awareness requirement on epistemic blamelessness (and, therefore, on any sort of justification which requires such blamelessness).

This page intentionally left blank

FOR EDUCATIONAL PURPOSES ONLY

PART II

DEFENDING EXTERNALISM

FOR EDUCATIONAL PURPOSES ONLY

This page intentionally left blank

FOR EDUCATIONAL PURPOSES ONLY

5

Proper Function

Part I of this book was devoted to my attack on internalism. The main argument of Chapter 1 ran as follows:

- I. An essential feature of internalism is that it makes a subject's actual or potential *awareness* of some justification-contributor a necessary condition for the justification of any belief held by that subject.
- II. The awareness required by internalism is either *strong* awareness or *weak* awareness.
- III. If the awareness required by internalism is strong awareness, then internalism has vicious regress problems leading to radical skepticism.
- IV. If the awareness required by internalism is weak awareness, then internalism is vulnerable to the SPO, in which case internalism loses its main motivation for imposing the awareness requirement.
- V. If internalism either leads to radical skepticism or loses its main motivation for imposing the awareness requirement (i.e. avoiding the SPO), then we should not endorse internalism.
- VI. Therefore, we should not endorse internalism.

Since premise (II) is uncontroversial, only four premises needed defending. In the first chapter I gave arguments for premises (III) and (IV) and offered a partial defense of premise (V). Chapter 2 was aimed at showing that not even those internalists sensitive to the similar-sounding *Sellarsian* dilemma for internalist foundationalism are able to escape the dilemma for internalism presented in Chapter 1. The purpose of Chapter 3 was twofold: I defended premise (I) by arguing against mentalist accounts of internalism which conflict with the account of internalism proposed in premise (I); then, having argued that mentalism is different from internalism, I argued against mentalism itself (thereby objecting to the main nonexternalist position besides internalism). Finally, in Chapter 4, I added to my defense of premise (V) by arguing that deontologism was an inadequate replacement for the main motivation for internalism.

Having argued in Part I against the two main competitors to externalism—internalism and mentalism—I will turn in Part II of this book to a defense of externalism. I won't defend externalism by identifying necessary

and sufficient conditions for being an externalist and arguing that in order for a theory to be correct it must satisfy those conditions. (In Chapter 1, I identified a necessary condition for being an internalist, indicating that I didn't think it was advisable to seek a complete list of necessary and sufficient conditions for being an internalist. I have the same reluctance about trying to identify a complete list of necessary and sufficient conditions for being an externalist.) Instead, I will defend externalism by defending a view that is clearly a version of externalism. By defending that view, I will be defending externalism itself, even though I'll be making no attempt to give an analysis of externalism.

According to the version of externalism I'll be defending, the two most important requirements on justification are the satisfaction of a *proper function* condition and the absence of a *defeater*. The first requirement will be the focus of this chapter. The topic of defeaters will be taken up in Chapter 6. After explaining and giving arguments for this version of externalism, I will turn, in Chapters 7 and 8, to a consideration of two objections that have been proposed against all versions of externalism: that externalists are committed to permitting an objectionable kind of epistemic circularity and that the externalist response to skepticism is inadequate in a variety of ways. These two objections can be viewed as further motivations for internalism if one assumes that internalists can avoid them. I will argue that switching one's allegiance from externalism to a nonexternalist position doesn't enable one to avoid those two objections (unless one adopts a position that, for other reasons, is wholly implausible). This provides still more support for premise (V) from Chapter 1 because it shows that these proposed motivations for internalism are also inadequate replacements for the main motivation debunked in Chapter 1. Moreover, I will argue in Chapters 7 and 8 that although some of the charges associated with these two objections do apply to externalism, they don't provide us with a good reason to reject externalism.

The goal of *this* chapter is to defend a proper function account of justification. In an attempt to get my internalist opponents on the road to recognizing the plausibility of my conclusion, I will begin with the assumption, congenial to internalists, that the following evidentialist thesis is, initially at least, an attractive account of justification:

EF: S's belief B is justified iff B is a *fitting* doxastic response to S's evidence.¹

¹ The more standard way of stating the evidentialist's thesis is to have the right-hand side of the 'iff' say something like 'B fits S's evidence' (see Feldman and Conee 1985). I'm assuming that doxastic response B is a fitting response to S's evidence only if (i) B fits S's evidence and (ii) B is formed in response to S's evidence. Clause (ii) is a way of adding a 'formed in the right way' requirement. Since justification, as I'm thinking of it in this book, is *doxastic* justification rather than merely *propositional* justification, the addition of such a requirement is quite unremarkable. (See Chapter 1, Section 1 for a discussion of the difference between doxastic and propositional justification.)

In Section 1 of this chapter, after saying a little about what the evidentialist position on justification is, I will argue that that position needs to be improved in three ways. Then, in Section 2, I will defend a proper function analysis of justification that captures the improvements recommended in Section 1. In the final section, I will respond to some objections to my proposed analysis of justification.

But before moving on, I'd like to highlight one additional benefit of this chapter: it shows that one of the standard objections to externalism is misguided. According to this objection, the unreliably formed beliefs of evil demon victims can be as justified as our own beliefs. This is supposed to create a problem for externalists because it is assumed that they impose a reliability condition on justification. But as I will make clear below, the proper function account of justification I give is an externalist account that doesn't impose a reliability condition on justification. In light of this we can see that this standard objection, while perhaps effective against reliabilists about justification, is not effective against externalism about justification.²

1 FROM EVIDENTIALISM TO PROPER FUNCTION

1.1 Evidentialism

Let's begin by trying to get a better understanding of evidentialism. What exactly counts as a subject's evidence? One proposal in the literature is that evidence is limited to things of which the subject is aware or potentially aware (i.e. to epistemically accessible things).³ Another proposal is that evidence is limited to the subject's mental states.⁴ Taking these as our guide, we may conclude that, in order for a belief to be a fitting doxastic response to a subject's evidence, it must be a fitting doxastic response to something of which that subject is (at least potentially) aware or to some of her mental states. I will assume that we have some grasp of what it is for a belief to *fit* one's evidence and that, with respect to certain clear cases, we have some fairly firm intuitions concerning whether the belief in question does or doesn't fit the subject's evidence. The very ease with which we make sense of the examples used by evidentialists to illustrate their position testifies to the plausibility of these last two assumptions. When, for example, Feldman and Conee (1985: 15) say that the belief that *there is something green before one* fits the evidence that a normal person has in ordinary circumstances when looking

² Nor is it effective against reliabilist accounts of warrant (that which makes the difference between knowledge and mere true belief) since demon beliefs are not typically supposed to be as *warranted* as our beliefs, even if they are typically supposed to be as *justified* as our beliefs. And given that reliabilists more often give analyses of warrant than of justification, this standard objection doesn't seem to apply to too many externalists.

³ See Feldman (1988a) for a discussion of this proposal.

⁴ Steup (2001a: 137). See also Conee and Feldman (2001: 231–9) where they suggest this reading.

at a plush green lawn in broad daylight or that the belief that *sugar is sour* does not fit our gustatory experience, we have no trouble understanding or accepting such claims.

A standard line taken by evidentialists involves endorsement of the following three claims:⁵

Nonreliability: the fittingness of doxastic response B to evidence E is not contingent upon E's being a reliable indicator of B's truth.

Objectivity: the fittingness of doxastic response B to evidence E is objective fittingness (in the sense that fittingness from the subject's perspective isn't sufficient for it).

Necessity: the fittingness of doxastic response B to evidence E is an essential property of that response to that evidence.⁶

Nonreliability is extremely plausible. Just as it seems possible for a demon victim who is my mental twin to have justified beliefs despite the fact that, in the twin's case, those beliefs are unreliably formed, so also it seems possible for a demon victim's beliefs to fit her evidence. The demon victim's problem has to do not with the fit between beliefs and evidence but with the connection between evidence and the world. So I join the evidentialist in accepting Nonreliability and in rejecting reliabilist accounts of justification. I also accept Objectivity (a position to be explained in greater detail below). Thus, I follow the evidentialist in thinking that there is an important sort of justification that is objective in the sense that innocently thinking all is well epistemically is insufficient for it. I made clear my agreement with evidentialists on Objectivity at the beginning of the book where I identified the sort of justification in which I was interested. Putting all this together, we can say that I join the evidentialist in denying that objective fittingness depends on reliability.

But then what *does* it depend on? The evidentialist's answer to this question is given by Necessity: the objective fittingness of doxastic response B to some evidence E is an *essential* property of that response to that evidence; it isn't a contingent property that depends on some other condition being satisfied. But notice that this answer isn't *forced* upon those who agree with evidentialists about Nonreliability and Objectivity. For the fact that the fittingness of doxastic response

⁵ Of these three claims—Nonreliability, Objectivity, and Necessity—the first is the one most obviously endorsed by evidentialists. Below I explain why I think Conee, Feldman, and other evidentialists believe in Objectivity and Necessity. Who counts as an evidentialist? Perhaps the clearest examples are Conee and Feldman (2004). But evidentialist sympathies can also be discerned in the writings of Chisholm (1977), Moser (1985, 1989), Fumerton (1995), Haack (1993, 1997), Russell (2001), and Steup (2001a).

⁶ The idea is that if B is a fitting response to E by itself, then, even if B could be an unfitting response to evidence that includes E *and more besides*, it couldn't be an unfitting response to E by itself.

B to evidence E isn't contingent upon E's being a reliable indicator of B's truth doesn't prove that it isn't contingent upon anything. So Nonreliability doesn't entail Necessity. I will be arguing that Necessity is false.⁷ In fact, the first of the three improvements to evidentialism that I'll be recommending below is that it drops its endorsement of Necessity.

In order to have a good understanding of my three recommended improvements to evidentialism, it will be helpful to recall the two questions employed in Chapter 3 for the purpose of classifying theories of justification:

- (A) What sorts of things can be appropriate inputs to our belief-forming processes?
- (B) Must the appropriateness of such an input be due to an essential property of that input?⁸

Evidentialists, by endorsing Necessity, give an affirmative answer to (B).⁹ My first recommended improvement to evidentialism is that it replace its affirmative answer to (B) with a negative answer to that question. The second recommended improvement arises out of my argument for the first improvement. The counterexamples I use to show that Necessity is false suggest that the fittingness mentioned in E_F should be understood in terms of proper function. Thus, my second recommended improvement is that we understand that fittingness in that way. The third improvement I will recommend makes it utterly clear, if it wasn't already, that if evidentialism is altered in accord with my recommendations, it will no longer be an *evidentialist* position. For my third recommendation is that we don't give an evidentialist answer to question (A). An evidentialist answer to (A) says that only evidential states can be appropriate inputs to our belief-forming processes. Given that only accessible states or only mental states can be evidence, the evidentialist answer to (A) says, in effect, that only accessible states or only mental states can be appropriate inputs to our belief-forming processes. I recommend that we reject those answers in place of a noninternalist, nonmentalistic answer that fits with externalism. The remainder of Section 1 will be devoted to defending these three recommended improvements to the evidentialist position.

⁷ Notice that in rejecting Necessity I am not saying that the relation of supervenience holding between justification and its supervenience base is contingent. For even if it were true that belief B's *justification* supervened (of necessity) on the fittingness of doxastic response B to the subject's evidence (which happened to be E), that wouldn't entail that the *fittingness* of doxastic response B to evidence E was an essential property of that response.

⁸ For an input to be appropriate is for it to be capable of resulting in a justified belief.

⁹ Feldman is a prime example. See the discussion in last two paragraphs of Section 1.3.2 of Chapter 3 where I describe how he defends his version of mentalism by highlighting what he takes to be the virtues of his affirmative answer to question (B). And given Feldman's remarks about what mentalism and internalism are (I discuss these remarks in Section 1.2.2 of Chapter 3), it is clear that he thinks an affirmative answer to question (B) is widely endorsed, since he thinks it is a crucial component of being an internalist and a mentalist.

1.2 The First Improvement: Drop Necessity

My argument for dropping Necessity involves the use of a counterexample that is best understood in light of some of the implications of Objectivity. I'll begin, therefore, by considering what some of those implications are.

1.2.1 Objectivity

According to Objectivity, the fittingness *from the subject's perspective* of doxastic response B to her evidence isn't sufficient for B's *actually* fitting her evidence. In order to explain this claim, it will be helpful to distinguish between the main evidence for a belief, on the one hand, and connectors and disconnectors on the other.

Consider a person Jane who has managed somehow to become an adult without any exposure (directly or via testimony) to the phenomenon of a straight object looking bent when immersed in water. And suppose that Jane for the first time comes upon a straight stick immersed in water and, upon seeing it, forms the false belief that it is bent. The natural thing to say is that Jane's evidence for this belief is her visual experience (which, we may assume, is like the visual experience you and I have when we see such a thing). Consider next Tim, a logically perceptive man who learns from an informer who delights in obfuscation that

(1) If John goes to the party then {if Judy goes to the party then {if Jan goes to the party then the party will be a lot of fun} }.

Upon learning this, Tim immediately forms the belief that

(2) If John and Judy and Jan go to the party, then the party will be a lot of fun.

In this case, the natural thing to say is that Tim's evidence for his belief that (2) is his belief that (1). I'll refer to evidence of this sort—the sort we are naturally inclined to point to when identifying a person's evidence—as 'the main evidence' for a belief.

Now consider the view that the main evidence for our beliefs doesn't always constitute our total relevant evidence for them. For example, in the Jane case, there may be, in addition to the visual experience on which she bases her belief that the stick is bent, the strong *felt* inclination to take her visual sensations as indicative of the truth of the belief in question.¹⁰ Those of us who are completely familiar with the phenomenon of water distorting the appearance of a straight stick lack this felt inclination (at least it isn't as strong in us). As a result, our total evidence is somewhat different from Jane's. We can call Jane's strong

¹⁰ A felt inclination to do X is an experience of feeling inclined to do X. It isn't merely a disposition.

felt inclination a ‘connector’ between her main evidence and her belief based on that evidence. It’s a connector we lack. Another (related) difference is that we are inclined to think that the main evidence on which Jane relies is *not* indicative of the truth of the belief she bases on it. This provides us with a *disconnecter* between the main evidence and the belief in question—a disconnecter Jane lacks.¹¹ So if we think of these connectors and disconnecters as parts of one’s total evidence, then *our* total evidence when we view a stick looking bent in water is different from *Jane’s* when she considers such a sight (even if the main evidence—the way it visually appears—is the same).

In the Tim case there could be a connector too. Let us suppose it is the strong felt inclination to take the truth of (1) to entail the truth of (2). It is easy to imagine those less logically perceptive than Tim failing to see this sort of connection. We could describe this difference by saying that although these others have the same main evidence Tim has, they don’t have the same total relevant evidence since they lack a connector he possesses.¹²

With this terminology at our disposal, we can return to our discussion of Objectivity. The view with which proponents of Objectivity disagree is the subjectivist view according to which a belief fits the subject’s evidence if (though perhaps not only if) the subject’s evidence consists of both her main evidence and a connector that connects her main evidence to her belief. In opposition to this suggestion, the supporter of Objectivity claims that if the subject’s belief fails to fit the subject’s main evidence, it won’t help merely to add to her evidence a connector connecting that main evidence with the belief. Consider, for example, the following belief and pieces of evidence:

Belief	Main Evidence	Connector
B1. The first person belief: ‘There is smallish hard round object in my hand’.	ME1. Tactile sensations of the type you experience when you grab a billiard ball.	C1. The strong felt inclination to take ME1 to be indicative of the truth of B1.
	ME2. Olfactory sensations of the type you experience when you smell a meadow full of flowers. ¹³	C2. The strong felt inclination to take ME2 to be indicative of the truth of B1.

¹¹ Thus, disconnecters will be a sort of undercutting defeater (see Chapter 6 for a discussion of undercutting defeaters). Notice that *lacking a connector* between a belief and one’s main evidence isn’t by itself sufficient for *having a disconnecter*.

¹² I am *not* suggesting that justification *requires* a connector between one’s beliefs and one’s main evidence. I’m just explaining what connectors are and noting that they may be present in some cases.

¹³ Since a meadow full of flowers usually has a number of different kinds of flower (as well as a variety of nonflowering plants) each of which has a distinctive odor, the type of olfactory

The subjectivist opponent of Objectivity might acknowledge that B1 fits ME1 and that it doesn't fit ME2. But she will add that B1 *does* fit the combined evidence of ME2 together with C2 because by adding C2 to ME2 (and adding no disconnectors¹⁴) we get subjective fittingness. The proponent of Objectivity rejects this suggestion. She says that since B1 doesn't fit ME2, it won't help merely to add C2 even if we stipulate that no disconnectors are present. The objective failure of B1 to fit the subject's evidence (ME2) isn't altered by adding C2 to the subject's evidence base; if B1 fails to fit ME2, it also fails to fit ME2 + C2 (the combination of sensation ME2 and connector C2).

Paradigm evidentialist Richard Feldman appears to be endorsing Objectivity so construed when he points out that one can't get justification for one's beliefs merely by thinking one has good reasons for them.¹⁵ Other evidentialists such as Bruce Russell, Paul Moser, and Richard Fumerton also seem to agree with Objectivity. For reliability isn't the only condition that they think is insufficient for E's being good evidence for B; they pay the same compliment to *thinking that E is good evidence for B*.¹⁶ In each case, the idea seems to be that if a belief doesn't fit one's main evidence, then merely adding a connector joining that belief with that main evidence won't help.

1.2.2 Learned and Unlearned Doxastic Responses

A possible concern about Objectivity arises in connection with its suggestion that if a belief doesn't fit the subject's main evidence then merely adding a connector won't help. Consider B1 and ME2. Why couldn't a person learn in some strange setting to associate olfactory experience ME2 with the truth of B1, in the way one learns to associate certain smells with gasoline or paint (i.e. by experiencing their constant conjunction)? Then one would have learned to have C2 whenever one has ME2 in which case B1 could be a fitting response to ME2 + C2 after all.

experience one has upon smelling a meadow full of flowers has many components. In this way it resembles the tactile experience one has upon grabbing a billiard ball, an experience that also has many components.

¹⁴ How could a person's total relevant evidence for B1 consist of only ME2 and C2 without any disconnectors? Wouldn't she glance at her hand or attempt to use the hand in which she thinks there is a small hard ball? Wouldn't others inform her that it is crazy for her to think, on the basis of ME2, that she is holding a small hard ball? These are certainly possibilities, likely ones even. But all we need is a possible example where no such disconnectors are present.

¹⁵ See Feldman (1988b: 411). See also Feldman and Conee (1985: sects. II–III) where they say that if B fails to fit one's evidence, it won't help to add subjective factors such as *trying one's best to hold only beliefs that fit one's evidence or being blameless in holding B*. They insist that even with those factors present, B still fails to fit the subject's evidence.

¹⁶ See Russell (2001: 37–8), especially his comments on the believers raised in the benighted religious community whom he thinks of as subjectively justified but not objectively justified. See also Moser (1989: 38–42, 47–52, 202–3) and Fumerton (1995: 8–20, 113–16, 183–224) where they emphasize that merely believing that one's evidence E is good evidence for B isn't sufficient for B's justification since E might not *in fact* make B objectively probable.

That concern about Objectivity strikes me as a sensible one. To handle it, we need to distinguish learned from unlearned doxastic responses. The distinction isn't easy to draw but it is something like this. *Learned* doxastic responses, such as an experienced birdwatcher's immediate bird identifications after a quick look (or listen), are ones a person comes to have only after first finding out independently (i.e. without relying in any essential way on other instances of that same type of doxastic response) that there is a correlation between the truth of such beliefs and the experiences to which they eventually become immediate responses. By contrast, an *unlearned* doxastic response to experience is a hardwired or automatic response that occurs (perhaps only after a certain level of cognitive development) without the subject first independently finding out that there is a correlation between the truth of the belief in question and the experience to which it is a response.¹⁷

We can also distinguish learned from unlearned *connectors*. Suppose C is a connector that connects main evidence ME with a belief B. If C is acquired by first independently finding out that there is a correlation between the truth of B and the occurrence of ME, then C is a *learned* connector. However, if C isn't acquired by first learning independently of a correlation between the truth of B and the occurrence of ME—if, instead, a person simply has C without learning independently that there is such a correlation—then C is an *unlearned* connector. With these distinctions in mind, we can handle the concern mentioned at the beginning of this subsection by explaining Objectivity more carefully as follows: some unlearned doxastic responses to one's evidence are unfitting and merely adding an *unlearned* connector to one's original evidence (one which connects that doxastic response to that original evidence) won't change that.

But there's a related concern that remains—one prompted by the thought that someone might have an unlearned connector which she wrongly takes to be a learned connector. Suppose C is an unlearned connector that connects S's main evidence ME with S's belief B. But suppose also that it (mistakenly) seems to S that she learned to connect ME with B by first learning independently of a correlation between the truth of B and the occurrence of ME. (In actual fact, she didn't learn this. She just came to have the connector in question without any learning going on.) Then although C is an unlearned connector, it seems to S to be a learned connector. We can call these sorts of unlearned connectors 'seemingly-learned unlearned connectors'. It is arguable that seemingly-learned unlearned connectors between ME and B are as helpful (or nearly as helpful) as learned connectors between ME and B: in each case, adding the connector in question to ME seems to provide S with something to which B is a fitting response.¹⁸

¹⁷ This distinction is similar to and inspired by Reid's distinction between original and acquired perception. See Reid ([1785] 2002: 235–9, 1997: 171–2). It is basically the same as Goldman's (1986: 93–5) distinction between beliefs produced by native processes and beliefs produced by acquired methods.

¹⁸ Thanks to Michael Rea for drawing this objection to my attention.

To handle this last concern, we can explain Objectivity even more carefully as follows: some unlearned doxastic responses to one's evidence are unfitting and merely adding an *unlearned* connector to one's original evidence—one that (a) connects that doxastic response to that original evidence and (b) isn't seemingly-learned—won't change that. Recall the original statement of Objectivity:

Objectivity: the fittingness of doxastic response B to evidence E is objective fittingness (in the sense that fittingness from the subject's perspective isn't sufficient for it).

Objectivity says that doxastic response B to E can seem appropriate to the subject—it can feel right—even when it is objectively unfitting. And the explanation at the beginning of this paragraph identifies the sort of 'feeling right' that is insufficient for objective fittingness. Sometimes a doxastic response to evidence feels right because it is a seemingly-learned response. Other times it feels right even though it isn't a seemingly-learned response. When an unfitting unlearned doxastic response comes (in an unlearned way) to feel right in a way that isn't seemingly-learned, that is not enough to make it a fitting doxastic response. That is the thought behind Objectivity. It opposes the view that a doxastic response to evidence is fitting so long as it feels right to the subject.

1.2.3 A Counterexample to Necessity

I accept Objectivity. And, as I noted earlier, because I also accept Nonreliability, I follow the evidentialist in denying that objective fittingness depends on reliability. This brings us to the all-important question: 'What *does* it depend on?' You will recall that the evidentialist's Necessity-inspired answer was that the objective fittingness of doxastic response B to some evidence E is an *essential* property of that response to that evidence, not a contingent property that depends on some other condition being satisfied. In this subsection, I will be presenting a counterexample to Necessity.

To do this, I will need to employ some assumptions implicit in Objectivity. In particular, I will be assuming that, just as there can be unlearned doxastic responses that are objectively fitting, so too there can be ones that are objectively *unfitting* (in the ways described in the previous two subsections). What follows is a description of a possible case in which the billiard ball belief B1 is a fitting unlearned doxastic response to ME2 and an unfitting unlearned doxastic response to ME1.¹⁹ Such a case is a counterexample to Necessity because in the

¹⁹ I call B1 a 'billiard ball belief' for convenience even though the concept of a billiard ball is not included in its content.

actual world, B1 is a fitting unlearned doxastic response to ME1 and an unfitting unlearned doxastic response to ME2.²⁰

Thomas Reid emphasized that there doesn't seem to be any logical connection between our sense experiences and the content of the beliefs based on them. For example, the tactile sensations we experience when touching a hard surface seem to have no logical relation to (nor do they resemble) the content of the hardness beliefs they prompt.²¹ In light of this he said that 'no man can give a reason why the sensations of smell, or taste, or sound, might not have indicated hardness' (Reid [1764] 1997: 57). Considering the matter in the abstract, tactile sensations do not seem to be any more suited than olfactory sensations to being indicators of hardness. Thus, it seems there could have been cognizers like us in outward appearance who experience, upon grabbing a billiard ball, a sensation that is qualitatively of the same type as one of our actual world sensations of smell.²² And it seems possible that the natural *unlearned* doxastic response of such a cognizer to that 'olfactory'²³ sensation is the first-person belief 'There is a smallish hard round object in my hand' (i.e. B1). There is nothing about the process of *grabbing a billiard ball then experiencing ME2 then holding B1* that makes it intrinsically less suitable (as a natural unlearned process for a cognizer to undergo) than the process of *grabbing a billiard ball then experiencing ME1 then holding B1*. In each case, there is a causal chain from an external stimulus to an experience to a belief. And in each case, the experience has the same functional role of connecting that stimulus with B1. The only difference is that in the one case, the experience playing this functional role is ME1 whereas in the other it is ME2. (I should add that the cognizers in this Reidian example are such that *all* of the experiences produced in them by the sorts of activities that produce tactile experiences in us are what we would call 'olfactory experiences'. So it isn't just

²⁰ Objection: Is B1 really an *unlearned* doxastic response to ME1 for normal humans? Perhaps psychological research will show (or has shown) that it is *learned*.

Reply: Maybe so. But suppose psychological research had shown that B1 is (as Reid believed) an *unlearned* doxastic response to ME1. Would that have given us a reason to take such beliefs to be unjustified? No. So it seems possible for there to be creatures for whom B1 is justified in virtue of being believed on the basis of ME1 even though B1 is (for them) an *unlearned* doxastic response to ME1. The possibility of such creatures is all I need to make my case.

²¹ Reid defends this view concerning hardness—along with a similar view concerning the relation of tactile (and proprioceptive) sensations to beliefs in extension, motion, and shape—in Reid ([1764] 1997: Chapter 5, Sections 2–6). For a superb discussion of Reid's views on perception see Wolterstorff (2001), especially Chapters 5 and 6.

²² Perhaps sensations supervene on brain states. Then we have to imagine, first, that sensations of the same qualitative type as our olfactory sensations could supervene on physical states (call them X-states) other than those on which ours supervene (cf. the Martians in Lewis 1980). Second, we imagine a creature who is like us in outward appearance and who, upon grabbing a billiard ball in its hand, causes itself to go into an X-state.

²³ I use quotation marks to indicate that although the sensation in question is qualitatively of the same type as *our* olfactory sensations, it may not, in the circumstances in question, be properly thought of as an olfactory sensation in the ordinary sense (since that sense may have functional role implications).

B1's ground that is different in this way. Likewise, none of the things that produce in us the experiences we call 'olfactory experiences' produce those sorts of experiences in them.)

This Reidian example suggests *two* things: first, that it is possible for billiard ball belief B1 to be an *unlearned* doxastic response to 'olfactory' sensation ME2; second, that it is possible for such an unlearned response to be natural for a cognizer, even an entire species of cognizers. The first claim seems relatively uncontroversial—the possibility of a certain kind of cognitive malfunction in humans entails it. But the second seems plausible too.²⁴ It is no less plausible than the possibility of a species of cognizers whose experienced color spectrum is inverted with respect to ours. (In the following subsection, I consider an objection to the possibility of the Reidian example—an objection posed by intentionalists about sensory experience.)

It might be helpful here to say something in support of a crucial feature of this objection to Necessity, namely, its suggestion that brain-damaged humans could have the *same* evidence base for B1 (i.e. ME2) as do the possible cognizers in the Reidian example. Consider first normal humans. It seems that, for them, B1 is an unlearned doxastic response to ME1—something that occurs without their first learning *independently* that hardness is correlated with such tactile experiences.²⁵ Similar remarks apply to the possible cognizers in the Reidian example: B1 is, for them, an unlearned doxastic response to ME2. And the same thing can be said of certain brain-damaged humans: due to some sort of injury, B1 is, for them, an *unlearned* doxastic response to ME2. You might think that normal humans also have, as a part of their evidence base for B1, an unlearned connector C1 (where what this amounts to is that they have a sense that B1 is the appropriate belief to hold given the circumstances, i.e. their experiencing ME1). But we could simply stipulate that the possible cognizers in the Reidian example have in their evidence base for B1 a similar sort of *unlearned* connector (only it is C2 rather than C1). And we could add that, due to the same injury that causes them to form B1 in response to ME2, the brain-damaged humans also have an *unlearned* connector C2—a sense that B1 is an entirely appropriate belief to hold in their circumstances, i.e. their experiencing ME2. Furthermore, we could stipulate that in neither the Reidian case nor the brain-damaged human case is there anything else in the subjects' evidence base that is relevant to their holding B1.²⁶

What should we say of a species of cognizers for whom the natural unlearned response to grabbing a billiard ball is to experience ME2 and then form B1? It seems we should say that for such cognizers, B1 is a *fitting* unlearned response to ME2 and an *unfitting* unlearned response to ME1. This shows that the fittingness of an unlearned doxastic response is a contingent feature of it, a feature that depends in some cases on the species of the cognizer who has the response.

²⁴ Cf. Greco (1999: 277; 2000: 173–4), Markie (2004: 530–3), and Plantinga (1993b: 54–63).

²⁵ See n. 20.

²⁶ See n. 14.

Necessity is, therefore, false. Hence my recommendation that the evidentialist drop it.²⁷

1.2.4 Defending the Counterexample Against Intentionalism

In my objection to Necessity in the previous subsection, I began by noting that for normal humans, the belief B1 is a fitting unlearned doxastic response to ME1 and an unfitting unlearned doxastic response to ME2. Then I gave a Reidian example in which B1 is a fitting unlearned doxastic response to ME2 and an unfitting unlearned doxastic response to ME1. And I claimed that that example is a possible one. In short, I claimed that it is possible that B1 is a fitting unlearned doxastic response to ME1 and also that it is possible that B1 is a fitting unlearned doxastic response to ME2. Philosophers who think sensory experiences have propositional content would take this to suggest the following proposal having to do with the *contents of experience*:

CE: It's possible (and actual) for normal humans to have experience ME1 with the propositional content of B1; and it's possible for there to be cognizers like those in the Reidian counterexample who have experience ME2 with the propositional content of B1.

But do sensory experiences like ME1 and ME2 *have* propositional content? If 'experience ME1 has the same propositional content as B1' is just shorthand for something like 'ME1 is an experience to which its subject is naturally inclined to form B1 as an unlearned automatic doxastic response' then I'd agree with CE. But I'm not convinced that sensory experiences have propositional content.²⁸ I agree, of course, that they can spontaneously elicit beliefs (which *do* have propositional content) as doxastic responses. But I'm not convinced that that is the same thing as the sensory experiences themselves having propositional content. It seems to me that the view that an experience E has propositional content *p* is

²⁷ The arguments in this section establish the following point on which I relied in Chapter 3: (iv) *The appropriateness of an input is due to a contingent feature, not an essential feature, of that input.* Cf. n. 32 in Chapter 3.

Plantinga's argument (1993b: 54–63) that the *warrant*-conferring power of a belief ground is a contingent feature of it is similar to my objection to Necessity. The main difference (aside from the fact that he focuses on warrant and I focus on justification) is that his examples are less convincing than the Reidian one I employed in this subsection. One problem with Plantinga's examples is that many of them involve *learned* doxastic responses. This leads to the concern discussed in Section 1.2.2—a concern that the Reidian example, which focuses on *unlearned* doxastic responses, sidesteps. Another problem is that one of Plantinga's examples (1993b: 62–3) is met with understandable resistance by Bruce Russell (2001: 45–6) who seems to think that Plantinga's example involves treating a doxastic response that must be *learned* as if it can be an appropriate *unlearned* doxastic response. I think Plantinga's example can survive Russell's criticisms. But the Reidian example is better because it involves a belief that clearly *can* be an appropriate unlearned doxastic response.

²⁸ See Alston (2005b) for some objections to the view that they do.

grounded in nothing more than the observation that a person having E is naturally inclined to form a belief that *p* as an automatic response to it. However, for the sake of convenience, I will proceed as if in giving my Reidian counterexample to Necessity, I am proposing CE, since that will make it easier to consider an important objection from intentionalism.²⁹

Intentionalism is the view that the phenomenal character of experience is entirely determined by its propositional content. One consequence of intentionalism is:

CI: No two possible experiences that differ in phenomenal character have the same propositional content.

Given the obvious fact that ME1 and ME2 are experiences with different phenomenal character, CI entails that CE is false. Hence, intentionalism entails that my counterexample to Necessity is metaphysically impossible.

There is a significant literature debating the truth of intentionalism.³⁰ Obviously I can't do justice to it all here. What I will do instead is consider an impressive recent argument for CI and respond to it. This will go some way toward defending CE, on which my objection to Necessity depends. The argument for CI that I have in mind is proposed by Alex Byrne (2001). The argument speaks of an idealized person, which is a human-like person with powers of introspection, memory, and perception. An idealized person is one of whom the following three things are true: she is *competent* at accessing memories (i.e. when she tries to access memories she succeeds), she *tries* to detect whether there has been a change in the phenomenal character of her experience, and her memory is *perfectly accurate* (Byrne 2001: 208–9). Here then is the argument:³¹

Byrne's Argument for CI

1. Necessarily, if the phenomenal character of an idealized person's experience changes, that person notices a change in the phenomenal character of her experience.³²

²⁹ In conceding this for the sake of convenience, I'll insist that experiences, like sentences, have their propositional content contingently (though only in the latter case is this contingent upon human convention). However, I think of beliefs as having their propositional content essentially. Beliefs are individuated by their propositional content whereas experiences are individuated by their phenomenal character.

³⁰ Those in support of intentionalism (also called 'representationalism') include Byrne (2001), Dretske (1995, 1996), Harman (1990, 1996), Lycan (1996), and Tye (1995). Those with anti-intentionalist sympathies include Alston (2005b), Block (1990, 1994, 1996, 1998, 1999, 2003), Burge (1997), Chalmers (1996), Levine (1997, 2001), Peacocke (1983, 1984), and Robinson (1998).

³¹ The formulation is my own, based on Byrne (2001: 206–17).

³² Byrne notes that one might object that, at best, this is true only for non-negligible changes in phenomenal character (i.e. only for changes in phenomenal character the subject can reliably discriminate). To that Byrne (2001: n. 19) replies that non-negligible changes in phenomenal character can be viewed (on a perfectly reasonable precisified sense of 'phenomenal character') as the

2. Necessarily, if a person notices a change in the phenomenal character of her experience, then the way things seem to her changes—i.e. the propositional content of her experience changes.
3. Therefore, necessarily, if the phenomenal character of an idealized person's experience changes, the propositional content of her experience changes. (From 1 and 2)
4. Necessarily, if a change in the phenomenal character of an idealized person's experience entails a change in the content of her experience, then a change in the phenomenal character of a nonidealized person's experience entails a change in the content of her experience.³³
5. Therefore, necessarily, if the phenomenal character of a person's experience changes (whether the person is idealized or not), the propositional content of her experience changes—in other words, same experiential content entails same phenomenal character. (From 3 and 4)
6. If it's possible for two people to have experiences with different phenomenal character but the same propositional content, then it's also possible for a single person to have changes in the phenomenal character of her experience without any change in the propositional content of her experience.³⁴
7. Therefore, it's not possible for two people to have experiences with different phenomenal character but the same propositional content. (From 5 and 6)
8. If it's not possible for two people to have experiences with different phenomenal character but the same propositional content, then no two possible experiences that differ in phenomenal character have the same propositional content.³⁵

only changes there are in phenomenal character. Whether this is a satisfactory response is something I won't take up here.

³³ Here is Byrne's defense of 4 (2001: 214–16). Suppose a nonidealized person has experiences *e* and *e** which differ in phenomenal character. It seems possible for an idealized person also to have experiences *e* and *e** with exactly the same propositional content and phenomenal character as, respectively, *e* and *e** had by the nonidealized person. Now suppose that 3 is true (i.e. suppose that if two experiences had by an idealized person differ in phenomenal character, then they differ in content). Then it follows that *e* and *e** had by the idealized person differ in propositional content. And since the nonidealized person has the same experiences with same phenomenal character and propositional content, her experiences *e* and *e** differ in propositional content too.

³⁴ Here is Byrne's defense of 6 (2001: 216–17). Suppose one person has an experience with phenomenal character *P* and propositional content *p* and another person has an experience with a different phenomenal character *Q* and propositional content *p*. Given some sort of recombination principle for determining what is possible, it seems to follow that it's also possible for a third person to have both experiences consecutively: first the one with character *P* and content *p* and then the other with character *Q* and content *p*.

³⁵ This premise is not made explicit or defended by Byrne. But he clearly relies on it. For (Byrne 2001: 216–17) argues for the conclusion that 'if two subjects enjoy *e* and *e**, respectively, and if the contents of *e* and *e** are the same, then so are their phenomenal characters'. This is equivalent to 7

9. No two possible experiences that differ in phenomenal character have the same propositional content. (From 7 and 8)

The conclusion of this argument would, by my lights, be very surprising if true. What it says is impossible seems perfectly possible to me—namely, that two possible experiences (perhaps in different possible worlds) that differ in phenomenal character can each naturally incline their subjects to form, as automatic responses, beliefs with the same content. Since the reasoning is valid, this makes me suspect that at least one of the premises is false. In my response to the argument, I'll be focusing my attention on premise 2 (which is not to concede that the other premises are true).

The first thing to notice is that premise 2 can be read in two ways:

- 2*. Necessarily, if a person notices a change in the phenomenal character of her experience, then there is a change in the way things other than the phenomenal character of her experience seem to her.
- 2**. Necessarily, if a person notices a change in the phenomenal character of her experience, then *either* there is a change in the way the phenomenal character of her experience seems to her *or* there is a change in the way things other than the phenomenal character of her experience seem to her.

2** is tautological (and, therefore, highly plausible) but not of any value for Byrne's argument. 2* is the version Byrne needs for his argument but it isn't nearly as plausible, though I'm sure Byrne thinks it is plausible enough. The reason I say 2** isn't of any value for Byrne's argument is that the contents of experience he has in mind are not propositions about the phenomenal character of the subject's experience but instead propositions about something else such as the subject's environment (Byrne 2001: 201–3). A *perceptual* experience (which will be our focus) represents the subject's *environment* as being a certain way; *that* is its content. It doesn't represent *itself* as having a certain phenomenal character. Henceforth, 2 will be understood as 2*.

My objection to 2* is that it is plausible *only if* it is restricted to a certain kind of perceptual experience. It isn't true in the fully general way it is stated.

above. After defending this conclusion, he then goes on to 'restate' it on (2001: 217) as Conclusion C++: 'For any two possible experiences *e* and *e**, if they differ in phenomenal character, then they differ in content'. This is equivalent to 9 above. It appears, therefore, that he must be relying on 8. But consider the contrapositive of 8 which can be stated as 8*: *If (a) it's possible for a subject to enjoy e and it's also possible for a subject to enjoy e* (where e and e* differ in phenomenal character and yet have the same propositional content), then (b) it's possible for there to be two subjects, one of whom enjoys e and the other of whom enjoys e* (where e and e* differ in phenomenal character and yet have the same propositional content).* 8* is of the form: If (a) A is possible and B is possible then (b) A&B is possible. And of course not all claims of that form are true (e.g. if A is *p* and B is $\sim p$). So it seems we would need some special reason to accept 8* (which is equivalent to 8). But Byrne doesn't give us any such reason. Indeed, he never even acknowledges that he is relying on 8. Whether this will ultimately cause a problem for Byrne's argument is something I won't consider here.

This is important because it shows that Byrne's argument is unsuccessful if its conclusion, CI, is not likewise restricted to the same sort of perceptual experience. Moreover, my Reidian counterexample to Necessity depends on the falsity of the *unrestricted* version of CI, not on the falsity of the restricted version. Consequently, my objection to Necessity is not harmed by Byrne's argument for CI and intentionalism.

I need to defend my claim that 2* is plausible only when restricted to perceptual experience of a certain kind. The kind of experience I have in mind is sensory experience which prompts belief in *secondary* (rather than primary) qualities of objects in our environment. Reid ([1785] 2002: 200–5) describes the difference between primary qualities (such as extension, divisibility, figure, motion, hardness, softness, solidity, and fluidity) and secondary qualities (such as sound, color, taste, smell, heat, and cold) as follows. Our sensory experience gives us a direct and distinct notion of primary qualities:

Every man capable of reflection may easily satisfy himself, that he has a perfectly clear and distinct notion of extension, divisibility, figure, and motion. The solidity of a body means no more, but that it excludes other bodies from occupying the same place at the same time. Hardness, softness, and fluidity, are different degrees of cohesion in the parts of a body. It is fluid, when it has no sensible cohesion; and hard when it is strong. Of the cause of this cohesion we are ignorant, but the thing itself we understand perfectly, being immediately informed of it by the sense of touch. It is evident, therefore, that of the primary qualities we have a clear and distinct notion; we know what they are, though we may be ignorant of their causes. (Reid [1785] 2002: 201)

By contrast, our sensory experience gives us only a relative and obscure notion of secondary qualities:

It is otherwise with secondary qualities. If you ask me, what is that quality or modification in a rose which I call its smell, I am at a loss to answer directly. Upon reflection I find, that I have a distinct notion of the sensation which it produces in my mind. But there can be nothing like to this sensation in the rose, because it is insentient. The quality in the rose is something which occasions the sensation in me; but what that something is, I know not. My senses give me no information upon this point. The only notion therefore my senses give is this, that smell in the rose is an unknown quality or modification, which is the cause or occasion of a sensation which I know well. The relation which this unknown quality bears to the sensation with which nature has connected it, is all I learn from the sense of smelling; but this is evidently a relative notion. The same reasoning will apply to every secondary quality. (Reid [1785] 2002: 202)³⁶

³⁶ Although Reid's account of the difference between primary and secondary qualities is similar to Locke's in ways, the two accounts are importantly different. One important difference is that Locke believed that our ideas of primary qualities *resemble* the qualities which cause them whereas our ideas of secondary qualities do not resemble the qualities which cause them. Reid, by contrast, was impressed by Berkeley's point that no idea or mental state can resemble a nonmental cause. From this Reid concluded that neither sensations caused by primary qualities of objects nor sensations caused by secondary qualities of objects resemble the qualities that cause them. See Reid ([1785] 2002: 207–10).

Thus, my sense of touch gives me a clear and direct (nonrelative) notion of the hardness of a billiard ball in my hand. (Of course, it doesn't provide a scientifically informed notion of what hardness is. But it does provide a grasp of hardness—i.e. the strong cohesion of the object's parts—that makes no reference to my sensory experience.) However, my sight gives me only an obscure and relative notion of a thing's color, one that *does* make reference to my sensory experience: a color is that (pre-scientifically unknown) quality in a thing which causes a certain color sensation in me. Notice that this distinction is an *epistemological* distinction concerning our untutored pre-scientific grasp of these properties, not a *metaphysical* distinction concerning a difference in their intrinsic nature.³⁷

In elaborating upon this distinction, Reid points out that the sensations prompting our beliefs in secondary qualities of things are naturally and easily made objects of our attention while those prompting our beliefs in primary qualities of things are not:

The first [secondary qualities] are not only signs of the object perceived, but they bear a capital part in the notion we form of it. We conceive it [the secondary quality] only as that which occasions such a sensation, and therefore cannot reflect upon it without thinking of the sensation which it occasions. We have no other mark whereby to distinguish it. . . . But having a clear and distinct conception of primary qualities, we have no need when we think of them to recall their sensations. When a primary quality is perceived, the sensation immediately leads our thought to the quality signified by it, and is itself forgotten. We have no occasion afterward to reflect upon it; and so we come to be as little acquainted with it, as if we had never felt it. . . . They [the tactile sensations producing beliefs in the primary qualities of objects] carry the thought to the external object, and immediately disappear and are forgotten. Nature intended them only as signs; and when they have served that purpose, they vanish. (Reid [1785] 2002: 204–5)³⁸

Reid's point here is that we have a difficult time attending to the tactile sensations that prompt our beliefs in primary qualities of objects and that we rarely make those sensations the object of our attention. This supports his conclusion that our notion of primary qualities, unlike our notion of secondary qualities, is not a relative notion concerning the sensation that quality produces in us.

³⁷ This should go some way toward tempering objections (to drawing a primary/secondary quality distinction) from those who think there is no *metaphysical* difference between primary and secondary qualities. Jennifer McKittrick argues that Reid's view of the metaphysics of secondary qualities of objects is that they are the 'causal basis of a disposition to produce a sensation' in us (McKittrick 2002: 488). And she notes in the same passage that Reid's view of the metaphysics of secondary qualities is similar to Frank Jackson's even though Jackson (2000) calls his view 'the primary quality view of color'. This helps us to see that Jackson's insistence that color is *metaphysically* similar to primary qualities doesn't count against Reid's point that our pre-scientific grasp of color (and other secondary qualities) is different from our pre-scientific grasp of the qualities typically identified as primary qualities.

³⁸ See also Reid ([1764] 1997: 55–8) for a similar discussion.

Let's consider now the implications of this for 2*. In particular, let's consider how 2* applies to experiences which have as their content propositions attributing secondary qualities such as colors to things. Given the truth of what Reid has said about our pre-scientific grasp of secondary qualities, a proposition saying of a thing in my environment that it is red says of that thing that it has a property which causes a red sensation in me.³⁹ Now consider a case of color-spectrum inversion which goes undetected because both subjects use the same words to refer to the same objects. For example, pointing to the same tomato, they both say of it that it is red, even though one person bases her judgment on what we would call a 'red sensation' whereas the other bases her judgment on what we would call a 'green sensation'. Do the experiences and judgments of these two people have the same propositional content? It seems not. The content of the experience and judgment of one person says the tomato has a quality which causes *this* sort of sensation (pointing inwardly to what we'd call a 'red sensation') while the content of the experience and judgment of the other person says the tomato has a quality which causes *that* sort of sensation (pointing inwardly to what we'd call a 'green sensation'). If we didn't know their color spectra were inverted relative to each other, we wouldn't know that their judgments and experiences had different contents, since we couldn't know which experiences they were inwardly pointing to (and so couldn't see that they were different). Even so, the difference in propositional content remains. The difference remains because their color spectra are inverted and the phenomenal character of their perceptual experiences plays a role (in the way Reid says) in their attribution of secondary qualities to things. Hence, when we consider 2* restricted to experiences with contents attributing secondary qualities, it seems to give the right result: if we notice a change in phenomenal character, then the content of the experience changes too.

But this doesn't show that the unrestricted version of 2*, on which Byrne's argument depends, is true. The considerations in the previous paragraph are silent concerning perceptual experiences with contents attributing *primary* qualities to things in our environments. Moreover, Byrne's discussion in defense of premise 2* (what he calls 'Premise B') focuses solely on experiences with contents attributing *color* to things (Byrne 2001: 210–11). My suspicion is that the reason Byrne is taken in by premise 2* is that his focus is narrowed in this way to such secondary-quality experiences. In order to test the truth of 2* in its fully general

³⁹ But is what Reid says about secondary qualities—concerning our grasp of them rather than concerning their intrinsic nature—true? I must say that I find it exceedingly plausible given how strongly supported it is by reflection on my own grasp of the qualities in question. For some excellent discussions of Reid's account of primary and secondary qualities—discussions which include defenses of Reid's view against criticisms as well as comparisons highlighting its advantages over other accounts—see Nichols (2003) and McKittrick (2002). See also Smith (1990) for a good survey of views on primary and secondary qualities and a defense (in Section VI) of a position similar to Reid's.

form, we need to consider its plausibility when applied to perceptual experiences with contents attributing *primary* qualities to things.

In order to facilitate our discussion, let us consider three different kinds of creature (two of which aren't possible according to intentionalism). Each of the creatures I have in mind is a perceptual cognizer resembling us in important ways: each is similar to us in physical appearance, each makes perceptual judgments based on sensory experience, and each is able to reason and form beliefs and draw conclusions. Now suppose that unlike us, these three creatures have no vision, no sense of taste, and no sense of smell. (Let's suppose they differ from us in outward appearance only by having no eyes or noses.) The first kind of creature has perceptual abilities of hearing and touch that operate in the same way ours do. Hence, the sensory experiences one of them has when hearing and touching have the same phenomenal character as our own when we hear and touch.

The second kind of creature makes the same perceptual judgments as the first kind of creature, when in the same circumstances. In particular, when a creature of this second kind makes judgments based on *touch*, they're the same (when in the same circumstances) as those of the first creatures. The only difference is that the phenomenal character of the sensory experience had by this second type of creature when she physically touches things is not at all like our tactile experience. Instead, it is like our olfactory experiences. (In this way, these creatures are like the cognizers in the Reidian example given in the previous subsection.)⁴⁰

The third kind of creature is the one that is most interesting for our purposes here. A creature of this kind also makes the same judgments based on touch (when in the same circumstances) as those made by creatures of the first kind and the second kind. Sometimes when grabbing a billiard ball, she experiences ME1 just as the first kind of creature does. On such occasions, her natural, unlearned inclination is to form a belief like B1 (again, just as it is for the first kind of creature). We may conclude from this that her experience, ME1, has the same propositional content as B1. However, on other occasions, when grabbing a billiard ball, she experiences ME2 (what we would call an 'olfactory experience'), just as the second kind of creature does. On such occasions, the natural, unlearned inclination of this third kind of creature is the same as it is for creatures of the second kind: it's to form a belief like B1. We may conclude from this that her experience, ME2, has the same propositional content as B1. What determines

⁴⁰ Notice, by the way, that for this second kind of creature, a sensory experience with the same phenomenal character as our olfactory sensations prompts a belief in *primary* qualities of objects whereas that type of sensory experience in us prompts a belief in *secondary* qualities of objects. Above I quoted and discussed a passage from Reid ([1785] 2002: 204–5) in which he notes that we rarely attend to sensory experiences that prompt beliefs in primary qualities of objects whereas it's the norm for us to focus our attention on sensory experiences that prompt beliefs in secondary qualities of objects. In light of this, we can assume that, although we make our olfactory sensations the focus of our attention, this second kind of creature rarely attends to its sensory experiences that have the same phenomenal character as our olfactory sensations.

whether grabbing the billiard ball results in her having ME1 or ME2? Let's say it's due to some indeterministic physical process that occurs when she grabs a billiard ball, with the result that about half the time she experiences ME1 (when she grabs a billiard ball) and half the time she experiences ME2. Let's also add that all of her activities of touching things have similar indeterministic results: half the time they result in what we'd call 'tactile experiences' and the other half they result in experiences we'd call 'olfactory experiences' (though the same physical stimulation is always correlated with one or the other member of the same pair of sensory experiences).

Now, the key question is whether it's possible for both ME1 and ME2 to have the same content for a single creature. It seems to me that it is possible. It seems clearly possible for the first kind of creature to form B1 as a natural and unlearned response to ME1. It also seems possible for the second kind of creature to form B1 as a natural and unlearned response to ME2. This leads me to think it's also possible for the third kind of creature to form B1 as a natural and unlearned response to ME1 and also as a natural and unlearned response to ME2. We can add that the creature notices the difference in phenomenal character of ME1 and ME2. But though she notices this, her natural unlearned inclination to form judgments about her environment is the same in each case. Admittedly, this would seem odd to you and me (to have a change in phenomenal character of our experience without a change in the judgments we're naturally inclined to form about our environments). But it doesn't seem odd in the least to her since this is how it has always been for her.

Now suppose that while grabbing a billiard ball, her experience changes from ME1 at t_1 to ME2 at t_2 . And suppose we ask her if her environment seems different at the two times (we can stipulate that no other experiential changes have taken place in that interlude). Then she will say that her environment seems exactly the same to her at t_2 as it did at t_1 . She might add that the phenomenal character of her experience has changed; but even so, she'll insist that the world as represented to her by her experience didn't seem to change at all. Again, this will strike us as odd because this sort of thing doesn't happen to us.⁴¹ But it won't seem odd to her. Nor does it seem to be impossible.

Notice that because B1 attributes primary qualities to things in her environment, there isn't the problem noted earlier with secondary qualities since, in the case of primary quality attribution, there's no reference to the phenomenal character of the subject's experience. So we haven't got *that* reason to think it impossible for the content of ME1 and ME2 to be the same. Do we have any other reason? I don't see that we do. Byrne, in defending premise 2*, says 'the subject can only discover the phenomenal character of her experience by attending to the world . . . as her experience represents it' (2001: 211). The idea is that we couldn't

⁴¹ I assume that this oddness plays a role in Byrne's conclusion that it is impossible. See his discussion in Byrne (2001: 206–7).

find out that the phenomenal character of our experience had changed if what it represented didn't change. But this seems flatly false, especially in light of the example we're considering. If my natural and unlearned inclination in response to both ME1 and ME2 is to form B1, then it seems that both of those sensory experiences have the same content—it seems that both represent the world in the very same way. But this doesn't prevent me from noticing that ME1 and ME2 differ in their phenomenal character.

I see no reason to think it's impossible for there to be a creature of this third kind who operates in the way described above. And if it is possible, then premise 2* from Byrne's argument is false. I conclude, therefore, that Byrne's argument for CI fails and that it presents us with no good reason to think my Reidian counterexample to Necessity fails.

1.3 The Second Improvement: Add Proper Function

Having rejected Necessity, we must return to the question that prompted our discussion of it, namely: 'If objective fittingness doesn't depend on reliability, what does it depend on?' The counterexample to Necessity considered in Section 1.2.3 suggests that the fittingness of a doxastic response depends, in some cases at least, on the species of the cognizer who has it. What is it about the species of a cognizer that determines such fittingness in those cases? An answer that immediately suggests itself is that what makes a belief a fitting unlearned doxastic response to an experience has to do with the way the cognitive faculties of the cognizer in question are *supposed to* function. For clearly that is something that can vary from species to species. *Our* cognitive faculties are supposed to function so that when we experience tactile sensation ME1, our unlearned doxastic response is B1. Not so the cognizers described in the Section 1.2.3. *Their* faculties are, we may assume, supposed to function so that when those cognizers experience 'olfactory' sensation ME2, their unlearned doxastic response is B1. The sense in which our faculties and theirs are *supposed to* function in the ways just specified is the same as the sense in which our hearts are supposed to function so that they beat less than 200 times a minute when we are at rest. And the 'supposed to' of heart function is clearly connected with the notion of proper function or healthy function. This suggests that the fittingness of a doxastic response to evidence is contingent upon the *proper function* of the cognitive faculties of the person in question. And this, in turn, suggests that the evidentialist claim E_F could be illuminatingly improved if it were changed to say something like the following (where a *PF-induced* doxastic response is one produced by the proper functioning of the subject's cognitive faculties):

E_{PF} : S's belief B is justified iff B is a *PF-induced* doxastic response to S's evidence.

That we replace E_F with E_{PF} is my second recommendation. It amounts to the suggestion that epistemic fittingness should be understood at least partly in terms of proper or healthy cognitive functioning.

1.4 The Third Improvement: Drop Mentalist and Internalist Implications

E_{PF} says that a belief is justified only if it is a PF-induced response to S 's *evidence*. As I noted earlier, this could suggest that some form of mentalism or internalism is true. For if evidence is something to which I have special epistemic access, then the proponent of E_{PF} will answer the question:

(A) What sorts of things can be appropriate inputs to our belief-forming processes?

by saying that only states that are accessible on reflection alone can be appropriate inputs. But to say this (while asserting E_{PF}) is to impose an awareness requirement on justification. And I've argued at length in Part I of this book against that position. If, on the other hand, evidence must be a mental state of the subject, then the proponent of E_{PF} will answer question (A) by saying that only mental states can be appropriate inputs. But there seems to be no good reason for that answer to question (A). Moreover, as I noted in Chapter 3, there are examples of possible scenarios in which beliefs produced without any mental states in their causal ancestry seem to be justified beliefs.⁴² These examples give us a reason to *reject* the mentalist answer to (A).

In order, therefore, not to give any hint of endorsing a form of internalism or mentalism, E_{PF} should be replaced with:

I_{PF} : S 's belief B is justified iff B is a PF-induced doxastic response to the *input* to S 's belief-forming systems.

This is my third recommendation. The only restrictions I_{PF} places on what counts as an appropriate input to one's belief-forming processes are (a) that the input is something to which the belief is a (causal) response and (b) that the input is something to which the subject is *supposed to* have the doxastic response in question (where the 'supposed to' has to do with proper function).⁴³

We began with the admittedly attractive E_F understood in accord with Nonreliability, Objectivity, and Necessity. Call that package 'the evidentialist position'.

⁴² See Section 1.3.4 of Chapter 3.

⁴³ In Bergmann (2004b: 43) I said that what count as candidates for appropriate inputs are all things that Pollock and Cruz call 'internal states' (see Section 2.1.1 of Chapter 3 where I explain what Pollock and Cruz say about internal states). I am now less confident that I know enough about what Pollock and Cruz have in mind by 'internal states' so I've decided not to use their term to explain my position.

After my recommended improvements, we now have I_{PF} understood in accord with Nonreliability, Objectivity, and the denial of Necessity. Call that package 'the proper function position'. Clearly the proper function position is significantly different from the evidentialist position. But despite this, it retains the attractiveness of the evidentialist position in addition to having the advantages of the three recommended improvements discussed above. The reason it retains the initial attractiveness of the evidentialist position is that things counting as evidence are perhaps the most prominent sort of input to *our* belief-forming systems. Consequently, there will be a wide range of *standard* cases in which the different implications of the evidentialist position and the proper function position are not noticeable. The result is that the *prima facie* plausibility of the evidentialist position, which arises from a consideration of these standard cases, carries over to the proper function position.

2 A PROPER FUNCTION ACCOUNT OF JUSTIFICATION

In Section 1 I argued that the initially appealing evidentialist position is made more plausible if it is improved to become the proper function position which denies Necessity and replaces E_F with I_{PF} . In this section I will lay out and defend an analysis of justification that can be viewed as a more careful statement of the idea behind I_{PF} .

2.1 The Proposed Analysis

Plantinga has proposed the following analysis of *warrant* (i.e. that which makes the difference between knowledge and mere true belief):

W_{PF} : S's belief B is warranted iff each of the following conditions is satisfied:

- (i) the cognitive faculties producing B are *functioning properly*
- (ii) the cognitive environment in which B is produced is sufficiently similar to the one for which S's faculties were 'designed'⁴⁴
- (iii) the modules of the 'design' plan governing the production of B are directly aimed at the production of *true* beliefs
- (iv) there is a high objective probability that a belief formed in accord with those modules in that sort of cognitive environment is true.⁴⁵

⁴⁴ The quotation marks are to indicate that the design in question needn't involve a literal designer. For discussion of how to make sense of design talk without postulating a literal designer, see Section 3.1 below.

⁴⁵ This is *not* a direct quotation; for his exact wording, see Plantinga (1993a: 19, 46–7). Plantinga insists that this is still only a first approximation. For further details, qualifications, and emendations see Plantinga (1993a, 1996, 1997, 2000: 156–61). In particular, see Plantinga (1996,

He adds condition (iv) because he recognizes that the first three conditions don't entail that the belief was reliably formed and he thinks warrant should entail reliability.

A similar line of reasoning leads me to think he should add a fifth condition requiring the absence of a defeater:

(v) S does not take B to be defeated.

The next chapter will be devoted to discussing defeaters in general and this condition in particular. For now I'll say only that just as it seems that warrant should entail reliability and that proper function conditions don't require reliability, so also it seems that warrant should entail the absence of a defeater and that proper function conditions don't do that. There could be cognizers designed by a literal creator to form beliefs in a reliable way but also to take each of their reliably formed beliefs to be defeated. Furthermore, they could have been designed so that although they took their beliefs to be defeated, they ignored this fact and continued to hold them. Such beliefs would *not* be warranted even if they satisfied Plantinga's four conditions.⁴⁶

My proposed account of justification is just my modified version of Plantinga's account of warrant (i.e. his account together with the fifth condition I just mentioned) *without* Plantinga's condition (ii)—the environmental condition. Once that condition is dropped, his account ceases to be reliability-entailing. Thus, to a first approximation,⁴⁷ we may say that:

J_{PF}: S's belief B is justified iff (i) S does not take B to be defeated and (ii) the cognitive faculties producing B are (a) functioning properly, (b) truth-aimed and (c) reliable in the environments for which they were 'designed'.⁴⁸

To see how this analysis is a refinement of I_{PF}, it's important to notice that clause (ii) is to be understood so that it entails that S's doxastic response to the

2000) where he notes that, in addition to the conditions (i)–(iv) noted above, we need to add a fifth condition called the 'resolution condition'. The purpose of this resolution condition is to guarantee that warrant entails truth, thereby enabling it to handle Gettier problems by ruling out cases of warranted belief that is *accidentally* true. On the question of whether warrant entails truth, see Merricks (1995).

⁴⁶ Plantinga says that warrant depends on the proper functioning of our defeater systems (1993a: 41). So he is aware of the concern I'm raising. However, what he doesn't acknowledge is that just as his conditions (i)–(iii) don't entail the satisfaction of a reliability condition, so also his conditions (i)–(iv) don't *entail* the satisfaction of condition (v).

⁴⁷ A more careful account would also explain how justification comes in degrees. For an attempt to use a proper function analysis to explain how *warrant* comes in degrees, see Plantinga (1993a: 7–9). Something similar could be said, *mutatis mutandis*, using the J_{PF} account of *justification*.

⁴⁸ Clauses (ii)(b) and (ii)(c) of J_{PF} are shorthand for their correlates in Plantinga's account of warrant given above. Thus, to say that the faculties producing B are truth-aimed is to say that 'the modules of the design plan governing the production of B are directly aimed at the production of *true* beliefs'. And to say that the faculties producing B are reliable in the environment for which they were designed is to say that 'there is a high objective probability that a belief formed in accord with those modules (i.e. those modules of the design plan governing the production of B) in that sort

input to her belief-forming systems is due to the proper functioning of S's faculties. As Plantinga points out (1993a: 22–4), when we say cognitive faculties are functioning properly, the basic idea is that their functioning results in the cognitively healthy doxastic response to the circumstances in which they are operating (which will include, rather prominently, the input to the subject's belief-forming systems).

The motivation for clause (i) of J_{PF} was discussed briefly above and will be discussed at length in the next chapter. The motivation for clause (ii)(a) of J_{PF}—the proper function condition—was discussed in Section 1 where I argued against Necessity and in support of understanding fittingness in terms of proper function. But nothing said so far addresses the motivation for adding clauses (ii)(b) and (ii)(c). So why does J_{PF} require that the belief be produced by *truth-aimed* cognitive faculties that are *reliable* in the environments for which they were 'designed'? Why not require, in addition to the absence of any believed defeater, merely that the belief be produced by *properly functioning* cognitive faculties?

The answer here is very much like the explanation Plantinga (1993a: 11–20, 26–8) gives for including clauses like (ii)(b) and (ii)(c) in his account of *warrant*. The main difference is that when Plantinga adds such clauses to his account of warrant, they entail that the belief is produced by a cognitive faculty that produces mostly true beliefs whereas they do not have that consequence when added to my account of justification. The reason that the addition of such clauses doesn't make my account reliability-entailing is that I've dropped the 'right environment' clause included in Plantinga's account. J_{PF}, therefore, is not a *reliabilist* account of justification despite the fact that it includes a clause having to do with reliability. Because it isn't a reliabilist account, it is able to avoid the standard evil demon objection to reliabilist accounts of justification. According to that objection, the fact that victims of evil demons could have justified beliefs despite the fact that their beliefs are unreliably formed shows that reliabilist accounts of justification are mistaken to say that reliability is necessary for justification.⁴⁹

To return, then, to the question at hand, let me explain why I've added clauses (ii)(b) and (ii)(c) which have to do with the cognitive faculty in question being truth-aimed and reliable in the right environment. My inclusion of clause (ii)(b), requiring that the faculty be truth-aimed, is due to the following sort of counterexample to the suggestion that clauses (i) and (ii)(a) are sufficient for justification. Suppose that the cognitive faculties of some cognizer were created by a being

of cognitive environment is true'. See Plantinga (1993a: 22–4) for an account of how to think of design plans in terms of the circumstance–response pairs (or the circumstance–response–purpose triples) they dictate. And see Plantinga (1993a: 38–40) for a discussion of the difference between being directly aimed at truth and being indirectly aimed at truth.

⁴⁹ The remarks in this paragraph establish the following point on which the argument of Chapter 3 relied: (iii) *Unlike reliabilism, the proper function account of justification handles the intuition that a belief can be justified even though formed in an unreliable way.* Cf. n. 32 in Chapter 3.

interested in that cognizer's comfort. But suppose also that the particular faculty producing the belief we are considering is intended by its creator not to produce *true* beliefs but, rather, to produce beliefs that will minimize psychological trauma (even if that involves regularly producing false beliefs). Then it seems that beliefs being produced by such a cognitive faculty won't be *epistemically* fitting responses to the input to the subject's belief-forming system though they may be appropriate in some other sense. This would be a case of a belief that isn't justified (since it isn't an epistemically fitting doxastic response) even though it is produced by a properly functioning cognitive faculty.⁵⁰

Or suppose some evil demon intentionally designs creatures to form mostly false beliefs. For example, suppose that this malevolent creator decides to create Vic in such a way that his natural unlearned doxastic response to ME2 (the 'olfactory' experience) is to form the billiard ball belief B1. Vic's cognitive faculties are, therefore, designed to function in much the same way as are the faculties of the possible cognizers described in my Reidian counterexample to Necessity in Section 1.2.3. The difference is that the beliefs of those cognizers are mostly true if their cognitive faculties are functioning properly in the environment for which they were 'designed' whereas Vic's beliefs (including B1) are not. Now we noted earlier that the beliefs of those possible cognizers mentioned in 1.2.3—beliefs like B1—are (epistemically) fitting responses to their evidence. But consider Vic who has been designed (by the demon who is intent on having Vic form mostly false beliefs) to form the false billiard ball belief B1 in response to the 'olfactory' experience ME2. Is Vic's belief B1 an *epistemically* fitting response to ME2? It seems not. It's true that we can't blame Vic for holding the belief. And his epistemic blamelessness might give his belief a *subjective* sort of epistemic justification. But that won't be enough to make his belief *objectively* epistemically fitting. Nor will the fact that he was created by a designer who *intended* that he respond in this non-truth-conducive way to his experiences. To avoid counterexamples like those given in this paragraph and the preceding paragraph, it is important to add clause (ii)(b) to J_{PF} requiring that the belief be formed by a properly functioning faculty that is *aimed at truth*.

To see the importance of clause (ii)(c), consider a creature designed by one of Hume's infant deities. And suppose that, although this incompetent creator was trying to make a believer with reliable faculties, it instead created one whose faculties produce mostly false beliefs when placed in the environment in which it was intended by its creator to produce true beliefs. For example, suppose this infant deity intentionally created Ric in such a way that, like the cognizers described in Section 1.2.3, his natural unlearned doxastic response to ME2 is

⁵⁰ Overly optimistic beliefs about chances of recovery in response to evidence of one's poor health are often suggested as examples of beliefs that might be the result of proper function even though they aren't epistemically fitting doxastic responses to the available evidence. Such beliefs might aid recovery, perhaps by reducing stress. They would, therefore, serve a purpose even if they aren't aimed at truth.

to form B1. But, contrary to what this bumbling creator had hoped, this belief (like most of the other beliefs Ric forms) is false when produced in the intended environment by faculties functioning as they were designed to function. Is Ric's belief B1 an epistemically fitting response to ME2? It seems not. And the same would be true if the infant deity had designed a more human-like creature to form B1 in response to ME1 rather than ME2, intending (but utterly failing to bring it about) that such a belief is true when formed in that way in that creature's intended environment. Despite the failed design attempt, the belief in question may be, in some sense, the output of properly functioning cognitive faculties (since those faculties are, in some sense, operating in the way they were designed to operate).⁵¹ But because the plan, when implemented as intended, didn't work, the belief doesn't seem to be an epistemically fitting response to the subject's evidence.

What the considerations in the previous paragraph suggest is that beliefs produced by properly functioning truth-aimed cognitive faculties do not result in epistemically fitting doxastic responses to one's evidence *if* the faculties in question aren't *successfully* aimed at truth—i.e. if they aren't likely to produce true beliefs when operating in the environment for which they were 'designed'. Hence, clause (ii)(c) of J_{PF}.

I should comment on the fact that what I've been saying in the last few paragraphs about demon cases might seem to conflict with my claim that evil demon counterexamples to reliabilism are plausible. According to those counterexamples, a belief can be justified, even objectively justified, in cases where the belief isn't reliably formed but is instead the result of misleading experiences produced by a deceptive demon. Yet in the last few paragraphs I point out the implausibility of the suggestion that the beliefs of a creature created by an evil demon to have unreliable beliefs are objectively justified. The explanation for this apparent conflict is that although I agree that in *some* evil demon cases (in which the subject's beliefs aren't reliably formed) the subject's beliefs *are* objectively justified, I don't think this is true in *all* such cases. What is true in most demon cases is that the subject's beliefs are subjectively justified, in the sense that the subject is epistemically blameless in holding them (since it is beyond her control to do otherwise). Furthermore, it is also true that we are sometimes quite sure that the demon victim is (epistemically) *supposed to* hold a certain perceptual belief in response to some sensory experience (as, for example, when an ordinary *human*, part way through her life, comes under the influence of a deceptive demon). If we *are* sure about such 'supposed to' claims, then, with respect to cases in which such a demon victim holds the beliefs we think she is supposed to hold (given the experiences she is having), we may conclude that those unreliably formed beliefs are objectively fitting doxastic responses and,

⁵¹ See Plantinga (1993a: 26–8) for a related discussion of two senses of working in accord with a design plan.

hence, justified, despite the fact that the experiences in question are illusory ones produced by the demon. However, when the demon victim is designed and created by the demon specifically to hold unreliably formed beliefs, things are different. For it is no longer clear that her beliefs are *epistemically* fitting responses to her evidence, even if there is some sense in which they are doxastic responses she is *supposed to* have.⁵²

2.2 The Virtues of the Analysis

Now that we have J_{PF} before us and understand some of the motivation for it, let's consider two of its virtues.

2.2.1 Handling of Cases

One virtue of J_{PF} is that it handles certain examples better than reliabilism or Necessity-endorsing internalism handles them. Some of these examples involve the reliable cognizers described in Section 1.2.3 whose natural unlearned response to ME2 is B1 and whose natural reaction to grabbing a billiard ball is to experience ME2. I'll call the complete description of the natural way of functioning for those alien cognizers 'design plan B'. And I'll call the complete description of our natural way of functioning 'the human design plan'.⁵³

Consider the following the six cases (in each case, the doxastic response is unlearned):

Case I: A human forms B1 in response to the tactile sensation ME1. B1 is a reliably formed belief produced by properly functioning faculties in an appropriate environment.

Case II: A human forms B1 in response to ME1. However, since this human is a demon victim whose perceptual beliefs are all artificially produced, ME1 is not produced by actual contact with a billiard ball (nothing like such contact occurs for this person). The result is that B1 is not reliably formed.

Case III: Due to cognitive malfunction caused by a radiation overdose, a human comes to have the tendency to form B1 in response to ME2. The same overdose also prevents her from ever noticing anything wrong with forming this belief in this way. Since ME2 isn't a reliable indicator of the truth of B1, the result is that B1 is not reliably formed.

⁵² What about a case where an incompetent demon sets out to create a cognizer that holds unreliably formed beliefs but bumbles things so badly that the result is a duplicate of an ordinary human in a standard human environment? For a discussion of examples like this one, see section 3.4 at the end of this chapter.

⁵³ As Plantinga points out (1993a: 22–4), we can think of 'design' plans as sets of circumstance–response pairs (or circumstance–response–purpose triples).

Case IV: A nonhuman cognizer with ‘design’ plan B forms B1 in response to ‘olfactory’ sensation ME2. This belief is a reliably formed belief produced by that cognizer’s properly functioning faculties in an appropriate environment for the operation of those faculties.

Case V: A nonhuman cognizer with ‘design’ plan B forms B1 in response to ME2. This belief is a properly functioning unlearned doxastic response for such a creature to ME2. However, since this cognizer is a demon victim whose perceptual beliefs are all artificially produced, ME2 is not produced by actual contact with a billiard ball (nothing like such contact occurs for this cognizer). The result is that B1 is not reliably formed.

Case VI: Due to cognitive malfunction caused by a radiation overdose, a non-human cognizer with ‘design’ plan B comes to have the tendency to form B1 in response to ‘tactile’ sensation ME1. The same overdose also prevents her from ever noticing anything wrong with forming this belief in this way. Since in this environment (which is an appropriate one for this cognizer) ME1 isn’t a reliable indicator of the truth of B1, the result is that B1 is not reliably formed.

In both case IV and case V, B1 seems to be justified. For in those cases, just as in cases I and II, B1 is a properly functioning and fitting unlearned response to the main evidence despite the fact that, in case V, the belief is formed in an unreliable way. And since case VI is like case III in that B1 is a malfunctioning and *unfitting* unlearned response to the subject’s main evidence, B1 seems to be as unjustified in case VI as it is in case III.

We can summarize all six cases as follows:⁵⁴

Case	Which ‘Design’ Plan?	Which Functioning Properly?	Which Environment?	Which Move From Ground to Belief?	Reliably Formed?	Justified?
I	human	yes	appropriate	ME1→B1	yes	yes
II	human	yes	demon	ME1→B1	no	yes
III	human	no	appropriate	ME2→B1	no	no
IV	plan B	yes	appropriate	ME2→B1	yes	yes
V	plan B	yes	demon	ME2→B1	no	yes
VI	plan B	no	appropriate	ME1→B1	no	no

⁵⁴ The six cases can also be described so that the belief ground in each case includes an *unlearned* connector (see Sections 1.2.1 and 1.2.2 of this chapter for an account of what an unlearned connector is). Just replace all appearances of ‘ME1’ (in the above descriptions of the six cases as well as in the table below) with ‘ME1 + C1’. Likewise, replace all appearances of ‘ME2’ with ‘ME2 + C2’. In the demon cases, the unlearned connector in question is caused by the demon; in the malfunction cases, the unlearned connector is caused by the radiation overdose.

Reliabilist accounts get cases II and V wrong. They say the beliefs should be unjustified because they are unreliably formed. Internalist accounts endorsing Necessity get cases IV, V, and VI wrong. For the belief ground in cases IV and V is the same as the belief ground in case III and yet, contrary to what such internalist accounts entail, B1 is *not* unjustified in cases IV and V even though it *is* unjustified in case III. Likewise, despite the fact that the belief ground in case VI is the same as the belief ground in cases I and II, B1 is justified in cases I and II but not in case VI. By contrast, J_{PF} gets all six cases right.⁵⁵

I noted in Chapter 3 that Feldman and Conee defend their evidentialist position by describing six pairs of examples their position handles accurately.⁵⁶

Example Pair 1: Bob and Ray both believe that it will be warm today. Ray bases his belief on reading yesterday's newspaper in an air-conditioned hotel lobby while Bob bases his belief on his doing that same thing together with his walking outside and feeling the heat. Bob's belief is more justified than Ray's.

Example Pair 2: Consider these two stories about Ray getting further information. Both begin with Bob, who is a pillar of integrity, coming in from the heat and telling Ray that it felt hot outside. The first story says that Ray thinks it just as likely that Bob is deceiving him as that he is telling the truth. The second story says that Ray has excellent reasons to think Bob is a pillar of integrity. Bob's belief about it being warm is more justified in the second story than in the first.

Example Pair 3: A novice bird watcher and an expert are together looking at a woodpecker and have (we'll stipulate for simplicity) the same visual experience of it. Both believe it is a woodpecker. But the expert has fully reasonable beliefs about what woodpeckers look like and the novice doesn't (nor does he have any other evidence—such as the expert's testimony—for thinking this is a woodpecker). The expert's belief is justified and the novice's is not.

Example Pair 4: A logic TA and a beginning logic student both are looking at the same sentence—one they both know is true. They are also considering a second sentence and they both believe it too is true. The logic TA can easily tell that the second sentence is logically equivalent to the first (this is why he believes it too) but the beginning logic student cannot and has no other information pertaining to the truth or falsity of the second sentence (not even a reason to think the TA accepts the second sentence). The logic TA's belief that the second sentence is true is justified whereas the beginning logic student's belief in that same claim is not.

⁵⁵ Given how easy it is to multiply examples like cases IV, V, and VI, the remarks in this paragraph establish the following point on which the argument of Chapter 3 relied: (ii) *There are many examples which are better handled by the proper function account of justification than by their mentalist view.* Cf. n. 32 of Chapter 3.

⁵⁶ See Conee and Feldman (2001: 236–8). For the sake of brevity, the descriptions given here use my wording, not theirs.

Example Pair 5: Smith believes, on the basis of good evidence presented to him by his officemate Jones, that Jones owns a Ford and that someone in the office owns a Ford (the latter conjunct is deduced by Smith from the first conjunct). In fact, the evidence was deceptive: Jones doesn't own a Ford though, unbeknownst to Smith, someone else in the office does. Later Smith comes to find out that Jones was only pretending to own a Ford but he persists in believing (now without any good evidence) that Jones owns a Ford and that someone in the office owns a Ford (again, the latter conjunct is deduced from the first conjunct). Smith's belief before the discovery that Jones is pretending is justified and Smith's belief after the discovery is unjustified.

Example Pair 6: Consider these two stories about Hilary, an envatted human brain recently abducted from a fully embodied ordinary environment. In both stories, his life now seems normal and he is artificially given the false memory that he had oatmeal for breakfast yesterday. In the first story, the artificial memory is faint, lacking in detail, and it seems to him an unlikely breakfast for him to have had. In the second, the memory is vivid and it seems a typical breakfast for him to have had. The former belief is unjustified and the latter belief is justified.

The first five pairs of examples each include one member that is basically like case I above: it is a reliably formed belief produced in accord with proper function in response to the subject's main evidence and, at least in some cases, a connector.⁵⁷ The other member in each of these first five pairs of examples is either less justified or not justified at all. And in each case this is due either to the absence of some main evidence (example pair 1), the absence of a connector (example pairs 2–4), or the addition of a disconnector (example pair 5). There is no reason to think a proper function account of justification would have any difficulty handling any of these cases. In each case, the justificatory difference would be due, at least in part, to the fact that the doxastic response is no longer in accord with proper function when the main evidence or connector is dropped or a disconnector is added.⁵⁸ In the last of Conee and Feldman's six pairs of examples, the belief is not reliably formed because it is produced in a brain in a vat via artificially induced experiences. But it is a human brain that is still responding to experience as a properly functioning human brain does. The well-justified member of this pair is like case II above. The other member is not well-justified due to the fact that the main evidence is different and there is a disconnector present. Although a reliabilist account of justification might have difficulty with this last

⁵⁷ See n. 54 above where I explain how my six cases can be read to include connectors.

⁵⁸ In Conee and Feldman's first two examples, the problematic member of the pair is still justified, but it is less justified. This is because the more strongly justified belief in each pair has two sources of justification and, in the case of the less justified belief, only one of these sources is missing (due to the absence of the main evidence or a connector). J_{PF} has no difficulty with appealing to the fact that the second source of justification is missing in order to explain the justification differences in these cases.

pair of examples (since it includes a case of an unreliably formed belief that seems justified), J_{PF} will have no trouble explaining the justificatory differences between the members of this pair.⁵⁹

J_{PF} differs from a reliabilist account of justification in at least two ways. First, unlike reliabilism, J_{PF} denies that reliability is *necessary* for justification. Because of that difference, J_{PF} handled cases II and V above better than reliabilism did. Second, unlike reliabilism, J_{PF} denies that reliability is *sufficient* for justification. It is for this reason that proponents of J_{PF} needn't be embarrassed by the examples of strange but reliable ways of forming beliefs that are so often employed against reliabilism. BonJour's examples of reliable clairvoyants which he uses against externalism won't work against J_{PF} unless it is stipulated that the cognizers in question form these clairvoyant beliefs when their faculties are functioning properly. But once that is stipulated, we have to admit that their clairvoyant beliefs are no more strange than our a priori or memory or perceptual beliefs. And then we lose the intuition that the beliefs in the example aren't justified. Similar remarks apply to Jonathan Vogel's (2000: 612 f.) example of a 'gas tank' belief. In this case, a person forms beliefs about how full the gas tank is by looking at the gauge. The gauge is a perfectly accurate indicator of how full the tank is but the person has no reason whatsoever to think so. It's difficult to take seriously the stipulation that the person has no reason at all to think the gauge is a reliable indicator of the amount of gas in the tank. But if we make an effort to take that stipulation seriously and to ponder the example as described, then we have a good reason to agree with Vogel that the beliefs so produced aren't justified. But this doesn't count against J_{PF} , even if it does count against reliabilism. For our assumption is that the person forming these beliefs is like us insofar as the proper functioning of her faculties does *not* dictate that she form beliefs about the fullness of the gas tank based (in an unlearned way) on what the gauge reads. Thus, J_{PF} handles these cases used to show reliability isn't sufficient for justification better than reliabilism handles them.

2.2.2 *Justification and Truth*

Another virtue of J_{PF} is that it provides an account of the justification—truth connection—one that doesn't lead to the usual troubles. Stewart Cohen (1984) has argued that although there seems to be a nontrivial connection between justification and truth, there seems to be no good account of that connection. He considers and, for good reasons, rejects two types of accounts: (i) reliabilist accounts, according to which there is an objective connection between justification and

⁵⁹ The remarks in this paragraph establish the following point on which the argument of Chapter 3 relied: (i) *The pairs of examples Conee and Feldman use in support of their mentalist view are handled easily by the proper function account of justification.* This, together with nn. 27, 49, and 55 in this chapter, completes the book-keeping that was promised in n. 32 of Chapter 3.

truth, and (ii) subjectivist accounts that say that what matters is that there is a justification—truth connection *from the subject's perspective*.⁶⁰

But there are several other accounts he doesn't consider—understandably since they hadn't been proposed at the time. One is J_{PF} . According to it, the connection between justification and truth comes through the notion of an appropriate environment: there is a high objective probability that a justified belief will be a true belief *if* the properly functioning faculties that produce it are operating in the environment *for which they were 'designed'*. This explains the connection between justification and truth while avoiding the evil demon counterexamples to reliabilism. It also explains our intuitions that the beliefs of the nonhuman alien cognizers with design plan B (discussed earlier in this chapter) are justified when formed in accord with proper function (and in a reliable way) in the environment for which they were 'designed', despite the fact that their way of forming beliefs is unreliable in our environment.

Another account is Sosa's virtue reliabilism, as it is developed in some of his earlier work. According to that early development of his position, 'justification is relative to environment' (Sosa 1991: 144). Thus, the demon victim's beliefs are justified relative to *our* environment because the virtuous dispositions producing her belief are reliable in our environment; but those same beliefs aren't justified relative to the victim's environment because the dispositions in question aren't reliable in the victim's environment. Presumably something similar can be said about the alien cognizers: their beliefs are justified relative to their environment but not relative to our own environment.⁶¹ But the advantage of J_{PF} over this version of Sosa's virtue reliabilism is that the proponent of J_{PF} can recognize a nonrelativized concept of justification whereas this position of Sosa's cannot. But it seems that we do have the notion of justification *simpliciter*. This is important because although it is natural to think that the beliefs of human victims of an evil demon aren't reliably formed in their environments, it isn't very natural at all to think of their beliefs as *unjustified relative to their environments*. Likewise, it isn't natural to think that the alien cognizers' beliefs are *unjustified relative to our environment*. Those relativized concepts of justification don't seem to be ones we have in our pretheoretical repertoire of concepts. What we think of the demon victims and the alien cognizers is just that their beliefs exemplify justification *simpliciter*.⁶² The problem is that Sosa needs, but doesn't seem to have, a principled way of selecting *the* environment in which reliability matters for nonrelativized justification. J_{PF} has a principled way of doing that: the environment that matters is the one for which the cognizer's faculties were 'designed'.

⁶⁰ Cohen (1984: sect II) objects to reliabilist accounts of justification on the grounds that they fail to attribute justification to demon victims. He also objects (Cohen 1984: sect III) to subjectivist accounts on the grounds that they implausibly require for justification what we rarely have, namely, further beliefs about the reliability of our beliefs.

⁶¹ This view is developed in Sosa (1991: 140–4, 284–90).

⁶² For a similar concern about views like Sosa's, see Goldman (1992: 161–3).

A third account not considered in Cohen's 1984 paper is Sosa's more recent view (BonJour and Sosa 2003: 156–61) according to which a belief is *adroit-justified* if the method by which it is formed is reliable in the actual world and it is *apt-justified* if that method is reliable in the subject's world. Thus, the demon victim's beliefs are adroit-justified but not apt-justified whereas the alien cognizer's beliefs are apt-justified but not adroit-justified. (Our own beliefs are justified in both senses.) What seems implausible about this position of Sosa's is that, according to it, there is no single sense of 'justification' in which our beliefs, the demon victims' beliefs, and the alien cognizers' beliefs are all justified. So, in order to explain our intuitions that all are justified, Sosa must attribute to us an unnoticed switch between using the term 'justified' to mean apt-justified and using it to mean adroit-justified. But it seems doubtful that this is what is going on when we make such justification attributions. According to J_{PF}, on the other hand, there is a single sense of 'justification' in which our beliefs, the demon victim's beliefs, and the alien cognizers' beliefs are justified.

A fourth account is proposed by Juan Comesaña (2002). Comesaña's account is interesting and complicated. I cannot do justice to it here. However, the upshot of it seems to be this: when we attribute justification to our beliefs, the demon's victim's beliefs, and the alien cognizers' beliefs, we are, each time we say 'their beliefs are justified,' making two claims. To use Sosa's terminology, we are saying that they are adroit-justified and that they are apt-justified. Thus, in our own case, we are completely right. But in the demon victim case we are partly mistaken since we assert both a truth (the demon victim's beliefs are adroit-justified) and a falsehood (they are apt-justified). Likewise, in the alien cognizer case we are partly mistaken since of that case too we assert both a truth (the alien's beliefs are apt-justified) and a falsehood (they are adroit-justified). These are the implications of Comesaña's position as I understand it. But J_{PF}, which accounts for the plain truth of each of these three univocal justification attributions (to us, to the demon victim, and to the alien cognizers) seems more plausible than Comesaña's position according to which we are at best partly right in such attributions due to the fact that with each attribution we're unwittingly making two assertions (one of them false) when we thought we were making only one.

3 OBJECTIONS

There are many objections that can be directed at the above proper function account of justification. Here I will address four that I take to be among the most important.⁶³

⁶³ For other ways to handle a variety of objections to talk about proper function and its role in analyzing epistemic concepts, see Plantinga (1993a, 1993c, 1995, 1996).

3.1 Proper Function and Naturalism

J_{PF} employs several notions in a way that might seem uncongenial to a naturalistic frame of mind. The three ‘offensive’ notions are the *proper function* of a cognitive faculty, its *aim*, and the environment for which it was *designed*. The proper function of a thing is the way it is supposed to function. The aim of a thing is the goal to which it is supposed to contribute. The environment for which a thing was ‘designed’ is the environment in which its functioning the way it is supposed to function *is supposed to* result in that thing’s contributing to the goal to which it is supposed to contribute. What makes J_{PF}’s employment of these notions potentially unacceptable to the naturalist is that, although they may not be problematic when applied to artifacts, they do seem problematic when applied to living organisms or their parts, which, according to naturalists, aren’t artifacts. Some people express this sort of worry by saying that proper function analyses commit their proponents to theism or some other view according to which living things, including humans, have an intelligent designer.

There is a large and growing literature on functions and goals.⁶⁴ Many naturalists are quite confident that the notion of a goal and of a function can be understood naturalistically and applied to living organisms (or their parts) even if these organisms are not the product of intelligent design. But when it comes to the idea of a function a thing is *supposed to* have or a goal to which it is *supposed to* contribute, some naturalists are more skeptical. Let’s examine some of the possible views one could hold concerning the normativity implied by such ‘supposed to’ talk.

First, consider the following three claims:

- the human heart is supposed to function so that it beats less than 200 times a minute when the person is at rest;
- the human heart is supposed to contribute to the goal of circulating blood and, ultimately, to the survival of the human in question;
- the type of environment in which a heart’s functioning that way is supposed to result in the survival of the human of which it is a part is an environment relevantly similar to the one in which we find ourselves.

These ‘supposed to’ claims seem very sensible and natural. But can they survive as more than just ordinary ways of speaking? Is it literally true that there is a function a human heart is *supposed to* have, that there is goal to which it is *supposed to* contribute and that there is such a thing as the type of environment in which its functioning in this way is *supposed to* contribute to that goal? And are such claims, taken literally and seriously, compatible with naturalism?

⁶⁴ See Nissen (1997), Boorse (unpublished), and Allen, Bekoff and Lauder (1998) for summaries of the relevant literature.

There are (at least) three positions one could take in response to these questions. Let's say that 'supposed to' talk is *reducible* to naturalistically acceptable talk just in case it is possible to give truth conditions for 'supposed to' claims (applied to parts of living organisms) using only naturalistically acceptable concepts. The *reductivist* position is that 'supposed to' talk is reducible to naturalistically acceptable talk. According to this view, the serious 'supposed to' claims mentioned in the previous paragraph are true and compatible with naturalism.⁶⁵ The other two positions—nonreductive realism and eliminativism—are nonreductivist views insofar as they agree that 'supposed to' talk is not reducible to naturalistically acceptable talk. According to *nonreductive realism*, serious 'supposed to' claims are true and incompatible with naturalism (so naturalism is false).⁶⁶ According to *eliminativism*, serious 'supposed to' claims are neither true nor compatible with naturalism.⁶⁷ The eliminativist with respect to the serious 'supposed to' claims says that, strictly speaking, there is no such thing as the *proper* function of organisms (or their parts) or the *right* environment for them; these normative notions need to be eliminated from our careful conversation which will instead employ notions that are more friendly to naturalism— notions such as *statistically normal* or *evolutionarily selected* functions or environments (or even functions *valued by us*).

Each of these three positions has its proponents; each has a legitimate claim to be taken seriously even by those who reject it. Which position is the correct one? Obviously, I can't resolve this issue here. But I will say this. In Sections 1.2.3 and 1.3 of this chapter I have defended the view that the notion of fittingness is conceptually tied to the idea of there being a *right* way for one's cognitive faculties to function—a way its faculties are *supposed to* function. If I'm right about *that*, then whatever position one takes with respect to the normativity of 'supposed to' claims about an organism's (or its parts') functions and goals—whether reductivism, nonreductive realism, or eliminativism—one ought to take that same position with respect to fittingness. But Objectivity is opposed to eliminativism with respect to fittingness. And Nonreliability is opposed to at least one strand of reductivism with respect to fittingness. So those who find Objectivity and Nonreliability plausible are forced to be nonreductive realists with respect to fittingness or to come up with some nonreliabilist way of being a reductivist about fittingness. Such people should, therefore, be either nonreductive realists or a certain kind of reductivist about the commonsense 'supposed to' claims concerning functions and goals; eliminativism is not an option for them.

⁶⁵ Karen Neander (1991a, 1991b) holds this view, as do Fred Dretske (1994, 2000) and Robert Koons (2000).

⁶⁶ This view is endorsed by Alvin Plantinga (1993a) and perhaps also by Bedau (1993).

⁶⁷ This view seems to be endorsed by John Searle (1992: 238) and by Christopher Boorse (unpublished).

In sum, my response to those naturalists who find the notion of proper function naturalistically unacceptable (i.e. my response to eliminativists about this sort of 'supposed to' talk) is twofold. First, my claim that the notion of epistemic fittingness is conceptually tied to the notion of proper function is compatible with eliminativism about the normativity of proper function. So at least that part of my argument avoids this naturalistic challenge. Second, upon accepting the conclusion that there is that conceptual tie between fittingness and proper function (and, therefore, between justification and proper function), one can take the plausibility of Objectivity and Nonreliability to provide us with a reason for rejecting eliminativism and certain versions of reductivism about proper function.

3.2 The Supervenience Thesis

The supervenience thesis (as it is sometimes called⁶⁸) says that normative properties supervene on nonnormative ones. Applied to epistemology, and, in particular, to justification, it says that justification supervenes on nonnormative properties—that there couldn't be a difference in justification without a difference in nonnormative properties. The plausibility of the supervenience thesis might give rise to the following complaint about proper function accounts of epistemic justification: 'The goal in understanding epistemic concepts such as justification is to give analyses of them in terms of nonnormative properties; but since *proper function* is a normative concept, analyses in terms of it are unsatisfying or uninteresting or in some other way unacceptable.'

Although this objection seems to be taken seriously by some people, I fail to see what the trouble is. The proponent of a proper function account of epistemic justification is like the Chisholmian who thinks epistemic concepts are reducible to ethical concepts.⁶⁹ Both think that epistemic evaluation is reducible (at least in part) to another sort of evaluation. But that in no way commits them to a rejection of the supervenience thesis. If epistemic properties are reducible to proper function (or ethical) properties and proper function (or ethical) properties are reducible to nonnormative properties, then the supervenience thesis holds. So I see no reason to think the supervenience thesis conflicts in any way with the analysis of justification given in this chapter. It's true that it would be nice to conjoin my proper function analysis of justification with an analysis of proper function in terms of nonnormative properties and that I haven't done that here.⁷⁰ But that doesn't make it uninteresting to explain one normative notion in terms of another in the way Chisholm does—not unless the *only* sort of analysis of interest

⁶⁸ See Steup (1996: 30–6).

⁶⁹ See Chisholm (1986a, 1991). For objections to Chisholm's position, see Firth (1978).

⁷⁰ But see the references in nn. 65 and 66 for attempts at and discussions of such analyses.

in epistemology is an analysis of epistemic properties in terms of nonnormative (rather than merely nonepistemic) ones.

3.3 Swampman

Against proper function analyses of epistemic properties, Ernest Sosa (1993) uses Donald Davidson's example of Swampman who comes into existence by accident as a result of a lightning strike. Swampman is supposed to be a molecule-for-molecule replica of Donald Davidson complete with Davidson's beliefs, memories, experiences, and dispositions. Given his origin, it seems that Swampman has no design plan, not even in the sense in which faculties 'designed' by evolution have a design plan. Thus, according to this objection, we can't say that his beliefs satisfy the proper function requirement. But since Davidson's beliefs are justified why not say the same of Swampman's?

Before responding to this objection, let's consider another example. Let's suppose that a lightning strike in a swamp results in a configuration of matter that in outward appearance and behavior almost exactly resembles a human woman. But Swampwoman differs from the Davidson-like Swampman not only because she resembles a woman rather than a man but also because, despite her outward appearance, she has a silicon-based brain and nervous system rather than a carbon-based one like our own. Now consider Swampwoman's unlearned doxastic response when grabbing a billiard ball in the dark. Upon grabbing the billiard ball she has tactile experience ME1 and forms the billiard ball belief B1 discussed earlier. Consider the following two ways to think of this case: (i) Swampwoman has our design plan and is believing in accord with proper function and (ii) Swampwoman has design plan B and is malfunctioning. The fact that forming B1 in response to ME1 seems to Swampwoman to be completely appropriate doesn't distinguish her from a person with design plan B who forms B1 in response to ME1 as a result of a malfunction that *also* causes this response to seem to that person to be entirely appropriate (despite the fact that for that person it is not a fitting one). She is just fortunate that, in addition to having this malfunctioning doxastic response to experience, she also has a malfunctioning experiential response to physical stimuli (having ME1 in response to grabbing a billiard ball is evidence of this latter sort of malfunction).

When we hear the Swampwoman case, we immediately impose our design plan on her and conclude that B1 is a fitting response to ME1. But suppose there were in fact silicon-based creatures, resembling humans in outward appearance and behavior, who had design plan B. What would they think of such an example? Perhaps they would conclude that Swampwoman, who they are told forms B1 in response to ME1, is malfunctioning. Why? Because they might impose *their* design plan on Swampwoman and according to that design plan, B1 does not fit ME1. In the same way, if *we* heard the Swampwoman story told

so that she formed B1 in response to *ME2* (rather than *ME1*), we would not be inclined to think that Swampwoman's belief fits her evidence.

So our judgments about the fittingness of Swampwoman's beliefs depend on which design plan we impose on her. What should we conclude about this case? One plausible-sounding response is to say that once we (a) see that our judgments about fittingness depend on which design plan we impose and (b) remember that Swampwoman doesn't seem to have a design plan, we should conclude that there is no doxastic response to *ME1* that is fitting for her—none that she is supposed to have. From this it follows that Swampwoman's belief B1 does not fit her evidence. And from that it follows that Swampwoman's belief B1 is not justified.⁷¹ The illusion that her beliefs are justified is, according to this response to the Swampwoman case, due to our imposing a design plan we have no good reason to impose given the description of the case.

Now return to the Swampman case. If we take the lesson learned from the above discussion of the Swampwoman case (i.e. that since there is no design plan, there is no sense in which the creature's belief fits the evidence and hence no sense in which her beliefs are justified) we should say that Swampman has no design plan either and that his beliefs aren't justified, despite the fact that he is a molecule-for-molecule duplicate of Davidson. If my position forced me to endorse this conclusion, this would, I think, strike many readers as a *reductio* of my position.

But this matter is more complicated than that. For consider what we'd say of Swampman's heart and lungs. Assuming Davidson's heart and lungs were in fine shape at the time of the accidental duplication, it seems that the thing to say about Swampman's heart and lungs is that they are healthy. Furthermore, it seems plausible to say that Swampman is a human (molecular replication implies DNA replication after all). We might feel *some* hesitation about saying Swampman's heart is healthy and that he's a human. But saying so seems far *more* plausible than saying that Swampman's heart and lungs are not healthy (either because they are unhealthy or because the concept of health doesn't apply) or that Swampman has Davidson's DNA but is not a human.

And if we find it plausible to say that Swampman's heart and lungs are healthy, it is reasonable also to conclude that his heart and lungs are functioning the way they are *supposed to* function. And if we can say that about his heart and lungs, we can say the same about his cognitive faculties. It's true that there is no literal designer and no evolutionary origin. But because of the physical similarities (down to the DNA level), we are inclined to think of Swampman as being human. And that leads us to think that Swampman is supposed to function in the way humans are supposed to function.⁷²

⁷¹ Given I_{PF} and that evidence is one sort of input to one's belief-forming systems. See the explanation, in Section 2.1 of this chapter, of how J_{PF} is a refinement of I_{PF} .

⁷² Cf. with Plantinga's response to the Swampman objection in Plantinga (1993c: 76–8).

What our discussion in this subsection shows is that our intuitions pull us in two different directions. We are, on the one hand, inclined to think that DNA matching indicates membership in the same natural kind. And we are inclined to think that what counts as proper function is essential to kind membership. Thus, we are inclined to think that what counts as proper function for Swampman is the same as what counts as proper function for us. That explains our intuition that Swampman's beliefs are as justified as Davidson's are.

But on the other hand, when we consider the Swampwoman case and imagine what silicon-based, human-resembling creatures with design plan B would think of her case, we realize that our reasons for assuming that Swampwoman has our design plan are no better than reasons they might have for assuming that Swampwoman has design plan B. And we are then led to conclude that the most plausible account of the Swampwoman example is that she has no design plan. This fits with our intuitions that creatures coming into existence by accident via lightning strikes have no design plans. And it has the consequence that Swampwoman is neither healthy nor unhealthy because there is nothing that counts as proper function for her. From that we may go on to conclude that none of her beliefs *fit* her evidence, in which case none of them are justified.

So which way should we go in interpreting the Swampman case? Should we say that none of his beliefs are justified or should we say his beliefs are as justified as Davidson's? I won't try to answer that question. However, I will point out that either way, we are preserving the connection between proper function and justification. For our intuitive judgments about whether Swampman has lots of justified beliefs or none waver together with our intuitive judgments about whether Swampman has a healthy heart and lungs or no healthy parts at all. If it makes sense to say that Swampman has a healthy heart and lungs, then it makes sense to say the concept of proper function applies and to say that his beliefs are fitting and justified. If it doesn't make sense to say that Swampman is a healthy creature with our design plan rather than an unhealthy creature with another design plan, then the concept of proper function doesn't apply and we shouldn't say his beliefs are either fitting or justified. Either way, the Swampman case fails to provide us with a counterexample to J_{PF}.

3.4 Skeptical Scenarios

There are a number of proposed counterexamples to proper function analyses of epistemic properties that employ strange skeptical scenarios. Some of them are dealt with elsewhere.⁷³ I'll take the time here to consider and respond to just two. One came up earlier in this chapter. The other has been proposed recently by Conee and Feldman.

⁷³ See responses by Plantinga to such objections applied to his proper function account of *warrant* in Plantinga (1993a).

Suppose an incompetent malevolent demon sets out to create a cognizer to hold unreliably formed beliefs but bungles things so badly that the result is a duplicate of an ordinary human in a standard human environment.⁷⁴ You might think that this creature's beliefs are justified even though there is an important sense in which, given the demon's wicked intentions, the creature's faculties aren't aimed at truth. But here, as in the Swampman case, we are pulled in two directions. On one way of looking at this case, it seems that the faculties aren't truth-aimed and that they are malfunctioning and only accidentally resulting in true beliefs. And that suggests that the beliefs aren't justified. But, if we stipulate that the creature turns out, by sheer accident, to be a molecule-for-molecule duplicate of a properly functioning human, then we are inclined to think that, perhaps due to DNA considerations, this creature is a human and has the same design plan humans have, no matter how it came into existence. We are inclined to think that it is *supposed to* function the way we are. In that case, insofar as its beliefs are formed after the manner of cognitively healthy human belief formation, we are inclined to think its beliefs are truth-aimed and justified.

Which is the right approach to the example? I won't try to answer that question. I'll simply note, as I did in my discussion of the Swampman case, that no matter which answer we give, we can see how our intuitions follow J_{PF} in linking justification with production by properly functioning truth-aimed faculties. For if it makes sense to say that the faculties are properly functioning truth-aimed faculties with the same design plan ours have (that the victim is forming the beliefs it is epistemically supposed to form), then it also makes sense to say that the beliefs of the lucky victim of the bungling evil demon are justified. And if it makes sense to say that the faculties are not properly functioning truth-aimed faculties, then it also makes sense to say that these beliefs aren't justified.

The other counterexample to my account of justification is proposed by Conee and Feldman:

If future cognitive scientists create human-like brains in vats for cognitive research, and design them to think as ordinary people do but to live in vat environments, then, as best as we can understand the situation, when these designed brains function properly their processes will not be reliable in the environments for which they were designed. This, we think, makes no difference to the justificatory status of their beliefs.⁷⁵

The idea is that this is supposed to count against J_{PF} because the beliefs of these brains can be justified even though they aren't produced by faculties that are reliable in the environments for which they were designed by the scientists.

The question we need to consider is how the brains in question are supposed to operate. More specifically, the question is how they are *epistemically* supposed

⁷⁴ This is the example that came up earlier in the chapter, in n. 52.

⁷⁵ This was part of their reply to my comments on their symposium paper 'Some Virtues of Evidentialism' presented at the Central Division Meeting of the American Philosophical Association in April 2005.

to operate. How something is supposed to operate is tied to its proper function; how it is *epistemically* supposed to operate is tied to the *truth-aimed* proper function of its *belief-forming faculties*.⁷⁶ So, how are these human-like brains in vats epistemically supposed to function? To consider that, we need to consider what the truth-aimed proper function of their belief-forming faculties is. And here, once again, we are pulled in two directions. On the one hand, if we think they are human-like enough to count as humans (perhaps because they have the same DNA as humans), then we think that since they are of the same natural kind as we are, the truth-aimed proper function of *their* belief-forming faculties is the same as the truth-aimed proper function of *our* belief-forming faculties. But if, on the other hand, we think that their proper function is determined by the cognitive scientists who created them, we must conclude that they do not have a *truth-aimed* proper function for their belief-forming faculties.

So which is it? Perhaps we can't say with certainty. But either way, J_{PF} seems to come out fine. If we think that what counts as their proper function is determined by the cognitive scientists who designed them, then it seems they do not have a *truth-aimed* proper function for their belief-forming faculties, in which case there isn't any way they are *epistemically* supposed to operate (even if there is some way they are supposed to operate). From this we may conclude that their beliefs don't admit of justification. But if we think that what counts as proper functioning and the right environment for *them* matches what counts as proper functioning and the right environment for *us*, then we can agree with Conee and Feldman that the beliefs of these brains can be justified despite the fact that they are unreliably formed and false. However, in that case, the beliefs will be formed by properly functioning truth-aimed faculties that are reliable when operating in their right environment. So there is no problem here for J_{PF} .

⁷⁶ See Section 2.1 of this chapter for a defense of the claim that epistemic fittingness is tied to *truth-aimed* proper function.

This page intentionally left blank

FOR EDUCATIONAL PURPOSES ONLY

6

Defeaters

In Chapter 5, I defended the following externalist account of justification:

J_{PF} : S's belief B is justified iff (i) S does not take B to be defeated and (ii) B is produced by cognitive faculties that are (a) functioning properly, (b) truth-aimed, and (c) reliable in the environments for which they were 'designed'.

My discussion there focused on the proper function requirement outlined in clause (ii). In this chapter, I want to discuss the first clause which says, in effect, that a belief is justified only if a certain sort of defeater is absent. The term 'defeater' has been used in a variety of ways by epistemologists. Thus, before I argue that clause (i) is necessary for justification, I'll take Section 1 to explain some of the different ways in which the term 'defeater' has been used. Then, in Section 2, after identifying the sort of defeater that must be absent according to clause (i) of J_{PF} , I'll defend the claim that the absence of such a defeater is necessary for justification. In the final section I will address several objections to my view on defeaters.

1 KINDS OF DEFEATERS

In surveying the kinds of defeaters there are, I'll be drawing two main distinctions. The first is the distinction between propositional defeaters and mental state defeaters. The second is between actual mental state defeaters and believed mental state defeaters. My only reason for discussing propositional defeaters is to make it clear that, although they play a significant role in many epistemological theories, they don't play any role in my account of justification. The sort of defeater that *is* the subject of clause (i) of J_{PF} is a believed mental state defeater. By saying that clause (i) of J_{PF} is required for justification, J_{PF} is saying, in effect, that all *believed* mental state defeaters are *actual* mental state defeaters. In order to give an adequate explanation of what this means, it will be important to give an account of both actual mental state defeaters (Section 1.1.2) and believed mental state defeaters (Section 1.2). In Section 1.1.3 I will clarify a difference between my views on defeaters and what Pollock says in his influential work on defeaters. These preliminary discussions will prevent us from jumping immediately into the main argument of the chapter. But because my defense of clause (i) in Sections 2

and 3 of this chapter relies on a somewhat detailed understanding of my views on defeaters, the delay will be necessary.

1.1 Propositional vs. Mental State Defeaters

I'll begin with a brief account of propositional defeaters and then continue with a more lengthy discussion of mental state defeaters.

1.1.1 Propositional Defeaters

Propositional defeaters are the sorts of defeaters had in mind by those, such as Peter Klein (1981), who propose defeasibility accounts of knowledge. Defeasibility accounts of knowledge were initially proposed to handle Gettier-cases such as the case in which you have a well-justified but *false* belief that Mr. Nogot (whom you see in the room) owns a Ford and you believe, on the basis of that justified false belief, that *someone* in the room owns a Ford. This latter belief of yours is *true* because Mr. Havit is also in the room and, unbeknownst to you, Havit owns a Ford. But clearly, your belief that someone in the room owns a Ford is not an instance of *knowledge* despite the fact that it is true and justified. Defeasibility accounts handle such Gettier-cases by adding to the 'justified true belief' account of knowledge the requirement that there be no true proposition that is a propositional defeater for the belief in question. The idea is that there must be no true (and nonmisleading) proposition which is such that if it were added to the subject's evidence base, the belief in question would cease to be justified.¹ In the case just described, the true proposition that *Nogot does not own a Ford* is a propositional defeater for your belief that someone in the room owns a Ford. According to defeasibility accounts, it doesn't matter that you don't *believe* this proposition that is a defeater. It is the mere truth of it that prevents your justified true belief that someone in the room owns a Ford from being an instance of knowledge.

1.1.2 Mental State Defeaters

Whereas propositional defeaters are propositions that are typically not believed by the subject, only mental states of the subject can be mental state defeaters.² Just as *evidence* for a belief is typically restricted to mental states of the subject,

¹ It is no easy matter to capture the difference between genuine and misleading propositional defeaters. The basic idea is that a misleading defeater is a true proposition you aren't aware of which misleadingly suggests that your belief is false or that your evidence for it isn't good evidence. For further discussion of these and related issues, see Annis (1973), Klein (1976, 1981), Lehrer and Paxson (1969), and Shope (1983: ch. 2).

² Notice that the terms 'propositional' and 'mental state' are intended to characterize the thing doing the defeating (the defeater) not the thing that gets defeated (the defeatee).

defeaters (of this sort) for a belief are also so restricted.³ The most natural examples of mental state defeaters are beliefs, but experiences can be mental state defeaters too, as can propositional attitudes other than beliefs—for example, withholdings.⁴ Because my focus will be on mental state defeaters and not propositional defeaters, I will from now on use the term ‘defeater’ to refer to mental state defeaters unless I indicate otherwise.

What are defeaters?⁵ The basic idea is that they are mental states of a person, *S*, that cause a justified belief of *S* to become unjustified.⁶ In order to get a more precise picture of what defeaters are, it will be helpful to begin by distinguishing several different kinds of defeaters. First, we need to distinguish newly acquired defeaters from continuing defeaters. A defeater newly acquired at *t* has the effect of making a belief that was justified just prior to *t* become unjustified at *t*. A continuing defeater is what you have when a newly acquired defeater continues to do its original defeating work. Second, we need to distinguish two kinds of newly acquired defeaters: newly acquired state defeaters and newly acquired power defeaters. Newly acquired state defeaters are newly acquired mental states. Newly acquired power defeaters are mental states the subject already had that gain the newly acquired power to function as a defeater. When thinking about defeaters, the examples that most naturally first come to mind are of newly acquired state defeaters. It is, therefore, common for those writing about defeaters to focus solely on defeaters of that type.⁷ For this reason, in defining defeaters below, I’ll begin by defining newly acquired state defeaters before going on to define the other kinds and defeaters generally.

Suppose I’m hiking on a trail in the mountains and that I read in a guide book, which I reasonably believe to be reliable, that upon reaching the peak at the end of the trail, one will not be able to see any lakes even when the weather is clear. I thus form the presumably justified belief that I won’t be able to see a lake from there. But when I get to the peak at the end of the trail, I look down and see

³ The fact that evidence and defeaters must be mental states of the subject shouldn’t be taken to imply that justification supervenes solely on the mental states of the subject. It implies only that factors other than the subject’s mental states which contribute to or prevent a belief’s justification aren’t properly called ‘evidence’ or ‘mental state defeaters’. See Chapter 3 (Sections 1.3 and 2.3) for objections to mentalism, the view that justification supervenes solely on the subject’s mental states.

⁴ To withhold a proposition is, roughly, to consider it and to resist both believing it and disbelieving it. See Bergmann (2005: 421) for further discussion.

⁵ Probably the best place to get started in looking for an answer to this question is Plantinga (unpublished: section III).

⁶ This is what full defeaters are. Partial defeaters are mental states that cause a belief to have *less* justification than it did. I will be ignoring partial defeaters. Also, this is what *justification*-defeaters are. One could instead focus on warrant-defeaters (where warrant is whatever makes the difference between knowledge and mere true belief). And, if one thought there were other epistemic virtues distinct from justification and warrant such as rationality of various kinds, one could choose instead to focus on rationality-defeaters. I will ignore these other kinds of defeaters as well and, for the sake of simplicity, stick with justification-defeaters.

⁷ See, for example, Plantinga (2000) and Bergmann (2005).

clearly what is obviously the lake from which the hiking trail began. That newly acquired visual experience makes my belief that one can't see any lake from the peak cease to be a justified one. This is an example of an experience that is a newly acquired state defeater. Beliefs too can be newly acquired state defeaters. While we're still a mile from the peak, my hiking companion might tell me that his new edition of the guide book says that the previous edition made the mistaken claim that one cannot see a lake from the peak in question. On this basis I form the belief that *my* guide book is mistaken on that point. Then, even before I reach the top and see the lake, this newly formed belief gives me a newly acquired state defeater for my previously held belief that one cannot see any lake from the mountain peak.⁸ I'll define this sort of defeater as follows:

NASD: D is a *newly acquired state defeater* at *t* for S's belief B iff (i) D is an experience or propositional attitude or a combination thereof, (ii) S *comes to have* D at *t*, and (iii) S's coming to have D at *t* makes B cease to be justified.⁹

Notice that this sort of defeater is defined in terms of justification. If we take newly acquired state defeaters to be, by definition, what NASD says they are, then it is analytic that a belief is justified only if there is no defeater of this sort for it.

Now consider another example. You are talking with your two brothers (one older than you, one younger) and your older brother whispers to you that when your sister comes in the room, she will be using a simple code to inform you of the whereabouts of your cousin Maggie: if your sister tells you that Maggie is downstairs, that is intended to indicate that she is in fact across town visiting her boyfriend. (You are told that the purpose of this ruse is to keep your uncle, who is in the room with you, from knowing that his daughter Maggie is off visiting her boyfriend.) Your sister comes upstairs from the basement and announces that she has just been talking with Maggie downstairs. According to the code that your older brother just whispered to you, this indicates that Maggie is across town visiting her boyfriend. You, however, aren't the least bit surprised to hear your sister say this because you already believed that Maggie was across town visiting

⁸ Propositional attitudes other than beliefs can also be newly acquired state defeaters. Take, for example, withholdings. Suppose, while hiking along the trail, I begin to reflect on my belief that the guidebook is reliable and I come to the conclusion that I really have no very powerful reasons to think that it is (though I have none for thinking it is unreliable either). If at that point I come to withhold the proposition that the guide book is reliable, I will have a defeater for my guide-book-based belief that no lakes are visible from the peak. After all, I realize then that the source isn't one I trust given that I withhold the proposition that it is reliable.

⁹ Plantinga's account of defeaters (2000) seems to assume that all defeaters are newly acquired state defeaters. The main differences between his account of defeaters there and my account here of newly acquired state defeaters have to do with the fact that he defines defeaters in terms of proper function whereas I define them in terms of justification. But given that I define justification in terms of proper function (see Chapter 5), these differences don't represent any serious disagreement between us. The benefit of my approach is that my account of defeaters can be used by those who don't define justification in terms of proper function whereas Plantinga's can't.

her boyfriend. Earlier in the day, you had spoken with Maggie's boyfriend and he told you Maggie would be at his place all day. Presumably then, even before you had this discussion with your brother and heard your sister, you had a justified belief that Maggie was at her boyfriend's place at this time.

But suppose that after this your younger brother—who, unlike your older brother, is not given to joking and in whom you have the utmost confidence—takes you aside and tells you that your older brother was pulling your leg. There had been no such arrangement for your sister to speak in code. Upon hearing this, your belief about what your sister said (that Maggie is in the basement) now gives you a defeater for your belief that Maggie is across town visiting her boyfriend. For you can see now that your sister wasn't in on any scheme to trick Maggie's father; she was simply reporting the mundane fact that she had just been speaking with Maggie in the basement. This is a case where a belief you already had (that your sister said Maggie is in the basement) gains the newly acquired power of a defeater (for your belief, arrived at earlier in conversation with Maggie's boyfriend, that Maggie is across town). We can define this sort of defeater as follows:

NAPD: *D* is a *newly acquired power defeater* at *t* for *S*'s belief *B* iff (i) *D* is an experience or propositional attitude or a combination thereof, (ii) *S* has had *D* constantly for some period of time up until and including *t*, and (iii) it comes to be the case at *t* that *S*'s having *D* makes *B* cease to be justified.

Notice that here too it is analytic that a belief is justified only if there is no defeater of this sort for it.

All newly acquired defeaters are either newly acquired mental states or mental states one already has with newly acquired defeating power. We can, therefore, define newly acquired defeaters as follows:

NAD: *D* is a *newly acquired defeater* at *t* for *S*'s belief *B* iff either (i) *D* is a newly acquired state defeater at *t* for *S*'s belief *B* or (ii) *D* is a newly acquired power defeater at *t* for *S*'s belief *B*.

Let's turn now to continuing defeaters.

Suppose that a person obtains some newly acquired defeater for a belief, which makes that belief unjustified, but continues to hold the unjustified belief. For example, suppose Jill believes with justification that her teenage son Tom is telling the truth when, in response to the accusation from Jill's neighbor that Tom has been smoking cigarettes and dropping the butts on the neighbor's lawn, Tom says that he has never smoked in his life. Then, a few minutes later, Jill looks out the window and sees Tom looking furtively around before lighting up a cigarette, smoking it, and eventually dropping the butt on the neighbor's lawn. Because Jill can't bear to believe that her son is lying to her, she makes up some lame excuse and continues to think Tom is an honest son. When Jill saw her son Tom smoking and dropping cigarette butts on the neighbor's lawn she obtained

a newly acquired defeater for her belief that Tom was an honest son. And even after she made up a lame excuse on Tom's behalf, her recently acquired defeater continued to cause her belief in her son's honesty to be unjustified. Thus, her *memory* of that visual experience of Tom smoking is a continuing defeater for her belief in her son's honesty. Or, alternatively, perhaps her *belief* that she saw Tom smoking shortly after he said he wasn't a smoker is a continuing defeater for her belief in her son's honesty. We can define a defeater of this sort as follows:

CD: D is a *continuing defeater* at t for S's belief B iff (i) at some time t^* prior to t , D (or the experience of which D is a memory) was a newly acquired defeater for S's belief B and (ii) throughout the time from t^* up until and including t , S's having D has continued to make B unjustified.

Once again, it is analytic that a belief is justified only if there is no defeater of this sort for it.

Putting all this together, we can define a (mental state) defeater as follows:

MSD: D is a (mental state) defeater at t for S's belief B iff D is either a *newly acquired* defeater at t for S's belief B or D is a *continuing* defeater at t for B.

To reiterate the point made above, the basic idea is that defeaters are mental states of a subject that cause a belief of hers that was once justified to be unjustified.¹⁰

I've identified several different types of mental state defeater: newly acquired state defeaters, newly acquired power defeaters and continuing defeaters. But so far I've said nothing about the types of mental state defeater most often mentioned in the literature: rebutting and undercutting defeaters. How are those last two defeater types related to the defeater types I've been discussing? In this subsection I've distinguished different types of mental state defeater in terms of *how long they've functioned as defeaters* (newly acquired vs. continuing defeaters) and in terms of *how long after the states were acquired they became defeaters* (newly acquired state defeaters vs. newly acquired power defeaters). But the more familiar way of distinguishing defeaters is in terms of *how they do their defeating*. And,

¹⁰ Let me note briefly two implications of the above definitions. The first has to do with the fact that, often, the appropriate response to having a defeater for a belief B is to cease holding B. But if one ceases holding a defeated belief B, then the above definitions of defeaters no longer apply to the mental states that defeated B since those definitions say only what it is for one to have a defeater for a belief one *has*. However, we can talk of mental states one still has which *were* defeaters for beliefs one once *had*. Second, the above definitions imply that if a mental state M1, which was a defeater for a belief B, ceases to make B unjustified (because, for example, some newly acquired mental state M2 neutralizes the defeating power of M1 with respect to B), then M1 is no longer a defeater for B. This implies that we don't need to distinguish (as philosophers sometimes do) between defeated and undefeated defeaters. That distinction is used to mark the difference between defeaters whose defeating power is neutralized or lost and defeaters whose defeating power remains effective. But given my definitions, there are no defeaters whose defeating power is neutralized; *all* mental states that are defeaters are undefeated defeaters. Nevertheless, I can mark the distinction in question by distinguishing mental states that *are* defeaters from those that *aren't* now defeaters for B (a belief one has) though they once were.

as already indicated, the most well-known categories of this sort are what Pollock calls ‘rebutting defeaters’ and ‘undercutting defeaters’ (Pollock 1986: 37–9).

For those unfamiliar with this latter distinction, rebutting defeaters do their work of defeating a belief by giving you a reason to think that belief is false; undercutting defeaters do their work of defeating a belief by giving you a reason to lose confidence in the source of that belief. Undercutting defeaters are often described as reasons to question whether your evidence or reasons or grounds for a belief actually indicate that the belief is true.¹¹ The easiest way to see the difference between rebutting and undercutting defeaters is to realize that you can have a reason to question the source of your belief without having a reason to think the belief is false. If you believe, on the testimony of your friend’s young child, that the door to your house is locked and then your friend tells you that her child usually can’t tell whether a door is locked, you don’t thereby learn that the door isn’t locked. Instead you learn only that the reason you had for thinking the door was locked (her child’s testimony) isn’t a good one. If, on the other hand, you walk over to the door and try it out yourself and find it unlocked, you have a rebutting defeater for your belief that the door is locked. You have not just a reason to question the source of your belief that the door is locked; you have, in addition, a reason to think your belief that it is locked is false.¹²

1.1.3 Defeatees

There are at least two different candidates for what gets defeated by defeaters: reasons and beliefs. It is obvious from the definitions in Section 1.1.2 that I’m

¹¹ See Bergmann (2005), Pollock (1986), and Plantinga (unpublished).

¹² In Bergmann (1997a: 99–103) I identified two other kinds of defeater, which, like rebutting and undercutting defeaters, are classified in terms of how they do their defeating. One I called a ‘reason-defeating defeater’ and another I called a ‘no-reason defeater’. As I noted above, according to the usual understanding of undercutting defeaters, they are reasons (mental states of some kind) for questioning whether one’s actual ground for a belief is indicative of that belief’s truth. So understood, they differ from reason-defeating defeaters which defeat only inferential beliefs. Reason-defeating defeaters work by defeating the belief on which an inferential belief is based. Once the belief on which the inferential belief is based loses its justification, the inferential belief loses its justification too. This sort of defeater for an inferential belief doesn’t cast any doubt on the suggestion that the belief on which the inferential belief is based is (if true) indicative of the inferential belief in question. That is why *reason-defeating* defeaters differ from undercutting defeaters as they are usually understood; for the latter *do* cast doubt on that suggestion. Moreover, reason-defeating defeaters differ from rebutting defeaters because the former do not give you a reason for thinking the defeated belief is false.

As for *no-reason* defeaters, they defeat only beliefs that are not based on any reasons or evidence. In order for there to be such defeaters, it must be possible for there to be beliefs that are justified even though they aren’t based on anything (see Sections 1.3.3–1.3.4 of Chapter 3 for a defense of that possibility). Assuming that is possible, no-reason defeaters work by making one lose one’s confidence in the source of a belief for which one has no reason. Here there is no defeat of a reason for an inferential belief (as in the case of a reason-defeater) and no questioning whether one’s belief ground—i.e. that on which the belief is based—indicates the truth of the belief based on it (as in the case of undercutting defeaters as they’re usually understood). Nor is there any reason for thinking the defeated belief is false (as in the case of a rebutting defeater).

thinking of defeatees as beliefs. According to the discussion there, it is a belief that gets defeated. And what that amounts to is that the belief in question ceases to be justified.

But according to some of John Pollock's work, what get defeated are reasons (1986: 38–9). Pollock thinks of reasons as things on which a justified belief can be based (thereby obtaining its justification). And he thinks that some beliefs are based not on other beliefs but on experiences (1986: 175–6). Hence, he thinks of reasons more broadly than some do, so that they include both beliefs and experiences.¹³ Pollock thinks of defeatees as reasons understood in this way. A defeater (which, like a Pollockian reason, can be either a belief or an experience) defeats a reason when it takes away that reason's power to confer justification on a belief.

There isn't much at stake in this difference of opinion about how to think of defeatees. Consider Pollock's example where I believe that the widgets moving along the conveyor belt are red and I believe this on the basis of my visual experience. Such a belief is, it seems, a justified belief. But suppose a shop superintendent comes along and tells me the widgets are being irradiated by a red light so that they will look red even if they aren't red. My new realization that the widgets will look red even if they aren't gives me a defeater. But what is it a defeater for? If you think defeatees are beliefs, you will say it is a defeater of my vision-based belief that the widgets are red: you will think that belief ceases to be justified. If you think defeatees are reasons, you will say it is a defeater for my reason for that belief, namely, my visual experience: you will think that visual experience loses its power to confer justification on my belief that the widgets are red. But notice that those who think defeatees are beliefs will agree that the visual experience loses its power to confer justification on the belief that the widgets are red; they'll say that that is why the belief is defeated. And those who think defeatees are reasons will agree that the belief that the widgets are red ceases to be justified; they'll say that that is the result of the visual experience being defeated. So there isn't really any substantive disagreement here as far as I can tell. It's just a difference in terminology. From now on, therefore, I will simply ignore this difference of opinion on what defeatees are and continue to assume that defeatees are beliefs.

1.2 Actual vs. Believed Defeaters

The distinction between actual and believed defeaters, like the distinction between actual and believed items of any sort, isn't the usual sort of distinction between two species of a single genus. First of all, because it is generally true that

¹³ Presumably he would allow propositional attitudes other than beliefs to count as reasons. For example, he would presumably allow that your all-consuming hope that your child comes to no harm is your reason for your introspective belief that you hope that your child comes to no harm.

every X is an actual X, it is true that every defeater is an actual defeater. It's not as if there are nonactual things in addition to the actual things there are. Just as everything that there *exists*, so also everything that there *is actual*. Or so I'll be assuming here.¹⁴ To be actual isn't just one among several ways for a defeater to be. It's the only possible way for a defeater to be.¹⁵

Second, a believed defeater is in an important way like an alleged crime. Thus, although we can sensibly argue that:

A1: given that every horse is an actual horse, we may conclude that every brown horse is an actual horse,

it isn't reasonable to argue that:

A2: given that every crime is an actual crime, we may conclude that every alleged crime is an actual crime.

In the same way, it isn't reasonable to argue that:

A3: given that every defeater is an actual defeater, we may conclude that every believed defeater is an actual defeater.

To say there is a believed defeater isn't to say there is a defeater with the property of *being believed to exist*; it's just to say that there is the belief that there is a defeater. It's true that in Section 2.2 below, I will be arguing that, given a certain account of what it is to be a believed defeater (an account to be presented in this subsection), we *can* conclude that every believed defeater is an actual defeater. But I will not be doing so via A3 which, for the reasons just noted, is a line of reasoning that is as disreputable as A2. So we can't understand the claim that there are believed defeaters as the claim that there are defeaters with the property of *being believed to exist*.

How, then, shall we understand the distinction between actual defeaters and believed defeaters? As I said, all defeaters are actual or existing defeaters. So actual defeaters are just the mental state defeaters described in MSD from Section 1.1.2. What then are believed defeaters? The basic idea is that they are things believed to be mental state defeaters, though that suggestion will need to be refined. To have a believed defeater for your belief B is to think your belief B is defeated. Thus, in giving my account of what it is to be a believed defeater, my goal will be to say what it is for a person to think her belief is defeated. I want it to turn out that

¹⁴ David Lewis (1986: 97–101) agrees that everything that there is exists but insists that not everything that there is is actual. I'll just record my position here that I think that view is deserving of the incredulous stares he so often felt directed at him.

¹⁵ This isn't to say that there couldn't be more or other defeaters than there in fact are (or were or will be)—as if there are no defeaters in other possible worlds that aren't in the actual world. It's simply to say that the possibility of there being more or other defeaters than there in fact are (or were or will be) just is the possibility of there being more or other *actual* defeaters.

children who have never heard of the term ‘defeater’ can think their beliefs are defeated and, therefore, have believed defeaters for their beliefs.

My proposal is that we think of a believed defeater as follows:

BD: S has a *believed defeater* for her belief B iff S takes B to be epistemically inappropriate.

What sort of conceptual abilities does S need if she is to take a belief of hers to be epistemically inappropriate? To answer this question, let’s compare it with two other questions:

Q1: What sort of conceptual abilities does S need if she is to take some action to be morally wrong?

Q2: What sort of conceptual abilities does S need if she is to take something to be aesthetically valuable?

A young (English-speaking) child can take some action of hers to be morally wrong even if she isn’t familiar with the word ‘morally’. It is enough that the child has in mind a kind of evaluation we would call ‘moral’ and that she thinks the action in question is to be evaluated poorly in that regard. Likewise, a small child can think something is aesthetically valuable even if the words ‘aesthetically valuable’ are unfamiliar to her. It is enough that she has in mind a kind of evaluation we would call ‘aesthetic’ and that she thinks the thing in question is to be evaluated highly in that regard. Using these two answers as our model, we can say that the right-hand side of BD can be satisfied by a small child’s belief if, in evaluating that belief, she has in mind a kind of evaluation we would call ‘epistemic’ and she thinks the belief does poorly along those lines.

2 NO-DEFEATER CONDITIONS

2.1 Kinds of No-Defeater Condition

In previous papers of mine, I’ve spoken of a requirement for justification and for warrant that I called ‘the no-defeater condition’.¹⁶ But it’s important to note that there are several different kinds of no-defeater conditions corresponding to the different kinds of defeaters mentioned in Section 1 of this chapter. In particular, there is the no-propositional-defeater condition, the no-actual-defeater condition and the no-believed-defeater condition.

The no-propositional-defeater condition is, as you might guess, the requirement that there be no propositional defeater. It is the sort of condition that is applied to knowledge by those proposing defeasibility accounts of knowledge. My account of justification has nothing to do with the no-propositional-defeater condition.

¹⁶ See Bergmann (1997b, 2000a).

The no-actual-defeater condition requires that there be no mental state defeater of the sort described in MSD:

MSD: D is a (mental state) defeater at t for S's belief B iff D is either a *newly acquired* defeater at t for S's belief B or D is a *continuing* defeater at t for B.

As I noted in Section 1.1.2, where MSD was first introduced, it is analytic that justification requires the absence of newly acquired defeaters and of continuing defeaters. But then, given MSD, it follows that it is analytic that there is a no-actual-defeater condition on justification. For it is analytic that a belief isn't justified if there is something making it unjustified. Clearly, therefore, clause (i) of J_{PF} can't be viewed as a no-actual-defeater condition on justification. For it isn't analytic that clause (i) of J_{PF} is necessary for justification. So clause (i) of J_{PF} imposes neither a no-*propositional*-defeater condition nor a no-*actual*-defeater condition.

What clause (i) of J_{PF} *does* say is that there is a no-*believed*-defeater condition on justification. That is, it says that it is necessary for B's justification that S has no believed defeater for B, where having a believed defeater for B is to be understood in accord with BD from Section 1.2:

BD: S has a *believed defeater* for her belief B iff S takes B to be epistemically inappropriate.

In saying this, the first clause of J_{PF} is saying that the mental states mentioned in the BD account of believed defeaters are *actual* defeaters. Or, to put it another way, according to clause (i) of J_{PF} , the no-actual-defeater condition, which is analytically true, entails (without analytically implying) the truth of the no-believed-defeater condition. In the next section, I will defend the necessity of clause (i) of J_{PF} for justification.

2.2 The No-Believed-Defeater Condition

Putting together what I've said so far in this chapter about how to understand clause (i) of J_{PF} , we can restate its claim—that there is a no-believed-defeater condition on justification—as follows:

NBDC: S's belief B is justified only if S doesn't take B to be epistemically inappropriate.

There are at least two ways in which one can take a belief B to be epistemically inappropriate: one can take oneself to have good reasons or evidence for the falsity of B or one can doubt the trustworthiness of the evidence or reasons on which B was based. In the former case, the believed defeater one has is an actual *rebutting* defeater; in the latter case it is an actual *undercutting* defeater.¹⁷

¹⁷ See n. 12 for a discussion of other kinds of defeaters.

But is NBDC true? That is the central question of this chapter. After explaining why it is tempting to think NBDC isn't true, I'll respond by defending it. The main worry about NBDC has to do with the fact that one can unjustifiedly believe that one's belief B is epistemically inappropriate. According to NBDC, this is *sufficient* for B's being unjustified. And many philosophers think that is a mistake because they think an unjustified belief cannot defeat a justified belief.

I've come across two lines of reasoning in support of this main worry about NBDC. One can be found in a paper by Alston (2002) where he is objecting to Plantinga's (2000: 364–5) claim that an irrationally formed belief can be a defeater for a warranted belief. Suppose that, while looking down at your hands, you form the perfectly sensible belief that you have hands. But then suppose you irrationally become convinced that you are a brain-in-a-vat under the control of some enemy of yours (a disgruntled former colleague) who is trying to mislead you into thinking you have hands when you don't. Plantinga thinks that this irrational belief about your being a brain-in-a-vat is a defeater for your belief that you have hands. Alston thinks that because it is irrational, it isn't a defeater for your justified belief that you have hands. Alston's reasoning, which appeals to how one ought rationally to *change* one's beliefs, can be represented as follows:

The 'Rational Change' Objection: Clearly, in the example just given, your belief that you have hands is more rational than your belief that you're a brain-in-a-vat. Moreover, each of these two beliefs provides you with a reason to stop holding the other. So which one gets defeated? Isn't it obvious that the one that is more rational defeats the one that is less rational? For clearly, if you are to *change* your belief set to avoid the problem created by this conflict, it would be better (more rational) for you to change it by giving up your irrational belief and keeping your rational one.¹⁸

Plantinga's discussions of defeaters, both in the passages to which Alston is objecting and in response to Alston's objection, invite this kind of argument.¹⁹ For in these discussions, Plantinga often speaks of what proper function requires concerning how we should *change* our noetic structure in response to acquired (irrational) defeaters.

But although Plantinga's discussion invites the 'Rational Change' Objection against NBDC, my own account of defeaters does not. My explanation of defeaters isn't given in terms of how one should *change* one's beliefs. Instead, it is in terms of what happens to the justificational status of one's belief upon acquiring (or continuing to have) a defeater for it. How one should change one's

¹⁸ See Alston's discussion of the madman case in Alston (2002: 198–9). It is clear that he is there relying on a principle he suggests earlier in the following passage: 'But why wouldn't the degree of warrant enjoyed by D [the belief Plantinga thinks is the defeater] and B [the belief Plantinga thinks is defeated] . . . be highly relevant to determining whether it would be more rational to reject B or to reject D? Isn't it clear . . . that in the case of incompatibility, it is most rational to stick with the more warranted contender?' (2002: 188).

¹⁹ Alston is objecting to Plantinga's discussion of defeaters in Plantinga (unpublished, 1993a, 2000). Plantinga's response to Alston is in Plantinga (2002: 272–5).

beliefs when one acquires a belief conflicting with one of them is one issue; what happens to the justificational status of the beliefs one already has when one acquires such a conflicting belief is another issue. The 'Rational Change' Objection improperly runs these two issues together.

The question concerning how one should change one's beliefs is a difficult one. As Plantinga emphasizes, the rational response to your conviction that you are a brain-in-a-vat is to give up your belief that you have hands. But as Alston emphasizes, given that it is irrational for you to think you are a brain-in-a-vat, the rational thing for you to do is to stop thinking you are. Plantinga focuses on what changes are rational *given that you think you are brain-in-a-vat*. Alston focuses on what changes *perfect* rationality would require. Whose emphasis is to be preferred?

Instead of answering that difficult question, I will turn instead to a different one: what happens to the justificational status of your belief that you have hands once you become convinced that you are a brain-in-a-vat? Is that hand belief defeated? Or is your belief that you are a brain-in-a-vat defeated? As for that last question, it's clear that your belief that you are a brain-in-a-vat doesn't get defeated because we are stipulating that it never had any justification to begin with. So it can't *lose* any justification. What about the belief that you have hands? It was justified. Does it lose its justification? I think the answer is 'yes'.

As I've already suggested, the first step to becoming convinced of this answer is to realize that the fact that Alston is right about the *change* that perfect rationality requires is a separate issue from whether your hand belief gets defeated. For we can agree with Alston about what the best change to your situation would be (namely, to give up your irrational belief that you are a brain-in-a-vat) while sensibly thinking that the hand belief loses its justification when you form your irrational belief about being a brain-in-a-vat. The second step to seeing that your hand belief loses its justification is to consider whether your belief that you have hands is an appropriate response to the rest of your evidence, which includes not only your perceptual experience but also your belief that that perceptual experience can't be relied on to indicate the truth about whether you have hands. And it seems intuitively that the belief that you have hands is *not* an appropriate response to that combined evidence.²⁰ Given all that, it seems that your belief that you have hands is an *inappropriate* response to the evidence you have. It is, therefore, unjustified in the circumstances.²¹

²⁰ Compare with a case where you come irrationally to think a thermometer you've been using to form temperature beliefs is unreliable. What does this do to your beliefs about the temperature formed using the thermometer as a guide? It seems to make those temperature beliefs (if you continue to hold them) unjustified. For you think (we can stipulate) that they were formed by relying on an unreliable thermometer. The fact that your belief that the thermometer is unreliable is itself unjustified does not affect this conclusion.

²¹ The fact that I'm appealing here to evidential considerations and taking them to be sufficient to defeat a belief shouldn't be interpreted as an admission that justification turns *only* on evidential

In short, then, the problem with the ‘Rational Change’ objection is that it fails to acknowledge that an unjustified belief can be a defeater for a justified one, even in cases where the best *change* to make in response to acquiring this unjustified defeater is not to drop your defeated belief but to give up the unjustified defeater.

The second line of reasoning against NBDC’s assumption that unjustified beliefs can be defeaters is something I’ve known philosophers to mention in conversation. It focuses on the alleged *arbitrariness* involved in treating defeaters differently from reasons:

The Arbitrariness Objection: Clearly, unjustified beliefs can’t confer any justification on other beliefs inferred from them, even if those inferred beliefs are properly inferred from those unjustified beliefs. Unjustified reasons, therefore, can’t confer justification. But then something similar must be true of defeaters: unjustified defeaters can’t remove justification. There is no good reason to treat reasons differently than defeaters in this regard. To say otherwise is *arbitrarily* to subject reasons to higher standards than defeaters.

By way of response, I’d like to begin by emphasizing that I agree entirely that an inferred belief can’t *obtain* any justification by virtue of its being inferred from an unjustified belief. The question is whether it follows from this that unjustified beliefs cannot be defeaters that result in the *loss* of justification for another belief. I don’t see how it does follow.

Notice first that the case of justification production in one belief via inference from another belief is different from the case of justification production in one belief via its being based on an experience. In the former case, we think the belief which is the inferential basis must be justified if the inferred belief is to be justified. But, in the latter case, we don’t think the *experience* must be justified in order for the belief based on it to be justified—the experience just isn’t the sort of thing that can be justified. So it’s not in general true that a belief’s epistemic status can be affected by another mental state only if that other mental state is justified. Likewise, some defeaters are themselves experiences and, therefore, incapable of having *any* justificational status. But this doesn’t prevent them from negatively affecting a belief’s epistemic status.

Moreover, there seems to be a relevant difference between a case of a belief B’s becoming justified via inference from another belief and a case of a belief B’s losing justification due to the subject’s coming to hold another belief which is a defeater for B. In the former case, when it works in the ideal way (i.e. when the inference is certain), we think the degree of justification had by the belief that is the inferential basis gets *transmitted* by recognized entailment so that the

considerations. What I’m admitting is just that evidential considerations *can* defeat a belief, not that they alone can or must justify a belief. (Evidence is, after all, something that counts as an *input* to the subject’s belief-forming and defeater systems.) See Chapter 3 for objections to the suggestion that justification depends only and essentially on what mental states one is in.

inferred belief has (at best, in the ideal case) that same degree of justification. We don't think any *extra* degree of justification, not already had by the belief that is the inferential basis, gets produced in the inferred belief due to the inference alone. But in the case of defeat, there is nothing relevantly like the transmission of some degree of epistemic status (possessed by the belief that is the inferential basis) by an ideal inference. It's not as if the belief which is a defeater has a degree of *unjustifiedness* which gets transmitted by an ideal inference to the defeated belief. Nor does the belief which is a defeater have a degree of justification which gets transmitted by an ideal inference. In cases of defeat, justification is *lost* not transmitted.

So, we have no reason to think it is *generally* the case that a belief's justificational status can be affected by another mental state only if that other mental state is justified. And we have good reason to think that the way justification gets produced via inference is importantly different from the way justification is defeated via beliefs that are defeaters. This leaves us without any good reason to accept the central inference employed in the Arbitrariness Objection—namely, that because an inferred belief can't obtain any justification by virtue of its being based on an unjustified belief, it follows that unjustified defeaters can't result in the loss of a belief's justification. That inference is unmotivated.

In addition, there is the independent intuition mentioned in response to the 'Rational Change' objection, namely, that it just seems intuitively that if you have a belief B that was justified on the basis of some evidence E and you acquire a further belief (whether justified or not) that E isn't a trustworthy guide to the truth about B, then B is an epistemically inappropriate response to your combined evidence (i.e. E together with this new belief that undercuts the perceived evidential value of E). This, together with the points summarized in the previous paragraph, gives us a reason to think that, when it comes to their epistemic effects, there is a relevant difference between what's required for beliefs to be justification-producing reasons and what's required for beliefs to be justification-destroying defeaters—in which case it isn't arbitrary to treat them differently.

I've considered and responded to two lines of reasoning in support of the view that unjustified beliefs can't be defeaters. But there's another objection to NBDC very similar to this one that allows that sometimes unjustified beliefs can be defeaters. According to this objection, although it is sometimes the case that unjustified beliefs can be defeaters, it's false that *thinking your belief B is epistemically inappropriate* will, of necessity, give you a defeater for B.²² The cases typically proposed as counterexamples are ones in which a subject has what we would normally take to be a justified perceptual belief while at the same time the subject is unable to refute some skeptical argument. For example, the subject may believe the car she is in is moving even though she can't think of any response to Zeno's arguments against motion. Or she may believe in material

²² See Greco (unpublished).

objects despite her inability to find a problem with Berkeley's defense of idealism. Or she may believe she has hands despite being unable to think of any good way to respond to Descartes' argument for skepticism. These are viewed as counterexamples because we are supposed to conclude that the commonsense beliefs in these cases clearly are justified even though the subject is troubled by the skeptical arguments in question.

But the cases, as described above, are ambiguous. We can understand them, on the one hand, as cases in which the subject recognizes her inability to respond to the skeptical arguments but concludes that there must be *something* wrong with the argument since she clearly does know she's moving or that there are material objects or that she has hands. If we understand the cases in that way, then I agree that the beliefs in question are justified. But that is consistent with NBDC since she doesn't take her beliefs to be epistemically inappropriate. However, we could, on the other hand, understand the cases in such a way that, because the subject is convinced by the skeptical arguments (or at least sufficiently shaken by them), she loses confidence in the source of her belief in motion or her belief in material objects or her belief that she has hands. If she really *does* lose confidence in the trustworthiness of those belief sources and yet continues to hold those beliefs, then it seems those beliefs aren't justified.²³ And that supports NBDC. It might also be that her confidence in the trustworthiness of her belief sources wavers. At some points, she really does mistrust them while continuing to hold the beliefs. And at those times, those beliefs are unjustified. But then she regains confidence in those sources (or loses her mistrust of them). At those times, the beliefs produced by those sources are not defeated (in that way). It seems, therefore, that we can handle these alleged counterexamples in a way that is consistent with NBDC.

This completes my defense of NBDC and, therefore, of clause (i) of J_{PF} which requires for justification that the person holding the belief not take that belief to be defeated (i.e. epistemically inappropriate). In the final section of this chapter, I will address four objections to various aspects of my views on defeaters.

3 OBJECTIONS

3.1 Concerns About Explanatory Circularity

The first objection I'd like to consider is that the position I've been laying out in this chapter and the previous one involves some sort of explanatory circularity. In the previous chapter, I proposed J_{PF} which explains justification in terms of the absence of a defeater whereas in this chapter, in Section 1.1.2, I explained defeaters as things that take away justification. What gain is there, one

²³ In support of this point, consider again the temperature belief example in n. 20.

might wonder, in explaining justification and defeaters in terms of each other? This concern is based on a misunderstanding. Because I focus on two kinds of defeaters—actual defeaters and believed defeaters—there is no circularity. J_{PF} in Chapter 5 explains justification in terms of *believed* defeaters whereas Section 1.1.2 of this chapter explains *actual* defeaters in terms of justification.

But doesn't that just show the circle is a little larger? For actual defeaters are explained in terms of justification, justification is explained in terms of believed defeaters, and believed defeaters are to be understood in terms of beliefs about *actual* defeaters. Isn't that a circular explanation? No. For although it is accurate enough to say that believed defeaters are explained in terms of *beliefs about* actual defeaters, it isn't right to say that believed defeaters are explained in terms of actual defeaters. It's important to recognize that ultimately, believed defeaters are explained in terms of a subject's beliefs. So here too there is no explanatory circularity.

3.2 Necessity Revisited

In Chapter 5, I objected to the evidentialist's endorsement of:

Necessity: the fittingness of doxastic response B to evidence E is an essential property of that response to that evidence.²⁴

Now consider the following three claims:

ALL: For any doxastic response B and any evidence E, the fittingness (or unfittingness) of doxastic response B to evidence E is an essential property of that response to that evidence.

NONE: For no doxastic response B and no evidence E is the fittingness (or unfittingness) of doxastic response B to evidence E an essential property of that response to that evidence.

SOME: For some doxastic response B and some evidence E, the fittingness (or unfittingness) of doxastic response B to evidence E is an essential property of that response to that evidence.

ALL says the same thing that Necessity says. One can deny ALL in two different ways: while endorsing NONE or while endorsing SOME. *My* way of denying ALL (i.e. Necessity) is to do so while endorsing SOME. This is made clear when I defend NBDC. For suppose that evidence E is *S's taking B to be epistemically inappropriate*. According to NBDC, it is an essential property of doxastic response B to E that B is an unfitting response to E.

²⁴ Recall that the idea is that if B is a fitting response to E by itself, then, even if B could be an unfitting response to evidence that includes E *and more besides*, it couldn't be an unfitting response to E by itself.

The objection that arises from these considerations is that once we endorse SOME rather than NONE, there's no telling how many cases there will be where the fittingness of a doxastic response to some evidence will be an essential property of that response to that evidence. This gives us reason to be skeptical of J_{PF} according to which there is only that one example.

Before responding, let me first give the context for my response by highlighting the fact that J_{PF} allows for two different ways in which a doxastic response to evidence can be unfitting: it can be *essentially* unfitting to that evidence or it can be *contingently* unfitting to that evidence. In order to explain how this works, it will be helpful to remind ourselves that an actual defeater is just a mental state defeater which we defined earlier as follows:

MSD: D is a (mental state) defeater at t for S's belief B iff D is either a *newly acquired* defeater at t for S's belief B or D is a *continuing* defeater at t for B.

In light of the discussion in Section 1.1.2 of newly acquired defeaters and continuing defeaters, we may conclude from this that an actual defeater for a belief B is a mental state S has that either begins or continues to make B unjustified. To put it another way, an actual defeater for a belief B is a mental state S has to which B is an unfitting doxastic response. I argued in Section 2.2 of this chapter that all believed defeaters are actual defeaters. In doing so, I argued that a belief B is an *essentially* unfitting doxastic response to a believed defeater for B. That's the first way in which J_{PF} allows for a doxastic response to be an unfitting response to one's evidence.

But J_{PF} also requires for a belief B's justification that the cognitive faculties involved in the production of B be functioning properly. These cognitive faculties will include any defeater systems whose operation is relevant to whether B is held. They too must be functioning properly in order for B to be justified.²⁵ So, even in a case where the person holding B has no believed defeater for B, it might be the case that she epistemically *should* have a believed defeater for B and that the reason she doesn't is that her defeater systems involved in the production of B aren't functioning properly. Consider an example. Suppose Jill has just been told, by someone she rightly considers to be a reputable source, that either Jill's mother or her father will be out of town for the entire day. And on that basis she infers that she will win her bet with her brother that her parents won't both be in town that day. And on the basis of that she quite happily believes that she will soon be receiving from her brother \$300, which will enable her to buy a new bicycle. A few moments later, she sees her mother and father walk into the room together,

²⁵ What are defeater systems? They're similar to belief-forming systems, except that the result of their operation is belief *loss* not belief *production*. When the perceptual component of a human's cognitive faculties is functioning properly, the person forms certain perceptual beliefs about the external world in response to sensory experience input. Likewise, when the defeater system component of a human's cognitive faculties is functioning properly, the person gives up certain beliefs in response to various kinds of experiential or doxastic input.

greeting her. This is a case where we think Jill has an actual defeater for her belief that she will soon be receiving \$300. And this is so even if Jill fails to put two and two together and, therefore, fails to have a believed defeater for her belief that she will soon be receiving the \$300. For Jill *should* have a believed defeater for that belief.²⁶ If she doesn't, it would seem that her defeater system isn't functioning properly.

But notice that it is a *contingent* feature of Jill's doxastic response (i.e. her belief that she will soon be receiving \$300) to that evidence (which includes her visual experience of seeing her parents as well as her recollection of how she came to the belief that she will soon be receiving \$300) that it is unfitting. It is a contingent feature because it depends on what counts as proper function for the cognizer having that doxastic response to that evidence. If the cognizer in question were a member of a species that had much weaker powers of reasoning than we humans do, then that doxastic response might not be unfitting, any more than our doxastic responses are unfitting when we fail to notice conflicts in our beliefs that are exceedingly difficult for us to detect. Consider, for example, a normal undergraduate philosophy major who is told by her rightly respected professor that from the premises the student has confessed to believing, it follows that *p*. The professor presents the reasoning from the student's confessed beliefs to *p* and the student is impressed with the reasoning and believes *p*: after all, her professor is an expert at detecting the validity of arguments and when the professor rehearsed the reasoning, it seemed impeccable to the student. However, unbeknownst to either of them, two of the premises together entail the denial of *p*. It doesn't follow from this that the student's belief that *p* is an unfitting doxastic response to her evidence. However, we can easily imagine a race of cognizers who are much more intellectually able than we are for whom a belief that *p* in response to the evidence the student had *would* be unfitting.

It is considerations such as these that explain why J_{PF} allows for doxastic responses that are *contingently* unfitting responses to the subject's evidence. The contingency of the fittingness of a doxastic response to evidence arises because of differences among kinds of cognizers in terms of what the proper functioning of their defeater systems requires. In some cases, the proper functioning of their defeater systems will require certain cognizers to have a believed defeater for the belief in question in light of the evidence for that believed defeater. In other cases, with the same belief and evidence, it won't, because the cognizers in these other cases have different design plans. Thus, although believed defeaters for B are, of necessity, actual defeaters for B, for other actual defeaters for B it will be only a contingent fact that they are actual defeaters for B. And it will be contingent upon what counts as the proper functioning of the defeater systems of the cognizer holding B.

²⁶ Here and elsewhere when I say someone 'should have a defeater', I mean she *epistemically should* have a defeater (i.e. that would be the epistemically appropriate thing to do).

Let's return now to the objection, namely, that perhaps it isn't only believed defeaters that are, of necessity, actual defeaters. Perhaps there are other actual defeaters that are, of necessity (regardless of proper function considerations), actual defeaters. If this were true, we'd have to add to J_{PF} other clauses besides (i) and (ii)—we'd have to add clauses which stipulate that those other defeaters (which aren't believed defeaters) are absent. We'd need to add these clauses because the absence of these other defeaters won't be guaranteed by the satisfaction of the proper function clause (ii) of J_{PF} , just as (for reasons given in Chapter 5 Section 2.1) clause (ii) of J_{PF} won't guarantee the absence of believed defeaters.²⁷

My only response to this objection is to say that no other doxastic responses to evidence come to mind as ones that are *essentially* fitting or unfitting to that evidence, regardless of what proper function requires for the defeater systems of the one holding the belief. For this reason, I'm inclined to think J_{PF} gets things right. I recognize that my failure to see any other clauses that must be added is rather weak evidence for the conclusion that there are no other clauses to be added. So I'll just record my willingness to add clauses to handle all convincing cases of doxastic responses to evidence that are essentially fitting or unfitting responses to that evidence. (One would hope that if there were in fact a large number of such cases, they could be handled by a single clause or perhaps a few clauses generalizing over common features of the cases.)

3.3 Fumerton on Conceptual Regresses

Richard Fumerton argues that the need to avoid a vicious conceptual regress creates problems for reliabilism and other forms of externalism.²⁸ He first highlights the fact that it is desirable for an account of justification to say what justification is without appealing to the concept of justification in the analysis. Then he mentions a case that will trouble reliabilists. It's a case where a person has a memory belief produced by a reliable process but she also has 'ample evidence (misleading as it turns out) indicating that these processes are unreliable' (Fumerton 1988: 182). The memory belief will be justified according to a simple version

²⁷ The reasoning given in Chapter 5 for the conclusion that the satisfaction of the proper function clause (ii) wouldn't guarantee the absence of believed defeaters went as follows. There could be cognizers designed by a literal creator to form beliefs in a reliable way but also to take each of their reliably formed beliefs to be defeated. Furthermore, they could have been designed so that although they took their beliefs to be defeated, they ignored this fact and continued to hold them. It seems that not only is this possible, but the beliefs so produced would be unjustified, for they would be held along with believed defeaters for them.

²⁸ Given that he thinks his argument 'provides reasons for returning to a very traditional form of foundationalism' (Fumerton 1988: 178), he clearly thinks it applies not just to reliabilists but other forms of externalism as well.

of reliabilism that says reliability is sufficient for justification. But that seems to be the wrong result.²⁹ That's why reliabilists, such as Goldman, who are sensitive to this problem, require for the justification of S's belief B—in addition to B's being produced by a reliable process—that S has no available reliable or conditionally reliable process which is such that, had S used it, she would not have held B. From Fumerton's perspective, the important part of this added requirement is its insistence that there be no available conditionally reliable *inferential* process which is such that, had it been used with other available beliefs of S's as inputs, S would not have held B.³⁰

The problem, says Fumerton, is that in light of these considerations, the externalist faces an unpleasant choice. In giving an analysis of justification, foundationalists will include both a base clause and a recursive clause in their account. The recursive clause will explain how one gets inferentially justified beliefs by basing them on other *justified* beliefs. Consequently, this clause will use the concept of justification in the analysans in referring to the justified beliefs which are the basis of justified inferential beliefs. The base clause will explain how a belief gets to be noninferentially justified. On pain of conceptual regress, the concept of justification cannot be used in the analysans of the *base* clause. Now consider the choice the externalist must face when she wants to add to her base clause the added requirement mentioned in the previous paragraph. On the one hand, she could require for your belief B's justification that there be no available and conditionally reliable inferential process that could be used on your beliefs which, if used, would have resulted in your not holding B. If the externalist requires that, then she will have to stipulate that the beliefs which the available inferential process can use as potential inputs are *justified* beliefs. For, says Fumerton: 'The availability of a conditionally reliable [inferential] process that would lead S to not [hold B] would only be relevant, I assume, if one had *justified* beliefs that the conditionally reliable process could use as input' (1988: 183). But then the externalist's account of justification will be explained, in part, by appeal to the concept of justification. For it will require, for the justification of *every* belief B—including noninferentially justified beliefs covered by the base clause—that the subject not have available any conditionally reliable inferential process which, if applied to other *justified* beliefs she has, would result in her not holding B. On the other hand, if we want to avoid this problem of explaining justification using the concept of justification, we have to stick with the initial crude version of reliabilism, which seemed to give an implausible result

²⁹ A similar problem case can be proposed for externalists who require proper function. It will be a case where a memory belief is in fact produced in accord with proper function and the person holding it has ample evidence that her memory is not to be trusted.

³⁰ An inferential process is conditionally reliable if it is likely to yield true beliefs as outputs when it takes true beliefs as inputs. See Fumerton (1988: 179) and Goldman (1979: 13).

for the memory belief example from the previous paragraph. So either way, the reliabilist runs into trouble.³¹

We can put the problem Fumerton has in mind as follows. It isn't only *believed* defeaters which are actual defeaters; there are also times when a person doesn't have a believed defeater but she *should* have one. In that case too she has an actual defeater. An example of this sort is the one given in Section 3.2 in which Jill's seeing her parents gives her an actual defeater for her belief that she will get \$300; this gives her an actual defeater for that belief even if she doesn't have a believed defeater for that belief. Another example is the case of the person with the reliable memory belief whose evidence makes it the case that she should have a believed defeater even though she doesn't. The problem is to figure out how we can say when it is that a person who doesn't have a believed defeater *should* have one, without appealing to the concept of justification.

The solution, according to J_{PF}, is to appeal to the notion of proper function. Whether a person *should* have a believed defeater she in fact lacks depends on what counts as the proper functioning of her cognitive faculties, in particular, her defeater systems. If the available evidence against the reliability of her memory is such that a properly functioning person in her situation (with her design plan) would believe that her memory is unreliable, then she should have a believed defeater for her memory beliefs. Likewise, if upon seeing her parents, Jill should (given what counts as proper function for her) believe she no longer has any good reason for thinking she will be receiving \$300, then she should have a believed defeater for that belief about the \$300. J_{PF} avoids the conceptual regress because it gives an account of epistemic normativity not in terms of epistemic normativity itself (that would invite trouble of the sort Fumerton rightly thinks we should avoid) but in terms of normativity of a different sort. At this point, objectors might complain that what we really want is an account of epistemic normativity in terms of something nonnormative on which epistemic normativity supervenes, not in terms of some nonepistemic sort of normativity. Defenders of J_{PF} can agree that it would be nice to have that while insisting that J_{PF} is both informative and helpful, even if it isn't as helpful as one would like.³² And they can point out that the conceptual regress problem that worried Fumerton has been avoided.

Before considering the final objection to what I've said about defeaters, let me address one other concern in the neighborhood of the objection from Fumerton that we just looked at. Fumerton was worried about actual defeaters in cases where the person should *have* a believed defeater that she in fact lacks. One might also worry about cases where one should *lack* a believed defeater that she in fact has. For you might think that S's belief isn't actually defeated by a believed

³¹ Fumerton would, presumably, think the proper functionalist account will run into the same sort of trouble. See nn. 28 and 29 above.

³² See Section 3.2 of Chapter 5 for more on this topic.

defeater if S shouldn't have that believed defeater (if, for example, S unjustifiedly believes she has a defeater). If you *did* think that, then you would conclude that it isn't enough for the justification of noninferential beliefs (i.e. the beliefs covered by the base clause of foundationalist analyses of justification) that they are reliably formed or produced by properly functioning faculties; you would think that, in addition, the subject must have no *justified* belief—i.e. none that is the product of a reliable process or a conditionally reliable process operating on *justified* beliefs—that she has a defeater for B. But then once again, there arises the problem of the concept of justification appearing in the base clause of the analyses.

But for this supposed problem we already have the solution. For, as I argued in Section 2.2 of this chapter, it is *not* only justified believed defeaters that count as actual defeaters. *All* believed defeaters count as actual defeaters. As a result, the problem described in the previous paragraph doesn't arise.

3.4 BonJour on Quasi-Externalism

In BonJour and Sosa (2003), BonJour says that internalism requires for justification that (i) the subject is aware of positive reasons for her belief and that (ii) the subject is not aware of any negative reasons against her belief (or for the unreliability of its source). According to BonJour, a qualified externalist³³ is one who accepts the latter no-believed-defeater condition but not the former awareness requirement on justification. BonJour thinks such an externalist is in effect accepting 'one part of the internalist view'; and he suggests that this results in 'an untenable halfway house' (BonJour and Sosa 2003: 31–2). To demonstrate the untenability of this position, he argues that the only reason for thinking there is a no-believed-defeater condition on justification is also a reason for thinking there is an awareness requirement on justification. From this he concludes that it is 'hard to see how the qualified externalist can defend one part of his position without undermining the other'.³⁴

By way of response, notice first that from the fact that internalism imposes a requirement R on some epistemic property, we can't conclude that every other view which imposes requirement R on that property is in some way conceding ground to internalism. For example, the fact that internalists think true belief is a necessary condition of knowledge doesn't show that anyone who thinks true belief is necessary for knowledge is making a damaging concession to internalists. What is most distinctive about internalism is its insistence on an awareness requirement on justification. But a no-believed-defeater condition

³³ A qualified externalist isn't an externalist with the requisite credentials for being an externalist. Rather, she is a person who is only a quasi-externalist rather than a full-fledged one.

³⁴ BonJour and Sosa (2003: 33). Greco (unpublished) takes seriously a similar charge. My response to BonJour applies to Greco as well.

isn't an awareness requirement. A no-believed-defeater condition doesn't require awareness of anything. At best, it requires the absence of awareness of something (i.e. defeaters). Furthermore, the no-believed-defeater condition seems to be an independent requirement on justification which, like the true belief requirement on knowledge, is accepted by many on both sides of the internalist–externalist divide. It isn't uncommon for paradigm externalists to think that a no-believed-defeater condition is necessary for justification or warrant.³⁵ Nor is it uncommon for internalists who recognize that externalists allow for this to say that thinking a no-believed-defeater condition is necessary for justification or warrant doesn't make you an internalist or disqualify you from being an externalist.³⁶ When internalists object to externalist views, they typically don't focus on the fact that externalists (at least some of them) think that a no-believed-defeater condition is necessary for justification. The correctness of the no-believed-defeater condition isn't what is at issue in the disagreement between internalists and externalists. So by allowing that there is a no-believed-defeater condition on justification, the externalist isn't conceding any ground to internalists or moving any closer to an internalist position.

What then of BonJour's argument that the only reason for thinking there is a no-believed-defeater condition is also a reason for thinking there is an awareness requirement? Consider what he says *is* the only reason for thinking there is a no-believed-defeater condition:

the only clear reason for the negative [no-believed-defeater] requirement is that accepting beliefs that are, as far as one can tell from one's own cognitive perspective, unlikely to be true *is plainly irrational* and irresponsible from an epistemic standpoint that aims at truth—even if those beliefs happen to be, unbeknownst to the person in question, reliably caused.³⁷

How does this reason for thinking there is a no-believed-defeater requirement on justification differ from saying that *it just seems intuitively* (perhaps after considering many examples) that a belief isn't justified if one has either a reason for thinking it false or a reason for doubting the reliability of its source—and that this is so whether or not the belief is in fact reliably formed? That is basically what I *did* say, in Section 2.2 of this chapter, when I explained the positive motivation for thinking NBDC is necessary for justification. And I can't see how it differs from the reason BonJour gives in the passage above (though if it does differ, then he's mistaken to say he's identified the *only* reason for thinking a no-believed-defeater condition is necessary for justification).

Thus, it seems that BonJour's complaint goes like this: 'The reason given for imposing a no-believed-defeater condition on justification is its intuitive

³⁵ Goldman (1986: 62–3, 111–12), Nozick (1981: 196), and Plantinga (1993a: 40–2).

³⁶ BonJour (1985: 37–41), Lehrer (1990: 165–6), and Moser (1985: 128–9).

³⁷ BonJour and Sosa (2003: 32), emphasis added.

obviousness. But the awareness requirement is also supported by intuitive obviousness. Since both requirements are supported by the same reason (i.e. their intuitive obviousness), it is untenable for the externalist to endorse the no-believed-defeater condition without also endorsing the awareness requirement.'

But that way of arguing isn't very convincing. One perfectly sensible externalist reply is to say that although the no-believed-defeater condition seems intuitively obvious, the awareness requirement does not. Another sensible response is to allow that, when one first considers the matter, there is intuitive support for the awareness requirement, but this intuitive support is completely outweighed by the argument in Chapter 1 for the conclusion that there is no such requirement on justification. And since there is no parallel argument that undermines the intuitive obviousness of the no-believed-defeater condition, qualified externalism is not an untenable halfway house. Indeed, such externalism isn't really even properly called 'qualified' since endorsement of a no-believed-defeater condition doesn't in any way force one to back off from the externalist rejection of awareness requirements or to make any concessions to internalism.

The main purpose of this chapter was to follow-up my Chapter 5 defense of clause (ii) of J_{PF} with a defense of clause (i) of J_{PF} . I defended clause (i) in Section 2.2 of this chapter where I argued that having a believed defeater prevents a belief from being justified, even if that belief that you have a defeater is itself unjustified. That defense was preceded by an overview of different kinds of defeaters and no-defeater conditions and it was followed by a response to various objections to my views on defeaters. Together, Chapters 5 and 6 constitute an initial presentation and defense of a Reid-inspired externalist view of justification. Given that the view I've defended in these two chapters is clearly an externalist view, a successful defense of it counts as a defense of the truth of externalism. In the remaining two chapters I will consider two objections to externalism, namely, that it commits one to approving of epistemic circularity and that its response to skepticism is inadequate in a variety of ways. In each case I will argue that the charges apply to nonexternalist views as well and, furthermore, that the alleged problems that *do* apply to externalism aren't genuine problems.

This page intentionally left blank

FOR EDUCATIONAL PURPOSES ONLY

Epistemic Circularity

In this chapter and the next, I'll be considering two objections to externalism. The objection that will be the focus of this chapter is that externalists are forced to approve of something that we shouldn't approve of, namely, epistemic circularity. In Section 1 I'll lay out this objection and in Sections 2 and 3, I'll respond to it by giving two arguments for the conclusion that externalists and nonexternalists alike should sometimes approve of epistemic circularity, in which case, the fact that externalists are forced to approve of it isn't a defect of their position. In Section 4, I'll explain why approving of epistemic circularity seems like such a bad thing even though it isn't. In the final section, I'll discuss the way I think epistemic circularity is likely to enter into our noetic structure in an acceptable manner.

1 THE 'EPISTEMIC CIRCULARITY' OBJECTION TO EXTERNALISM

The 'epistemic circularity' objection to externalism is just the complaint, highlighted above, that externalists are forced to approve of epistemic circularity and that this counts as a reason to reject their view. In this first section, after saying what epistemic circularity is, I will explain the reasoning behind that objection.

1.1 Epistemic Circularity

Before saying what epistemic circularity is, it will be good to begin with a couple of examples to help us to understand what it is. Consider the following dialogue:

- Juror #1: You know that witness named Hank? I have doubts about his trustworthiness.
- Juror #2: Well perhaps this will help you. Yesterday I overheard Hank claiming to be a trustworthy witness.
- Juror #1: So Hank claimed to be trustworthy did he? Well, that settles it then. I'm now convinced that Hank is trustworthy.

Is the belief of Juror #1 that Hank is a trustworthy witness justified? Most of us would be inclined to say it isn't. Juror #1 begins by having some doubts about

Hank's trustworthiness, and then he comes to believe that Hank is trustworthy. The problem is that he arrives at this belief on the basis of Hank's own testimony. That isn't reasonable. You can't sensibly come to trust a doubted witness on the basis of that very witness's testimony on his own behalf.¹

Now consider the following soliloquy:

Doubter: I have some doubts about the trustworthiness of my senses. After all, for all I know, they are deceiving me. Let's see . . . Hey, wait a minute. They are trustworthy! I recall many occasions in the past when I was inclined to hold certain perceptual beliefs. On each of those occasions, the beliefs I formed were true. I know that because the people I was with confirmed to me that they were true. By inductive reasoning, I can safely conclude from those past cases that my senses are trustworthy. There we go. It feels good to have those doubts about my senses behind me.

Doubter seems to be making the same mistake Juror #1 was making. He is relying on a doubted source to confirm its own trustworthiness. For he depends on the testimony of the people around him to confirm the trustworthiness of his senses. But it was only by relying on his senses that he believed there were people with him and that they confirmed what his senses told him.

Doubter's belief is an example of a belief infected with epistemic circularity. (I'll call beliefs so infected 'EC-beliefs'.) But what exactly is epistemic circularity? Epistemic circularity arises in connection with beliefs about the trustworthiness or reliability of one's own belief sources—beliefs such as 'my belief source X is trustworthy' or 'that belief of mine wasn't formed in that unreliable way'. Epistemic circularity can afflict both arguments and beliefs. Suppose I hold a belief—either inferentially or noninferentially—in the trustworthiness of one of my belief sources, X. If, in holding that belief, I *depend upon X*, then it is an EC-belief. (To depend upon a belief source X in holding a belief B is for B either to be an output of X or to be held on the basis of an *actually employed* inference chain leading back to an output of X.²) Likewise, if I propose an *argument* for the conclusion that X is trustworthy and my belief in one or more of the premises depends upon X, then that argument is epistemically circular.³ And if I base my

¹ You might think that by using a lie detector test, you can come to trust a doubted witness on the basis of that witness's testimony on his own behalf. But that is a case where you have testimony from another source confirming that person's testimony. The other source is that person's heart rate etc. which is monitored in the lie detector test and which can either confirm or disconfirm claims made by that person. Likewise, some people think that they can look a person in the eye and tell whether that person is lying. What's happening in such cases is similar to what happens when using a lie detector test. The listener is using another source (body language, eye appearance) to confirm or disconfirm the witness's claim that he is trustworthy.

² What matters in determining what a belief is based on is *not* the manner in which the belief was originally acquired but rather with the way it is currently being sustained.

³ There are at least two different kinds of epistemically circular arguments that a person, S, may rely on. Both have as a conclusion a proposition asserting the trustworthiness of one of S's own

belief in the argument's conclusion on such an argument, then my belief in that conclusion is an EC-belief. Although the most common examples of EC-beliefs are like Doubter's in that they are based on epistemically circular *arguments*, it is possible (as we will see in Section 5) for there to be noninferential EC-beliefs that aren't based on arguments of any sort.⁴

Our understanding of epistemically circular arguments can benefit by contrasting them with logically circular arguments. An argument is *logically* circular, in the sense I have in mind, only if it includes its conclusion—the very same claim—as one of its essential premises (i.e. as a premise without which the argument would be invalid).⁵ Suppose that, as seems plausible, a belief sustained via source X is warranted (i.e. it has what it takes to turn true belief into knowledge) only if X is a trustworthy source. Then we may distinguish logically circular arguments from epistemically circular arguments as follows: a logically circular argument's conclusion is *identical with* one of the premises whereas an epistemically circular argument's conclusion is such that *its truth is required for warranted belief in* one of the premises. (Notice that for epistemically circular arguments, it is the truth of the conclusion, not belief in it, that is required for the warrant of belief in one of the premises.)⁶

Logically circular and epistemically circular arguments face a similar charge. The charge is that if an argument, A, is circular (in the relevant sense), a belief in the conclusion of A cannot be justified in virtue of being based on A. This charge is clearly correct in the case of *logically* circular arguments. For, prior to the believer's inference to the conclusion, the belief in the premise that is repeated as a conclusion will be either justified or unjustified. If that premise belief is unjustified, then the belief in the argument's conclusion won't be justified in virtue of being based on that argument (since you can't get justification to transfer from

belief sources, X. With the first kind, S's belief in one of the premises of the argument depends on X (this is the kind Doubter relies on). With the second kind, S's act of inferring the conclusion from the premises is an instance of S depending upon X. An example of this second kind of epistemically circular argument is an inductive argument for the reliability of induction. In such an argument, one cites many examples of inductive arguments (with true premises) whose conclusions have been independently confirmed and then inductively infers that induction is trustworthy. Belief in the premises of such an argument doesn't depend on induction. But the inference from the premises to the conclusion does. So the belief in this conclusion, when formed in this way, is afflicted with epistemic circularity as defined above in the text because a person is depending upon a belief source (in this case, inductive reasoning) in forming the belief that that source is trustworthy. See Braithwaite (1953: ch. 7) and Van Cleve (1984) for sympathetic discussions of inductive defenses of induction.

⁴ See Alston (1989: 326–9, 1993b: 15) for a characterization of epistemic circularity that is in some ways similar to the one given here.

⁵ Notice that including a premise that *entails* the conclusion is not the same thing as including the conclusion itself as one of the premises.

⁶ This comment applies only to those epistemically circular arguments that are afflicted with the *first* of the kinds of epistemic circularity identified in n. 3 of this chapter.

premises to conclusion if belief in one of the argument's essential premises is unjustified).⁷ If, on the other hand, that premise belief is justified independently of the inference to the conclusion, then, although belief in the conclusion will be justified, it won't be justified *in virtue of being based on* the argument, since it was already justified as a premise belief. However, in the case of *epistemically* circular arguments, although the charge is the same, there is no simple parallel line of reasoning that can be used to show that the charge is correct when applied to them.

Doubter's belief that his senses are trustworthy—like the belief of Juror #1 that Hank is to be trusted—seems to be unjustified.⁸ And the reason *seems* to be that it is an EC-belief. That suggests that epistemic circularity is a bad thing—in the sense that it prevents beliefs infected by it from being justified. As I noted in the introduction to this chapter, I'll be arguing in Sections 2 and 3 that, appearances to the contrary notwithstanding, epistemic circularity is *not*, in itself, a bad thing in this sense. But first let's consider how the assumption that it is a bad thing can be used in an argument against externalism.

⁷ Sorenson (1991) identifies several arguments that appear at first glance to be counterexamples to this claim. Each is logically circular and can be used to get a justified belief in the conclusion even if, prior to using that argument to get a justified belief in the conclusion, one didn't have a justified belief in the premise. Here are a few examples he gives (1991: 248–50):

Some arguments are written in black ink.
Some arguments are written in black ink.

There are at least two tokens of an eleven-word sentence.
There are at least two tokens of an eleven-word sentence.

Some arguments are composed solely of existential generalizations.
Some arguments are composed solely of existential generalizations.

But these are not counterexamples to my claim that belief in the conclusion of an argument can't be justified *in virtue of being based on* the argument if belief in one or more of the premises isn't already justified. For in the examples above, one's use of the argument to get a justified belief in the conclusion isn't a case of the conclusion being *based on* the argument in question. A belief B is *based on an argument A* iff the content of B is A's conclusion and B is inferred from A's premises via the reasoning A lays out. Thus, a belief in A's conclusion is not *based on A* when it is *not* inferred from belief in A's premises but is based instead solely on an awareness of some other feature of A (or a particular tokening of A), which is what is happening in Sorenson's examples above.

⁸ Remember that the examples are to be understood so that just as Juror #1 is relying on Hank's testimony in arriving at his belief in Hank's trustworthiness, so also, Doubter is relying on the reasoning noted in arriving at his belief in the trustworthiness of his senses. Prior to that reliance, Juror #1 had doubts about Hank's trustworthiness and Doubter had doubts about the trustworthiness of his senses. Furthermore, I'll be assuming that the beliefs in these examples aren't causally overdetermined by two or more different belief sources (or combinations of sources). Thus, if Doubter's belief that his senses are trustworthy is based on his senses, we may conclude that that belief isn't *also* produced (in an overdetermining way) by a distinct set of belief sources that *doesn't* include his senses.

1.2 An Objection to Externalism

Richard Fumerton and Jonathan Vogel have argued that reliabilists are committed to approving of epistemically circular track record arguments for the trustworthiness of our faculties.⁹ A track record argument is an argument of the sort employed by Doubter in the soliloquy presented above. In such an argument, one infers from the past success of a belief source in producing true beliefs—i.e. its track record—that it is a trustworthy source. Epistemically circular track record arguments are ones that include premises (used to confirm that the source in question did produce true beliefs on particular past occasions), belief in which is produced by that very source. This objection to reliabilism concludes that, because reliabilists are committed to approving of epistemically circular track record arguments, reliabilism is to be rejected. Obviously, this objection takes for granted that epistemic circularity is a bad thing.

Why think reliabilists are so committed? Take as an example a simplistic version of reliabilism according to which a belief is justified if and only if it is reliably formed. From this it follows that a perceptual belief is justified if it is reliably formed, even if the person holding it doesn't hold the belief that her sense perception is reliable. Consider the implications of this position for an epistemically circular track record argument that your sense perception is reliable:

1. I formed perceptual belief B1 at t_1 .
2. B1 was true at t_1 .
3. I formed perceptual belief B2 at t_2 .
4. B2 was true at t_2 .
- ...
- n. Therefore, most of my perceptual beliefs have been true.
- C. Therefore, my perception is reliable.

Suppose that you believe each of the premises of such an argument and that each of those beliefs is reliably formed, including the one(s) formed via sense perception. (Let's say that beliefs in the even-numbered premises were based on testimony, in which perception plays a role.) It follows that (a) you've got a rather simple and inductively valid argument (that isn't logically circular) for the conclusion that sense perception is reliable and that (b) your beliefs in its premises are justified. What could keep you from forming, on the basis of this argument, a *justified* belief in the conclusion? According to Fumerton and Vogel, it isn't reasonable to deny that justification can be transferred from premises to conclusion via this simple inductively valid argument. This is why reliabilists are viewed,

⁹ See Fumerton (1995: ch. 6) and Vogel (2000). Fumerton (2006) has since modified his presentation of this sort of objection to reliabilism.

by critics such as Fumerton and Vogel, as being committed to approving of epistemically circular track record arguments, that is, as committed to thinking that beliefs based on such arguments can be justified beliefs.

The above objection is directed at reliabilism but it applies equally to other versions of externalism such as my own proper function account given in Chapter 5. Suppose I have an epistemically circular track record argument for the conclusion that my sense perception is reliable. And suppose that I believe the premises of such an argument and that each of those beliefs, including the beliefs produced by sense perception itself, is produced by properly functioning faculties with no believed defeaters. It follows again that (a) I've got a rather simple and inductively valid argument (that isn't logically circular) for the conclusion that sense perception is reliable and that (b) my beliefs in its premises are justified. What could keep me from forming, on the basis of this argument, a *justified* belief in the conclusion? Again, it isn't reasonable to deny that justification can be transferred from premises to conclusion via this simple inductively valid argument. These remarks could be expanded to cover other versions of externalism too and I'm sure that Fumerton and Vogel intended this objection to apply to externalism generally. That's why I've cast the argument as a complaint about *externalists* being committed to approving of epistemic circularity.

A crucial assumption of the objection is, as I've noted, that epistemic circularity is obviously a bad thing (i.e. beliefs infected by it aren't justified). Given that externalists are committed to approving of it in at least some cases, they are committed to approving of a bad thing (i.e. they are committed to the false claim that EC-beliefs *can*, in at least some cases, be justified). In response to this objection, I will, in the next two sections, be arguing that epistemic circularity needn't be a bad thing. For we have good reason to conclude that it's possible for beliefs infected with epistemic circularity to be justified.

2 THE FIRST ARGUMENT IN SUPPORT OF EPISTEMIC CIRCULARITY

In this first argument defending epistemic circularity, my strategy will be to show that it isn't only externalists who are committed to approving of it. So I'll be agreeing with Fumerton and Vogel that reliabilists and externalists are committed to approving of epistemic circularity. But I'll be arguing that this is because almost all epistemologists are committed to approving of it. This is because almost all epistemologists endorse the following foundationalist thesis:

F: There can be noninferentially justified beliefs.¹⁰

¹⁰ A belief is *noninferentially* justified iff it is justified but not in virtue of being inferred from or based on another belief.

The reason almost all epistemologists endorse F is that it is exceedingly plausible, for reasons I will explain below. But given that acceptance of this exceedingly plausible view commits one to approving of (at least some instances of) epistemic circularity, we may conclude that epistemic circularity isn't, in itself, a bad thing. Of course, if we thought that approval of epistemic circularity was even more objectionable than rejecting F, then our response would be to reject F. But, in light of the defense of F to be given below, I think it is safe to assume that rejecting F is more objectionable than approving of epistemic circularity. Hence, the sensible response, upon seeing that endorsement of F commits you to approving of epistemic circularity, is to conclude that epistemic circularity isn't in itself a bad thing. That, in a nutshell, is the first argument. In spelling it out more carefully, I'll begin by arguing that F is an exceedingly plausible view and then I'll argue that proponents of F are committed to approving of epistemic circularity. (In saying someone is committed to approving of epistemic circularity, the suggestion isn't that the person is committed to approving of *all instances* of epistemic circularity but only that the person is committed to approving of *some* possible instance of it.)

2.1 In Defense of Foundationalism

Perhaps endorsement of foundationalism requires more than endorsement of F. If so, then I am here defending only one aspect of it, for I take it for granted that foundationalism includes, at the very least, a commitment to F.

To deny F is to say that all justification must be inferential—i.e. that a belief can be justified only in virtue of being inferred from another belief. But the denial of F together with a few widely accepted premises leads directly to the radical skeptical position that none of our beliefs are justified:

1. A belief can be justified only if it is inferentially justified. [i.e. $\sim F$]
2. A belief can be inferentially justified only if the belief from which it is inferred is a justified belief.
3. Therefore, a belief is justified only if it is justified via logically circular reasoning or it is justified via an infinite chain of reasoning. [from 1 and 2]
4. No beliefs can be justified via logically circular reasoning.
5. None of our beliefs are justified via infinite chains of reasoning.
6. Therefore, none of our beliefs are justified. [from 3, 4, and 5]

Since this argument is valid and premises 2, 4, and 5 are so widely accepted, the conclusion that denying F implies radical skepticism will also seem very plausible. And since the radical skeptical conclusion seems so implausible, F has been a very popular view among contemporary epistemologists.

But can this age-old regress argument for foundationalism really be any good?¹¹ Hasn't foundationalism been refuted or at least been shown to be on shaky ground? Aren't there at least defensible versions of coherentism? Aren't there at least some defenders of infinitism? In response to the second of these four questions, so-called refutations of foundationalism are typically objections to versions of foundationalism making claims in addition to F—e.g. that all justified beliefs are either believed noninferentially with absolute certainty or deductively inferred with certainty from such beliefs. Rarely do we find any attempts to refute F itself.¹² And one is hard-pressed to find coherentists who think circular *reasoning* gives rise to justification.¹³ The only versions of coherentism worth taking seriously accept premises 2, 4, and 5 while rejecting the skeptical conclusion above. But that commits them to F. Several philosophers have argued (plausibly I think) that such versions of coherentism are best thought of as foundationalist views of a special sort.¹⁴ But even if we don't want to classify such views as versions of foundationalism, it is safe to say that their proponents accept F.¹⁵

As for infinitism, its main defender is Peter Klein.¹⁶ And while he has many things to say in defense of infinitism about *propositional* justification, when he speaks of *doxastic* justification (which is our focus in this book), his responses to objectors seem more limited. In fact, I don't think his view about doxastic justification counts as a version of infinitism. For he thinks (Klein forthcoming b)

¹¹ The regress argument can be formulated as follows: As the argument above shows, it's impossible for 1, 2, 4, 5, and ~ 6 to be true at once. Therefore, at least one of them is false. 2, 4, and 5 are exceedingly plausible. And ~ 6 is more plausible than 1. Hence, 1 is false (i.e. F is true).

¹² BonJour (1985: ch. 2) might have given the impression that he was arguing against F. But in fact he was insisting only that all *empirically* justified beliefs are inferentially justified, not that *all* justified beliefs are justified inferentially (since he allowed in that book for noninferentially justified *a priori* beliefs). The problem (see Howard-Snyder 1998) is that his argument for this less general claim seemed to rely on intuitions which, if correct, would support the more general claim that all justification is inferential (see especially premises 2–4 in his BonJour 1985: 32). That might help to explain why he has now given up on those intuitions, as is evidenced by the fact that he now rejects even the more restricted claim that all empirical justification is inferential (BonJour and Sosa 2003: ch. 4).

¹³ It's true they think mutual *support* among believed propositions is helpful. But they deny that circular *basing* (which is what is involved in circular reasoning) is justification-producing. See Plantinga (1993b: ch. 4) for an excellent discussion of what is wrong with circular reasoning and how it differs from mutual support.

¹⁴ See Plantinga (1993b: ch. 4), Klein (1999: 298), and Sosa (1991: 180).

¹⁵ Susan Haack, who proposes a view she calls foundherentism, is also committed to F. As she puts it using her crossword puzzle analogy (where the clues represent experiential evidence): 'As in the case of a crossword we eventually reach the clues, so with empirical justification eventually we reach experiential evidence' (Haack 1997: 289). Ultimately, because she accepts 2, 4, and 5 from the argument above in the text, she thinks there would be no inferential justification for any of our beliefs unless there were first some noninferential justification for beliefs obtained independently of their being based on other beliefs.

¹⁶ See Klein (1999, 2000, forthcoming a, forthcoming b).

that a belief can be justified in virtue of being based on an actual chain of inference leading back to a belief that isn't itself justified (and for which one doesn't have available an infinite chain of reasons). But this is just to reject premise 2 above and allow that a belief can be inferentially justified by being based on an unjustified belief. I find this denial of 2 to be quite out of character for someone who thinks having a good reason for one's beliefs is essential for justification. In fact, without 2, there seems to be no motivation for endorsing infinitism in the first place. But although this denial of 2 strikes me as wholly implausible, infinitism itself seems even worse. For on that view, a belief is (doxastically) justified only if it is actually based on an infinite chain of reasons (each member of which is based on another). Whether or not that is something we are capable of—and I doubt very much that it is—it seems utterly clear than none of our beliefs is *in fact* held in that way. And that is enough to establish premise 5 above.¹⁷

2.2 A Foundationalist Defense of Epistemic Circularity

Having argued that F is an exceedingly plausible view, my goal in this section is to show that endorsing F commits one to approving of (at least some instances of) epistemic circularity, from which we can conclude that epistemic circularity isn't, in itself, a bad thing.

My argument that endorsing F commits one to approving of epistemic circularity depends on the claim that a sensible proponent of F will think the following four claims are *compossible*:

- (a) A subject S has belief sources, X_1 – X_n , each of which directly produces non-inferentially justified beliefs.
- (b) On the basis of the noninferentially justified beliefs produced by the belief sources mentioned in (a)—including beliefs produced by X_1 —S relies on valid inductive reasoning to infer (and, thereby, to come to believe on that basis and for the first time) that source X_1 is reliable.¹⁸

¹⁷ For further discussion of these issues, see Bergmann (forthcoming b) and Klein (forthcoming a, forthcoming b).

¹⁸ Both here and in (c), I'm understanding *valid inductive reasoning* in such a way that, in addition to satisfying the obvious formal constraints on valid inductive reasoning, it must also satisfy the following two conditions: (i) it doesn't involve the projection of unprojectable properties (such as Goodman's grueness and bleenness) and (ii) it isn't impossible for counterexamples (to the inductively inferred conclusion) that are in fact possible, to show up in the sample class on which the induction is based (so we can't conclude from the fact that all ravens in our sample class have been observed that all ravens have been observed, since, even if there are unobserved ravens, it is impossible for unobserved ravens to show up in our sample class which, of necessity, contains only what is observed). See Goodman (1983: chs. 3–4), Plantinga (1993a: ch. 7), and Pollock (1990: chs. 3–5) for a discussion of these conditions on appropriate inductive reasoning.

- (c) Justification transfers from an argument's premises to its conclusion when S relies on valid inductive reasoning to infer the conclusion (which is then believed on that basis and for the first time)¹⁹ from the already justified premises (*unless* in making that inference S comes to take—or epistemically *should* come to take—her belief in one of the premises or her reliance on that inference to be epistemically inappropriate).²⁰
- (d) In making the inference mentioned in (b), it's false that S comes to take—or epistemically *should* come to take—her belief in one of the premises mentioned in (b) or her reliance on the inference mentioned in (b) to be epistemically inappropriate.

Notice that (a)–(c) together entail that, so long as the 'unless' clause in (c) isn't satisfied, there exists a justification-producing epistemically circular track record argument. And (d) tells us that the 'unless' clause isn't satisfied in the case in question. Thus, if you allow for the compossibility of (a)–(d), you will be committed to thinking that it is possible for there to be justified EC-beliefs. From this we may conclude that any proponent of F who thinks (a)–(d) are compossible is committed to thinking that epistemic circularity needn't be a bad thing.

Before considering whether sensible proponents of F are committed to the compossibility of (a)–(d), it will be helpful to understand the motivation for adding the 'unless' clause in (c). What would be wrong with (c) if we dropped the 'unless' clause? Why not say that justification *always* transfers from premises to conclusion when one believes the conclusion on the basis of a valid inference from justified premises? Consider the following example. Suppose that you justifiedly believe that *p* and that *q* but that you've never before connected those beliefs in your mind or considered what they together imply. Now suppose someone gets you to consider the proposition that *p* & *q* entails *r* and that, upon considering that conditional, you conclude that it seems obviously true and you justifiedly believe it. It is then pointed out to you that you now have a good argument for *r*. But when you consider *r*, you find it utterly absurd—so absurd that

¹⁹ Notice that this parenthetical clause prevents (c) from saying that justification *transfers via inference* from premises to conclusion in logically circular arguments. In logically circular arguments, the conclusion is just a restatement of one of the premises. So coming to believe that conclusion on the basis of the premises of the logically circular argument won't be a case of coming to believe the conclusion for the first time.

This parenthetical clause also prevents (c) from saying that justification transfers in the cases Wright (2002) wants to disallow. As I understand him, he thinks that if the justification for the premises depends on one's already justifiably believing the conclusion, then the justification can't transfer from the premises to the conclusion via the inference in question (2002: 338). But in that case, believing the conclusion on the basis of those premises won't be a case of coming to believe the conclusion for the first time.

²⁰ To take a belief to be epistemically inappropriate is to have a believed defeater for it. See Chapter 6, Section 1.2 for a discussion of what believed defeaters are. And to take reliance on an inference to be epistemically inappropriate will involve taking the belief so inferred to be epistemically inappropriate.

you are far more certain of the falsity of r than you are of the truth of p or of q . In such a situation, upon recognizing that two of your beliefs—your belief that p and your belief that q —together entail something you consider to be utterly absurd, you will naturally think that you have a defeater for at least one of those two beliefs. Suppose that what happens is that you get a believed defeater for Bp , which is the belief (of those two) in which you're least confident. Under these circumstances, what does the original argument you were given for r do for you? Would the justification you had for Bp , Bq , and $B[(p \& q) \rightarrow r]$ transfer, via that argument, to a belief that r ? Certainly not. If, as a result of making the inference in question, you get a believed defeater for one of the premise beliefs, you thereby lose your justification for that premise belief (since, as I argued in Chapter 6, believed defeaters are actual defeaters and actual defeaters prevent justification). The result is that the justification doesn't transfer because it no longer exists for one of the premise beliefs.

What happens if making the inference epistemically *should* have resulted in your having a believed defeater (due to what is required by the proper functioning of your defeater systems), but you don't in fact get one?²¹ In this case too you fail to have justification for your belief in the conclusion, so the justification doesn't transfer. But this time, the reason is that your belief formed via inference wasn't produced by properly functioning cognitive faculties. Your defeater systems weren't working properly and that resulted in your not having a believed defeater you should have had.²² So, if in making the inference you either *have* a believed defeater for one of the premise beliefs (or the inference) or you *should have* a believed defeater but you don't, justification doesn't transfer. This is the motivation for the 'unless' clause in (c).²³

Now that we understand the motivation for the 'unless' clause in (c), let's consider whether a sensible proponent of F will think (a)–(d) are compossible. Notice first that (d) entails (b) since (d) says that the inference mentioned in (b) occurs without a defeater. Notice also that (b) entails (a), since (b) says an

²¹ See Chapter 6, Section 3.2 for a discussion of the proper functioning of our defeater systems and how this relates to cases where we epistemically should have a believed defeater but don't.

²² As I noted in the previous chapter, when I say here and elsewhere that someone 'should have a defeater,' I mean she *epistemically should* have a defeater (i.e. that would be the epistemically appropriate thing to do).

²³ In Bergmann (2004a), I stated the 'unless' clause in (c) as follows: 'unless the denial of the conclusion seems or should seem to S, when she considers it, to be more plausible than the premises'. But this takes care of only the kind of rebutting defeater discussed above, not of all kinds. For example, it doesn't take care of the situation where, in making the inference, you come to have doubts about the reliability of the source of one of your premise beliefs without thinking you have any reason to think any of your premise beliefs are false. There too you have a believed defeater for a premise (an undercutting one, not a rebutting one) and lose justification for one of your premise beliefs; there too there is a failure of justification transfer. Another example: suppose that, upon making the inference, you come to take the inference itself to be epistemically inappropriate. In that case too the inference will not transfer justification from premises to conclusion, this time because you take yourself to have an undercutting defeater for a belief in the conclusion.

inference is made from beliefs produced in the way (a) describes. Moreover, given that (c) is a noncontingent proposition, its possibility entails its necessity. Thus, if we can show that each of (a)–(d) is individually possible, we may conclude that they are compossible. For since (d) entails (b) and (b) entails (a), it follows that if each of (d), (b), and (a) are individually possible then their conjunction is possible. And since (c)'s possibility entails its necessity—and, hence, its compossibility with whatever is possible—it follows that if (c) is possible and the conjunction of (a), (b), and (d) is possible, then (a)–(d) are compossible. Thus, to determine whether a sensible proponent of F will think (a)–(d) are compossible, it will be sufficient to consider whether a sensible proponent of F will think each of (a)–(d) is individually possible.

Let's start with:

(a) A subject S has belief sources, $X_1 - X_n$, each of which directly produces noninferentially justified beliefs.

The only thing that (a) adds to

F: There can be noninferentially justified beliefs

is the suggestion that noninferentially justified beliefs can be produced by more than one belief source in the same person. Any sensible person who thinks F is true will, therefore, think that (a) is possible too.

Consider next the possibility of:

(b) On the basis of the noninferentially justified beliefs produced by the belief sources mentioned in (a)—including beliefs produced by X_1 —S relies on valid inductive reasoning to infer (and, thereby, to come to believe on that basis and for the first time) that source X_1 is reliable.

One way to question (b)'s possibility is to say that it isn't possible for the reasoning in question to be valid. But it seems clear that the inductive reasoning used in epistemically circular track record arguments *is* inductively valid. It is modeled exactly after paradigm cases of inductively valid arguments. Another way to question (b)'s possibility is to say that it is impossible to have a belief that is the output of a belief source X without also believing that X is reliable. The idea would be that by relying on the belief source (which, of course, you do in holding a belief produced by that source), you are, *of necessity*, implicitly believing that the source is reliable. This suggestion strikes me as implausible. Presumably small children can hold beliefs without also believing that the sources of those beliefs are reliable. For although they hold beliefs, some might not even be capable of recognizing that having beliefs involves having a belief source or that such a belief source may or may not be reliable. The same thing might be true of unsophisticated adults

who are unaccustomed to reflection. Sensible proponents of F would not, for this reason, deny the possibility of (b).²⁴

Let's turn next to:

(c) Justification transfers from an argument's premises to its conclusion when S relies on valid inductive reasoning to infer the conclusion (which is then believed on that basis and for the first time) from the already justified premises (*unless* in making that inference S comes to take—or epistemically *should* come to take—her belief in one of the premises or her reliance on that inference to be epistemically inappropriate).²⁵

To see the plausibility of (c), let's begin by noting that valid inference from justified beliefs seems, in most cases, to yield justified beliefs. Next, notice that (c) stipulates that the inferred belief has never been believed in the past and that making the inference neither does nor should lead the subject to take one of the premise beliefs or the inference itself to be epistemically inappropriate. How, in such a circumstance, could justification fail to transfer? There are certainly many cases like this in which justification *would* transfer. The worry is that justification won't transfer in a case like the one mentioned in (b). But if the premise beliefs are justified and the inference is valid and the subject neither does nor should take the premise beliefs or her reliance on the inference to be epistemically inappropriate, how could the justification fail to transfer? It would seem that anyone tempted to think there is a problem with (c)—when it is applied to inferences like the one mentioned in (b)—is in fact concerned either about the validity of such inferences or about whether the subject should take such premise beliefs or her reliance on such inferences to be epistemically inappropriate. That is, worries about (c) seem to be misplaced worries about (b) or (d). I conclude, therefore, that sensible proponents of F will think (c) is true and hence possible.

Consider finally:

(d) In making the inference mentioned in (b), it's false that S comes to take—or epistemically *should* come to take—her belief in one of the premises mentioned in (b) or her reliance on the inference mentioned in (b) to be epistemically inappropriate.

²⁴ Moreover, this particular objection to (b)'s possibility could be avoided by replacing (b) with (b*): *On the basis of the noninferentially justified beliefs produced by the belief sources mentioned in (a)—including beliefs produced by X₁—S relies on valid inductive reasoning to infer (and, thereby, to come to believe on that basis) that source X₁ is reliable, something she has never before justifiedly believed.* The possibility of (b*) is consistent with the (implausible) claim that it is impossible to have a belief that is an output of X without also believing X is reliable.

²⁵ In the previous footnote I said we could change (b) to (b*). If we did that, we'd have to change (c) to (c*): *Justification transfers from an argument's premises to its conclusion when S relies on valid inductive reasoning to infer the conclusion (which is then believed on that basis and justifiedly believed for the first time) from the already justified premises (unless in making that inference S comes to take—or should come to take—her belief in one of the premises or her reliance on that inference to be epistemically inappropriate).* The points I make here in defense of (c) apply, *mutatis mutandis*, to (c*).

Clearly, it's possibly false that, in making the inference mentioned in (b), *S does* take either her premise beliefs or her reliance on the inference to be epistemically inappropriate. The question is whether it's possibly false that, in making the inference mentioned in (b), *S should* take either her premise beliefs or her reliance on the inference to be epistemically inappropriate. Those objecting to my claims about (d)'s possibility must insist that sensible proponents of F will think that, necessarily, in making the inference in (b), *S should* take either her premise beliefs or her reliance on the inference to be epistemically inappropriate.

Consider first the question of whether, in making that inference, she should take her *premise beliefs* to be epistemically inappropriate. It seems not. Why should her recognition of the fact that these premise beliefs validly imply that at least one of those premise beliefs was formed in a reliable way give her a believed defeater for one of those premise beliefs (for which she had no defeater *before* she noticed the implication in question)? That certainly doesn't seem to be a bad-making feature of the premise beliefs.

Consider next the question whether, in making the inference mentioned in (b), *S should* take that very inference to be epistemically inappropriate. Given that it's a valid inference, it is difficult to take seriously the suggestion that she should take it to be epistemically inappropriate. In fact, if you think that, in making the inference mentioned in (b), *S should* take the inference itself to be epistemically inappropriate, then it would seem that your real concern has to do with the validity of the inference and, hence, with the possibility of (b). But since, as I mentioned above in discussing the possibility of (b), it seems possible for there to be an inductively *valid* inference of the sort mentioned in (b), it seems false that, of necessity, one *should* take an inference like that to be epistemically inappropriate. So (d), like (a)–(c), seems to be something a sensible proponent of F will think is possible. And, as I noted earlier, from that we may conclude that a sensible proponent of F will think (a)–(d) are compossible.

2.3 A Dilemma for the 'Epistemic Circularity' Objection to Externalism

So far in Section 2 I've defended the conclusion that F is true and that sensible proponents of F are committed to thinking (a)–(d) are compossible, which, in turn, commits them to approving of epistemic circularity. From this—together with the fact that rejecting F is more objectionable than approving of epistemic circularity—I've concluded that epistemic circularity isn't, in itself, a bad thing and that the 'epistemic circularity' objection to externalism fails. But suppose opponents of epistemic circularity think that my argument for the compossibility of (a)–(d) can be resisted. Will that help the proponent of the 'epistemic circularity' objection to externalism? No. For it seems that the proponent of that objection is faced with the following dilemma.

On the one hand, she can deny that it is sensible to think (a)–(d) are compossible (this will require her to object to my argument given above in defense of that compossibility claim). But if she does that, she will lose her argument that externalists are committed to approving of epistemic circularity. That argument (laid out in Section 1.2) was just a simplified version of my argument in Section 2.2 that proponents of F are committed to approving of epistemic circularity. In arguing that sensible proponents of F will think (a)–(d) are compossible, I was simply tightening up the sort of reasoning on which the proponent of the ‘epistemic circularity’ objection to externalism relied. If my more careful version of that argument fails, then so does their less careful version. If sensible proponents of F can sensibly resist the compossibility of (a)–(d) and, thereby, avoid commitment to approving of epistemic circularity, then externalists can avoid such a commitment too. To avoid this conclusion, one will have to think of an objection to the compossibility of (a)–(d) that has the following two features: (i) sensible proponents of F will accept it and (ii) it can’t be employed by an externalist, without her ceasing to be an externalist.

On the other hand, if an advocate of the ‘epistemic circularity’ objection to externalism thinks it is sensible to accept that (a)–(d) are compossible—perhaps because she agrees that F is true and that sensible proponents of F will think (a)–(d) are compossible—she will be forced herself to approve of epistemic circularity. But then she will have to acknowledge that being committed to approving of epistemic circularity needn’t be a bad thing, in which case such a commitment on the part of externalists needn’t be a cause for concern. In short, it seems that either all who endorse F are committed to approving of some instances of epistemic circularity or externalists aren’t so committed. Either way, the ‘epistemic circularity’ objection to externalism fails.

3 THE SECOND ARGUMENT IN SUPPORT OF EPISTEMIC CIRCULARITY

The first argument in support of epistemic circularity was modeled after the argument, used by objectors to externalism, to demonstrate the externalist’s commitment to approving of epistemic circularity. It concluded that all who support the foundationalist thesis, F, are committed to approving of epistemic circularity. Although I think that first argument is a good argument, I noted at the end of the previous section that whether or not it is a good argument, the externalist’s ‘epistemic circularity’ objection, which relied on the same sort of reasoning, fails. In this section I want to offer a different reason for approving of some instances of epistemic circularity.

This second argument in support of epistemic circularity defends it by attacking the alternatives. We can acknowledge that our initial inclination is to think EC-beliefs *can’t* be justified. But we should reject that initial inclination once we

realize how implausible all possible alternatives to it are. Any intuitions against the view that there can be justified EC-beliefs can be countered by stronger intuitions against the implications of the view that there can't be justified EC-beliefs. Moreover, as I will argue in Section 4 of this chapter, there are good explanations for why it mistakenly seems like epistemic circularity must be a bad thing.

This second argument will actually be defending the conjunction of:

E: EC-beliefs can be justified

with the claim that:

F: There can be noninferentially justified beliefs.

If the conjunction, E & F, is false, at least one of the following four positions must be true:

Alternative 1: $\sim F$

Alternative 2: $\sim E \ \& \ F$ & it's impossible to have justified beliefs in the trustworthiness of *any* of one's belief sources.

Alternative 3: $\sim E \ \& \ F$ & it's possible to have justified beliefs in the trustworthiness of *some* of one's belief sources & it's not possible to have justified beliefs in the trustworthiness of *all* of one's belief sources.

Alternative 4: $\sim E \ \& \ F$ & it's possible to have justified beliefs in the trustworthiness of *all* of one's belief sources.

In Section 2.1 I explained why alternative 1 is to be rejected. In the remainder of this section, I'll explain why alternatives 2–4 are so implausible, thereby defending the conjunction E & F.

Let's begin with alternative 2. It is extremely plausible to think we can come to have a justified belief in the reliability of some scientific instrument. Why, then, would it be impossible for us to come to have a justified belief in the reliability of one of our natural faculties such as hearing? Assuming F is true, it seems that it wouldn't be difficult for me to come to justifiedly believe that my beliefs produced by hearing are reliably formed. I could do this without relying on my hearing at all, perhaps without relying on the hearing of anyone. This seems exactly analogous to learning of the reliability of some scientific instrument. So alternative 2 can be ruled out as implausible in view of fact that it's implausible to jointly affirm its last two conjuncts.

Consider next alternative 3. According to it, although it's possible to have justified beliefs that some of one's faculties are reliable, we can't, if E is false, justifiedly believe of each of our faculties that it is reliable. But why not? Consider a possible cognizer with ten belief sources, each of which could have its reliability established using some of the other nine in the way that we could use our other faculties to confirm the reliability of hearing. If we can justifiedly believe that our hearing is reliable using our other faculties, then what is the problem with this

sort of confirmation happening for more than one faculty? And if there were a possible cognizer who could confirm the reliability of each of its faculties *individually* in this way, why couldn't it thereby come to have justified beliefs about the reliability of *all* of its faculties?

You might think that in doing so, this cognizer would have EC-beliefs, thereby violating our supposition in alternative 3 that E is false. But that isn't true. Consider a case of a cognizer with belief sources X1 and X2. Since we're assuming the truth of F, we can suppose, to simplify the case, that X1 has as an output the noninferentially justified belief that *X2 is reliable* and that X2 has as an output the noninferentially justified belief that *X1 is reliable*. Our question is whether either belief is an EC-belief. It seems not. It's true that if the X2-produced belief that *X1 is reliable* were based on the further belief (produced by X1) that *X2 is reliable*, then we'd have a case of epistemic circularity. For we'd have a case of a belief—namely, the belief that *X1 is reliable*—being such that, in holding it, the cognizer depends on X1 itself.²⁶ But consider the case described at the beginning of this paragraph, where X1 *noninferentially* produces the justified belief in X2's trustworthiness and X2 *noninferentially* produces the justified belief in X1's trustworthiness. In such a case there is no epistemic circularity because neither belief depends on the other and so the cognizer doesn't depend on a belief source to arrive at a justified belief in that source's own trustworthiness.

So the cognizer in our example believes of each of his ten faculties that it is reliable (call these beliefs B1–B10). And, for reasons of the sort noted in the previous paragraph, none of these beliefs is infected with epistemic circularity (i.e. none is an EC-belief). Moreover, B1 is held in such a way that if the other nine weren't held, B1 would be a justified belief. The same is true of each of the other beliefs (B2–B10): each is held in such a way that if the other nine weren't held, it would be justified. So why would the fact that the other nine beliefs are held prevent B1 from being justified? B1 has the very same basis it would have had if the other nine weren't held. And none of those other nine, or even all of them together, seem, of necessity, to constitute an actual defeater for B1. I conclude, therefore, that alternative 3 is no less objectionable than alternative 2. Just as it's implausible to jointly affirm the last two conjuncts of alternative 2, it seems also to be implausible to jointly affirm the final three conjuncts of alternative 3.

Consider finally alternative 4. I've just argued that having belief sources, each of which is confirmed using some of the others, needn't result in any EC-beliefs. That is, I've just argued that the first conjunct of alternative 4 (i.e. $\sim E$) doesn't explicitly contradict the joint truth of its last two conjuncts. So why not take alternative 4? The problem with this alternative is that, although the first conjunct doesn't explicitly contradict the joint truth of its final two conjuncts, it

²⁶ Recall that I said earlier that to depend upon a belief source X in holding a belief B is for B either to be an output of X or to be held on the basis of an *actually employed* inference chain leading back to an output of X.

seems implausible to affirm the first conjunct while also jointly affirming the last two. We can see this as follows. According to alternative 4, it's perfectly acceptable to say (in light of its last two conjuncts and the example two paragraphs back) that:

P1: It's possible to depend on X1 to justifiedly believe in the reliability of the source by which you justifiedly believe *X1 is reliable*.

However, given that E is false according to the first conjunct, alternative 4 also tells us that it is *unacceptable* to say that:

P2: It's possible to depend on X1 to justifiedly believe that X1 is reliable.

The problem is that P1 and P2, though different, seem so close to being equally plausible that anyone who balks at P2 should, it seems, balk at P1 as well. Likewise, anyone willing to accept P1 should be equally willing to accept P2. It just doesn't seem defensible to accept one and not the other. So alternative 4 should be rejected, along with alternatives 2 and 3; each of them seems to be less plausible than E.²⁷

The argument above can be restated as follows. Consider the following series of claims:

- (i) One can justifiedly believe (without epistemic circularity) in the reliability of a scientific instrument.
- (ii) One can justifiedly believe (without epistemic circularity) in the reliability of some of one's belief sources.
- (iii) One can justifiedly believe (without epistemic circularity) in the reliability of all of one's belief sources.
- (iv) One can have justified EC-beliefs.

I've argued that, for a supporter of F, each of the following responses should seem implausible:

²⁷ Remarks similar to these apply to Stewart Cohen's position on epistemic circularity. Cohen thinks it is fine to say P3: *It's possible to know via source X, without first verifying X's reliability, that X is reliable (so long as this is among the first things known via X)*. Yet he rejects the claim that P4: *It's possible to know via source X, without first verifying X's reliability, that X is reliable (even if this is not among the first things known via X)*. We see his acceptance of P3 in his suggestion that 'we can know on the basis of holistic support that holistic support is reliable' (Cohen 2002: 323). And we see his rejection of P4 in his endorsement of a principle he calls 'KR' according to which 'A potential knowledge source K can yield knowledge for S only if S knows [either first or at the same time] that K is reliable' (Cohen 2002: 309). But it seems that anyone who balks at P4 should also balk at P3; and anyone who finds P3 acceptable should find P4 acceptable too. For P3, like P4, permits one to use a source X in coming to know X is reliable. And P3, like P4, permits one to gain knowledge via a source X without first verifying X's reliability (and without depending upon the knowledge that X is reliable). The only difference between the two is that according to P4, the proposition that *X is reliable* needn't be among the first things known via the so far unverified X, whereas according to P3, it must be.

Response 1: reject (i).

Response 2: accept (i) without accepting (ii).

Response 3: accept (ii) without accepting (iii).

Response 4: accept (iii) without accepting (iv).

Given the plausibility of F and that each of those four responses (as well as their disjunction) seems less plausible than accepting (iv), we may conclude that epistemic circularity needn't be a bad thing.

In Sections 2 and 3 I've argued that all proponents of F (whether externalist or nonexternalist) should acknowledge that EC-beliefs can be justified. This shows that we can't use the 'Epistemic Circularity' Objection to externalism as a motivation for endorsing internalism. For, first of all, nonexternalist positions endorsing F are themselves committed to permitting epistemic circularity in some circumstances. So the alleged problem can't be avoided by giving up on externalism. Secondly, given that supporters of F are committed to admitting justified EC-beliefs and that the alternatives to admitting justified EC-beliefs are more unpalatable than admitting them, we've got good reason to think it isn't a defect of a position that it is committed to admitting that there could be justified EC-beliefs. Hence, this supposed motivation for internalism (that it enables one to avoid the 'Epistemic Circularity' Objection to externalism), isn't an adequate one either.

4 WHY EPISTEMIC CIRCULARITY SEEMS LIKE A BAD THING WHEN IT ISN'T

I've argued in the previous two sections that epistemic circularity needn't be a bad thing. But if that conclusion is right, what should we say about the cases of Juror #1 and Doubter? Surely their beliefs were unjustified. And surely that had to do with the fact that they were EC-beliefs.²⁸ So how can I defend epistemic circularity when, in light of those examples, it seems like such a bad thing?

In this section I will explain why it (misleadingly) *seems* that epistemic circularity is, in itself, a bad thing.²⁹ The explanation comes down to this: there are some situations (e.g. Doubter's situation) in which EC-infection is malignant; and, for a variety of reasons, when we think about epistemic circularity, we are inclined to think about its occurrence in such situations. I'll begin, in Section 4.1, by saying what makes the difference between malignant and benign

²⁸ Or at least this is so in Doubter's case. Strictly speaking, the belief of Juror #1 isn't an EC-belief since it isn't a belief in the trustworthiness of one of the subject's own belief sources.

²⁹ For another account, similar in spirit to my own, of why epistemically circular arguments seem defective even though they aren't, see Pryor (2004).

epistemic circularity. Then, in Section 4.2, I'll explain why we tend to focus only on cases of malignant epistemic circularity, ignoring cases of benign epistemic circularity.

4.1 Malignant and Benign Epistemic Circularity

The sort of belief that can get infected with epistemic circularity is a belief such as that one's belief source, X, is trustworthy or that one's belief, B, wasn't formed in an unreliable way. Now there are two kinds of situation in which a person can form EC-infected beliefs about a source X or a belief B:

QD-situations: Situations where, prior to the EC-belief's formation, the subject *is or should be* seriously questioning or doubting the trustworthiness of X or the reliability of B's formation.³⁰

Non-QD-Situations: Situations where, prior to the EC-belief's formation, the subject *neither is nor should be* seriously questioning or doubting the trustworthiness of X or the reliability of B's formation.

To *seriously* question or doubt the trustworthiness or reliability of something is to question or doubt it to the point where one withholds³¹ or disbelieves the claim that the thing is trustworthy or reliable.

Here's an example of an EC-belief formed in a QD-situation. Suppose that Tom—who has recently been persuaded, by some skeptical argument, to have serious questions about the reliability of his sense perception—is considering an argument that has been proposed to convince him that his perception is trustworthy after all. And suppose that he can see that the argument is epistemically circular. He will, if he's sensible, consider this argument to be useless as a means to help him *regain lost confidence* in his perception. The reason is simple. Tom doesn't trust his perception. He is uncertain about whether it is reliable. So long as that is the case, it wouldn't be reasonable for him to depend on the testimony of perception to learn things. What he should be looking for is some *other* testimony—some testimony that is *independent* of perception—that supports the reliability of his sense perception. This is just the sort of context that Doubter is in, which is why we conclude that epistemic circularity is a bad thing in his case.

³⁰ In Bergmann (2004a), I focused on questioned-source contexts rather than on QD-situations. The difference is that questioned-source contexts are ones where the source's trustworthiness is in fact questioned whereas a QD-situation is one where a source's trustworthiness *or the reliability of a particular belief's formation* either is *or should be* questioned or doubted. So although questioned-source contexts are QD-situations, not all QD-situations need be questioned-source contexts. Consequently, the claim that EC-infection is malignant in all QD-situations is different from the claim that it is malignant in all questioned-source contexts. This makes my proposal here (concerning when EC-infection is malignant) slightly different from what it was in Bergmann (2004a).

³¹ To withhold a proposition is, roughly, to consider it and to (successfully) resist both believing it and disbelieving it. See Bergmann (2005: 421) for further discussion of withholding.

But not all situations in which EC-beliefs are formed are like that. Consider the following example of a *non*-QD-situation. Becky has never had any questions or doubts at all about the trustworthiness of her sense perception. In fact, she never considered the proposition that her sense perception is reliable before she came to believe, on this occasion, that her sense perception *is* reliable. But although Becky has no questions or doubts about the reliability of her sense perception, she is curious about how she came to believe in its reliability. She is wondering about the origin of this belief, not its legitimacy (about which, as I've already indicated, she has no doubt or uncertainty). Now suppose she discovers that she formed her belief about the reliability of her senses in a way that involved epistemic circularity. Since she wasn't trying to allay her worries about this source by looking for some *independent* verification of its reliability, this discovery doesn't give rise to any doubts or questions in her mind concerning the reliability of her senses. Becky was merely curious about how it was that she came to hold the belief that her sense perception is reliable—a belief about whose credentials she has no questions or doubts. If it's not the case that this discovery *should* make Becky have serious questions or doubts about its reliability, then this is a non-QD-situation—a situation in which EC-infection *doesn't* seem to be a bad thing.³²

My initial rough-and-ready proposal, then, is that EC-infection is malignant in QD-situations and that it is benign in non-QD-situations. What explains this difference? The answer comes in two parts, for there are two kinds of QD-situations: those in which the subject *does* seriously question or doubt the trustworthiness of a belief's source or the reliability of its formation, and those in which she doesn't but *should*. Consider, first, cases of the first sort, where the subject *does* seriously question or doubt. In a QD-situation of that sort, an EC-belief in a source X's trustworthiness is produced, at least in part, by a source (i.e. X) whose trustworthiness the subject *does* seriously question or doubt. And typically, due to the fact that she seriously questions or doubts X's trustworthiness, she has a believed defeater of the undercutting sort for all her beliefs produced (even in part) by source X, including the belief that X is a trustworthy source. Because of clause (i) of J_{PF} (discussed in Chapters 5 and 6), this believed defeater keeps the belief in X's trustworthiness from being justified.

Now consider QD-situations of the second sort, where the subject doesn't doubt or question but she *should*. In such a situation, the subject *should* (due to what is required by the proper functioning of her defeater systems) have a believed defeater for that belief. Hence, she avoids having a believed defeater for her EC-infected belief only because her cognitive faculties (in particular, her defeater systems) aren't functioning properly in the formation of that belief. In this case too, the EC-infected belief isn't justified, this time because the proper function condition—clause (ii) of J_{PF}—isn't satisfied.

³² The antecedent of this conditional will be resisted by some readers. I will address that resistance below in Section 4.2.2. For the moment, I will suppose that the antecedent is true.

Now that we've seen the connection between QD-situations and believed defeaters, we can replace the initial rough-and-ready distinction between malignant and benign EC-infection (given in terms of QD-situations) with something more careful and precise in terms of believed defeaters. EC-infection is malignant in situations where it either does or should give rise to a believed defeater for the belief so infected. EC-infection is benign in situations where it neither does nor should give rise to a believed defeater for the belief so infected. This gives us approximately the same result as we get when we characterize EC-infection as malignant in QD-situations and benign in non-QD-situations. For this reason, and because the distinction between malignant and benign EC-infection in terms of QD-situations is, for certain purposes, more illuminating, I will use the two methods of distinguishing malignant and benign EC-infection interchangeably.

We can use this distinction (in terms of QD-situations) between malignant and benign epistemic circularity to explain why people mistakenly think epistemic circularity is, in itself, a bad thing. In QD-situations, epistemic circularity *is* a bad thing; and when people think about epistemic circularity, they tend to think about it only in QD-situations, ignoring cases of epistemic circularity in non-QD-situations. What remains to be explained is why people tend to ignore cases of EC-beliefs in non-QD-situations.

4.2 Why Benign Epistemic Circularity is Ignored

There are at least three reasons why cases of benign epistemic circularity, involving EC-beliefs in non-QD-situations, are ignored. First, philosophers don't find such cases interesting. Second, it's difficult to imagine the most commonly discussed sort of EC-belief—namely, EC-belief produced by an epistemically circular track record argument—occurring in a non-QD-situation. Third, EC-beliefs in non-QD-situations seem, to some people, to be impossible (in which case ignoring them would be ignoring what is impossible). Let's consider each of these in turn.

4.2.1 *The First Two Reasons: Philosophical Interest and Track Record Arguments*

The first reason has to do with what philosophers find interesting. In most discussions in the literature concerning beliefs in a source's trustworthiness, philosophers have in mind situations in which the believer has serious questions or doubts about the source's trustworthiness and is trying to resolve them.³³ Juror #1 and Doubter are both in situations like this. And even if the philosophers discussing these cases have no doubts about the trustworthiness of sense perception, when they are discussing how one can justifiedly come to believe that sense perception

³³ The cases in Cohen (2002) involving conversations with his son who expresses uncertainty about the trustworthiness of a belief source are typical.

is reliable they are interested in ways a person who *does* seriously question or doubt the reliability of perception can come to hold such a belief. Perhaps this is because they are seeking to identify a way that *anyone*— even someone with serious questions or doubts about the trustworthiness of perception—can sensibly come to believe in its trustworthiness. Perhaps it just doesn't seem very interesting to most epistemologists to explain how a person with no doubts at all about the trustworthiness of sense perception can come to justifiedly believe that her perception is reliable. At any rate, because such discussions almost always focus on situations in which the believer has these sorts of questions or doubts, there is a strong inclination to think epistemic circularity is a bad thing. And since epistemic circularity *is* a bad thing in such contexts, that inclination is understandable. But, of course, the fact that philosophers aren't typically *interested* in EC-beliefs in non-QD-situations doesn't show that there couldn't be any.

The second reason people ignore benign epistemic circularity is the tendency, when thinking about EC-beliefs, to focus on the ones resulting from track record arguments. The problem with focusing on EC-beliefs of this sort is that they are believed on the basis of *arguments*. And an argument is typically evaluated in terms of how effective it would be at convincing someone who initially questions or doubts its conclusion. An epistemically circular track record argument fails abysmally by that standard. It is of no use whatsoever to anyone who begins with serious questions or doubts about its conclusion. For this reason, epistemically circular arguments (including epistemically circular track record arguments) seem to be pathetic arguments.³⁴ And it is difficult to imagine a context in which it would be a good thing to depend upon a pathetic argument.³⁵

³⁴ Although he defends epistemic circularity, Alston (1993b: 15–17) seems to feel uncomfortable with the suggestion that we know our faculties are reliable via an epistemically circular track-record argument. Fumerton (1995: 178–80) picks up on this and takes it as evidence that externalists tend to feel uncomfortable with the epistemic circularity to which their views force them. But, in light of my comments about epistemic circularity being malignant in QD-situations, we can think of Alston's discomfort as simply a discomfort with reliance on epistemically circular *arguments*. We have this discomfort because of our understandable tendency to assess the value of arguments based on how they fare in QD-situations, where epistemically circular arguments fare poorly. See Section 5 of this chapter for an account of how we can justifiably believe our faculties are reliable without relying on arguments of any kind. And see Section 3.1 of Chapter 8 for further discussion of Fumerton's critique of Alston.

³⁵ Notice that if you thought that you had to add—to the original premises of an epistemically circular argument—further premises stating that the sources of your beliefs in the original premises are trustworthy, then, because the original argument is epistemically circular, you'd have to add the *conclusion* as one of the premises. This would transform the epistemically circular argument under discussion into a logically circular argument. This provides another possible explanation for why people are inclined to think epistemic circularity is a bad thing. For they may be inclined to think there is some pressure to add (or to be able to add) to the premises of an epistemically circular argument further premises about the trustworthiness of your belief sources. Since doing so turns an epistemically circular argument into a logically circular argument (and since it seems that a belief cannot be justified in virtue of its being based on a logically circular argument), this might incline such people to think epistemic circularity is, in itself, a bad thing.

But it is important to distinguish two kinds of case: those in which an argument is thought of as an instrument for persuading people who question its conclusion and those in which someone just happens to notice what her justified beliefs entail and, as a result, infers a conclusion about which she has never had any questions or doubts. An argument might be useless in cases of the first kind (i.e. convincing those who question the conclusion) while at the same time being completely effective in producing justified beliefs in cases of the second kind (where there is no serious doubt—in effect, no believed defeater to overcome). Although arguments in cases of the first sort are interesting only in QD-situations, arguments in cases of the second kind can (even if epistemically circular) be interesting in non-QD-situations.

4.2.2 *The Third Reason: The Seeming Impossibility of EC-Beliefs in Non-QD-Situations*

Let's turn now to the last of the three reasons why cases of benign epistemic circularity are ignored. In Section 4.1, I said the case of Becky's belief was an example of an EC-infected belief in a non-QD-situation. But is it really a non-QD-situation—a situation in which Becky neither does nor should seriously question or doubt the trustworthiness of the source of the EC-belief in question (i.e. that her perception is reliable)? It's clearly false that she *does* question or doubt the reliability of her sense perception. (I stipulated that she had no questions or doubts about the trustworthiness of her perception, even though she recognized that her belief in its reliability was formed in an epistemically circular way. Since this is clearly possible, it's perfectly appropriate to make this stipulation.) But is it also false that, in such a situation, she *should* question or doubt the reliability of her perception? Objectors might say this isn't false. They might insist instead on the following 'should' claim:

S1: It is necessarily the case that a person with an EC-infected belief *should* seriously question or doubt the trustworthiness of its source.

If S1 is true, then it is impossible to have EC-beliefs in non-QD-situations. In the remainder of this subsection, I'll evaluate and reject some reasons for thinking S1 is true. The appeal of these reasons to some people helps to explain why S1 *seems* to be true and, therefore, why it *seems* impossible to have EC-beliefs in non-QD-situations. My criticisms of these reasons explain why these seemings are misleading: they're misleading because, despite initial appearances to the contrary, the reasons in question aren't good reasons. It is, therefore, a mistake—though perhaps an understandable mistake—to ignore possible examples of EC-beliefs in non-QD-situations on the grounds that such examples aren't possible.

The first reason I want to consider for thinking S1 is true is that it is entailed by the following 'should' claim, which some people might find attractive:

S2: It is necessarily the case that, for every belief a person has, she *should* seriously question or doubt its source.

It shouldn't be surprising that some philosophers are tempted by S2. Many philosophers (perhaps traditional epistemologists in particular) take it to be their professional duty to question or doubt the sources of their beliefs. To do otherwise is, they think, to live the unexamined life, something not worth living. Thus, they take the philosophical technique of asking difficult questions about things taken for granted and exalt it to the status of an epistemic obligation that applies to all beliefs of all believers. This, apparently, is how they arrive at S2.

But the problem with S2 is that, if it were true, complete skepticism would follow. For S2 says we should have serious questions or doubts about *all* our belief sources. But that would prevent any of our beliefs from being justified. I've already explained, in Section 4.1, why having such questions or doubts about a source's trustworthiness prevents beliefs produced by that source from being justified (until the reliability of the source is independently verified). Here I want to point out that a person who does what S2 recommends couldn't escape her plight of having only unjustified beliefs by appealing to further beliefs verifying the reliability of her questioned or doubted belief sources. For these additional beliefs would be helpful only if they were justified. But none of them would be justified given that, according to S2, we *should* have doubts about the trustworthiness of their sources too. The problem is that if we seriously question all of our belief sources, we have no way to independently verify any of them. Because this extreme skeptical result is so implausible, I conclude that S2 is implausible and, therefore, of no help in defending S1.³⁶

Let's focus a bit more narrowly than S2 does. Consider these two kinds of beliefs:

Type (i): beliefs which are such that those holding them have considered the proposition that their source is reliable.

Type (ii): beliefs which are such that those holding them haven't yet considered the proposition that their source is reliable.

³⁶ It is sometimes suggested that we can conclude that our faculties are reliable by appealing to an inference to the best explanation: the idea is that the best explanation of their *apparent* reliability is their *actual* reliability. But the source of this belief that our faculties are reliable is inference to the best explanation. And according to S2, we should seriously question or doubt the reliability of *that* way of forming beliefs too, thereby preventing beliefs produced by it from being justified. In addition, all inferential beliefs, including those produced by inference to the best explanation, are justified only if the beliefs on which they are based are justified. Thus, in the case at hand, one would need justified beliefs in the apparent reliability of one's faculties in order for the inferred belief about actual reliability to be justified. The problem, of course, is that the belief about apparent reliability has a source as well and, according to S2, that source too should be seriously questioned or doubted. But if it is, then beliefs produced by it won't be justified, in which case the belief inferred via inference to the best explanation won't be justified either.

Proponents of F will agree that beliefs of type (ii) can be justified noninferentially. That is, they will agree that a person's belief source can have a justified noninferential output even if that person has never considered any proposition about the reliability of that belief source. But if a person can have a noninferentially justified type (ii) belief, then it is *false* that, for all beliefs of type (ii), one *should* have serious questions or doubts about their reliability. For if the beliefs are of type (ii), then the subject doesn't have serious questions or doubts about their reliability (since she has never even considered the proposition that their source is reliable). And if it were the case that a person should have such doubts but doesn't, then her beliefs produced by that source wouldn't be justified after all. So proponents of F (a view defended above in Section 2.1) will deny that, for all type (ii) beliefs, we should have serious questions or doubts about their reliability.

Now what about beliefs of type (i)? Is it true that for any belief of that type, one *should* have serious questions or doubts about its reliability? Suppose the correct answer to this question were 'yes'. Then if a belief were to change from being a type (ii) belief to being a type (i) belief, it would automatically become (if it weren't already) a belief whose reliability one should seriously question or doubt. In other words, if one should have serious questions or doubts about beliefs of type (i)—even though it's false that one should have serious doubts or questions about all beliefs of type (ii)—then the following 'should' claim would be true:

S3: It is necessarily the case that the mere fact that one considers the proposition that one's belief source is reliable makes it the case that one *should* seriously question or doubt the truth of that proposition.

If S3 were true, then S1 would be true too.³⁷ So perhaps the reason S1 seems appealing to some people is that they find S3, which entails S1, appealing. This is the second reason I want to consider for thinking S1 is true.

The problem is that on reflection S3 seems false. For contrary to what S3 says, it seems implausible to think that it is fine *not* to have serious questions or doubts about a belief source when one hasn't even considered the issue of its reliability but that, as soon as one considers the issue, one *should* have serious questions or doubts. Why think that merely considering that proposition makes it the case that one should seriously question or doubt it whereas (as I argued two paragraphs back) prior to considering it there was no such obligation? In cases where no questions or doubts about a source's trustworthiness have been brought to

³⁷ Proof: S1 says we should question or doubt the reliability of the source of EC-beliefs. S3 says we should question or doubt the reliability of a belief source if we have considered whether it is reliable. So if the following conditional is true—if *B* is an EC-belief, then the person holding *B* has considered whether its source is reliable—then S3 entails S1. And that conditional is true. For if a person has an EC-belief, she will believe that one of her belief sources is reliable (since those are the sorts of beliefs that can be infected by epistemic circularity). In fact, since that belief is an EC-belief, she will believe that *its own source* (at least one component of it) is reliable. Thus, she will have considered the proposition that the source of that very belief is reliable.

the person's attention—cases in which the person believes in the source's trustworthiness immediately upon considering it—this suggestion seems especially implausible. So S3 won't be useful either in defending S1.

The third reason for S1 that I want to look at focuses on something more narrow even than S3. It employs the following 'should' claim to single out only recognized EC-beliefs for consideration, rather than *all* beliefs or all type (i) beliefs:

S4: It is necessarily the case that realizing that one of your beliefs is an EC-belief makes it the case that you *should* seriously question or doubt the trustworthiness of its source.

As we noted in rejecting S3, merely considering whether a source is reliable doesn't make it the case that you should seriously question or doubt its trustworthiness. So suppose you consider whether a source is reliable and that, upon doing so, you don't question or doubt its trustworthiness in the least; instead you believe confidently in its trustworthiness and reliability. So far there is no reason to think you *should* be seriously questioning or doubting that source's trustworthiness. But then suppose that you notice that this source confirms its own reliability. Should that make you seriously doubt or question the trustworthiness of that source? It's hard to see why. That it confirms itself isn't in any way indicative of its untrustworthiness. Self-confirmation isn't a vice of a belief source.

But suppose you go on to notice that *your* belief in this source's trustworthiness is produced, at least in part, by the source itself. Should *that* make you seriously doubt or question its trustworthiness? If so, why? It's true that if you *already had* serious doubts or questions about that source's trustworthiness, then you'd be troubled by the fact that your belief in its trustworthiness was produced by that very source. You'd be concerned because, in having serious doubts about that source, you'd have a believed defeater for all beliefs you recognize as ones produced by that source, including the belief that the source is reliable. But if you (appropriately) *didn't already* have such doubts or questions, why should you now all of a sudden begin having them upon recognizing that your belief in the source's trustworthiness is produced, at least in part, by the source itself?

Perhaps the thought is this: If you recognize that your belief is an EC-belief, you can see that, without gaining *independent* support for its reliability from other belief sources (something it can't do while remaining an EC-belief), its justification won't survive the onslaught of serious questions or doubts about its reliability. And the concern is that that sort of vulnerability is a bad thing—so bad that beliefs vulnerable in this way can't be justified. Let's call this 'the vulnerability charge' against EC-beliefs.

I agree that our EC-beliefs have that sort of vulnerability and that it seems like a bad thing for them to be vulnerable in that way. I can even see how these facts might tempt one to think that EC-beliefs can't be justified. But although

I understand this temptation, there are two reasons why I think it would be a mistake to give in to it. First, consider again my response above to S2. There I pointed out that if we seriously question all of our belief sources, we have no way to independently verify any of them and skepticism follows. This suggests that our belief sources collectively face a vulnerability charge that is very much like the vulnerability charge just considered against EC-beliefs: justification is lost if questions or doubts of a certain kind arise. And given that the vulnerability afflicting our belief sources collectively doesn't prevent the beliefs they produce from being justified (if it did, extreme skepticism would follow), that suggests that the similar vulnerability afflicting our EC-beliefs needn't prevent them from being justified either.

The second reason for rejecting the vulnerability charge against EC-beliefs emerges when we consider what can be said for and against epistemic circularity. Against the suggestion that EC-beliefs can be justified is the fact that they have the undesirable feature of being vulnerable in the way just described. But in support of the view that EC-beliefs can be justified, we have the arguments given in Sections 2 and 3 of this chapter. And those considerations in support of the possibility of benign epistemic circularity seem to me to outweigh the considerations just given against its possibility. In particular, I find it more plausible to think that a belief with the vulnerability described above can be justified than to endorse any of the four alternatives to that suggestion mentioned in Section 3 where I laid out the second argument in support of epistemic circularity. Being *vulnerable* to loss of justification is a bad thing; but it isn't sufficient by itself for *actual* loss of justification.

I conclude, therefore, that S4 isn't plausible either. This (together with my reasons given above for finding S2 and S3 implausible) gives us a good reason for rejecting S1. But then it seems plausible to think that there *could* be EC-beliefs in non-QD-situations. And since there could be EC-beliefs in non-QD-situations, it is a mistake—even if it is an understandable mistake—to ignore them and to pass judgment against epistemic circularity based on its obvious badness in QD-situations.

5 REALISTIC BENIGN EPISTEMIC CIRCULARITY

The chapter so far has dealt largely with necessities and possibilities. I've argued that it's *possible* for there to be justified EC-beliefs. So epistemic circularity isn't *necessarily* a bad thing. But it would be nice to step out of the realm of mere possibility and have an account of some realistic-sounding case of benign epistemic circularity. EC-beliefs which are produced by epistemically circular track-record arguments in non-QD-situations may be possible cases of benign epistemic circularity. But they still don't seem like realistic cases. We just don't seem to form our beliefs in the trustworthiness of our senses using such epistemically circular

track-record arguments. Perhaps this is because we find it difficult to form beliefs based on arguments without any concern for the value of the argument when its conclusion is questioned or doubted (and, as noted above, epistemically circular arguments are worthless if their conclusions are questioned or doubted). Whatever the reason, the examples of epistemically circular track-record arguments don't seem like realistic accounts of how we come to form beliefs in the reliability of our faculties. So how do we come to form beliefs in the reliability of our faculties?

Before answering that question, it's important to emphasize that proponents of F will think that we don't *need* to form those higher-level beliefs in the reliability of our faculties in order for the outputs of those faculties to be justified. Nevertheless, many people with leisure for philosophical reflection—and this includes nonphilosophers too—do believe that their faculties are reliable. So although we don't need to form such beliefs, we often do. And what we're interested in is how we come to form such beliefs. In answering that question, I'll be endorsing a Reidian position according to which many of our beliefs in the reliability of our faculties are *noninferential* beliefs that are justified because they satisfy the conditions of justification mentioned in J_{PF}.³⁸ And in the process of explaining that Reidian position, I'll offer a realistic account of how benign epistemic circularity can enter into our noetic structure, without the use of any epistemically circular track-record arguments.

According to Reid, it is a first principle that our faculties are reliable.³⁹ And first principles, says Reid, are properly believed *noninferentially* (Reid [1785] 2002: 452). Just as we have noninferential knowledge about our immediate physical environment by means of sense perception and about our past by means of memory and about our own minds by means of introspection, so also we have a faculty by means of which we have noninferential knowledge of first principles.⁴⁰ Reid thinks of first principles as self-evident truths. He thinks some are contingent and some are necessary. The one mentioned above (concerning the reliability of our natural faculties) is contingent. And the faculty by which we know these first principles (whether necessary or contingent) he calls 'common sense'. Thus, Reid, as I understand him, disagrees with Alston's (1993b) conclusion that one can't know that sense perception is reliable without relying on sense

³⁸ I am more interested in developing a line of thought that seems to me to be suggested by Reid than in defending my belief that Reid intended to propose it. I'd be attracted to the ideas that I see in Reid even if I became convinced that I haven't understood him correctly.

³⁹ 'Another first principle is, that the natural faculties [e.g. sense perception, memory, introspection, etc.], by which we distinguish truth from error are not fallacious' (Reid [1785] 2002: 480).

⁴⁰ 'We ascribe to reason two offices, or two degrees. The first is to judge of things self-evident; the second to draw conclusions that are not self-evident from those that are. The first of these is the province, and the sole province, of common sense . . . and is only another name for one branch or degree of reason' (Reid [1785] 2002: 433).

perception. According to Reid, one *can* know directly and noninferentially, via the faculty of common sense, that sense perception is reliable.⁴¹

It is important not to be put off by Reid's name—i.e. 'common sense'—for the faculty by which we know first principles. We tend to classify as 'common sense beliefs' beliefs that are peculiar to our own culture or upbringing. Reid doesn't (or at least he doesn't want to). His intention is to include only propositions that almost everyone believes (and knows) noninferentially—things that are immediately accepted by sane persons once considered and understood. That $2 + 2 = 4$, that *modus ponens* is a valid form of inference, that the thoughts of which I am conscious are my thoughts, that I have some degree of control over my actions—these are examples of what Reid considers the dictates of common sense. The first two are examples of necessary truths known by common sense; the latter two are contingent truths. A more familiar name for the faculty by which we have noninferential knowledge of necessary truths is *a priori intuition*. So the branch of reason Reid calls 'the faculty of common sense' encompasses both what we call 'a priori intuition' and something akin to it that produces beliefs in contingent rather than necessary truths.

How exactly does this faculty of common sense work? What is the process by which it leads us to beliefs in first principles? Sense perception seems to work as follows: we have sensory experiences and, on the basis of such experiential evidence, we form noninferential perceptual beliefs. Something similar can be said about the faculty of common sense:

We may observe, that opinions which contradict first principles are distinguished from other errors by this; that they are not only false, but absurd: and, to discountenance absurdity, nature has given us a particular emotion, to wit, that of ridicule, which seems intended for this very purpose of putting out of countenance what is absurd, either in opinion or practice. (Reid [1785] 2002: 462)

The idea is that when we entertain the contrary of a first principle, we experience what Reid calls 'the emotion of ridicule'. On the basis of this experience we do two things: we dismiss as absurd the contrary of the first principle and we believe the first principle itself. Thus, noninferential common sense beliefs, like noninferential perceptual beliefs, are based on experiential evidence.

According to Reid, how is it that we come to know that our faculty of common sense is reliable? Recall that the first principle Reid mentioned earlier was that *all* our natural faculties are reliable; this includes the faculty of common sense itself. Thus, as I understand him, Reid thinks that one noninferential output of the faculty of common sense is the experience-based belief that the faculty of

⁴¹ However, as we shall see below, Reid agrees with Alston's other conclusion, namely, that in knowing that all our faculties are reliable, we eventually get involved in epistemic circularity of some sort.

common sense itself is reliable.⁴² This is where the epistemic circularity enters Reid's account. And so long as one doesn't question or doubt one's faculty of common sense in forming the belief that common sense itself is reliable, one is in a non-QD-situation and the epistemic circularity is benign. Moreover, if that belief that common sense is reliable is produced by properly functioning faculties and one has no believed defeater for it, then, according to J_{PF} which I defended in Chapters 5 and 6, it is a justified EC-belief.

I've offered this Reidian account of how we know our faculties are reliable as an account that is more realistic than accounts according to which we know this via a track-record argument. It's also more realistic than accounts by philosophers such as Bonjour and Feldman according to which we know our faculties are reliable via an *argument* of some sort.⁴³ My account is more realistic than these for two reasons. First, the suggestion that our EC-beliefs in the trustworthiness of our belief sources are formed noninferentially is much more in accord with what actually seems to take place than is the suggestion that such beliefs are based on arguments. We don't seem to formulate to ourselves arguments that our faculties are reliable. Rather, we seem just to take it as obvious (without inference) that our faculties are reliable. Second, if these arguments are epistemically circular (as they will be if they are to yield belief in the reliability of *all* our faculties), then we will be inclined to offer a dismal evaluation of these arguments. For such arguments are worthless in QD-situations where their conclusions are questioned or doubted, and we tend to evaluate arguments in light of how they fare in QD-situations.⁴⁴

But isn't my account in terms of a faculty of common sense even more unrealistic? How can we take seriously this Reidian alternative to the argument-based accounts of how we know our faculties are reliable? It's important to recognize that we can take this suggestion seriously without defending all of Reid's views on common sense. One certainly doesn't need to agree with Reid about which propositions are first principles; for our purposes here, what matters is that we justifiably believe in the reliability of our faculties in something like the way Reid suggests. Nor does one need to hold Reid's views on the details of how one comes to believe in the reliability of our faculties—details concerning whether we have a faculty of common sense employing the emotion of ridicule. All we need, in order to have a sensible alternative to the argument-based accounts of how we

⁴² See Reid ([1785] 2002: 481) where he says: 'How then come we to be assured of this fundamental truth on which all others rest [i.e. the truth that our natural faculties are not fallacious]? Perhaps evidence, as in many other respects it resembles light, so in this also, that as light, which is the discoverer of all visible objects, discovers itself at the same time: so evidence, which is the voucher for all truth, vouches for itself at the same time'.

⁴³ See Bonjour and Sosa (2003: 83–96) and Feldman (2003: 148–52).

⁴⁴ Another problem with the abductive inference method of believing our faculties are reliable (endorsed by Bonjour and Feldman) is that it is doubtful that a nonskeptical position really is a better explanation than its skeptical rivals. See Alston (1993b: ch. 4).

know our faculties are reliable, is a defense of the claim that *we noninferentially and justifiedly believe, perhaps on the basis of some sort of nonpropositional evidence, that our faculties are reliable*. The question, then, is whether this italicized claim can be taken seriously.

As a matter of fact, it *is* taken seriously. The particularist approach to epistemology is currently quite popular among analytic philosophers. Those who employ it rely heavily on the noninferential knowledge they have of Moorean truths—truths such as that we aren't the victims of massive deception about the external world or about the past and that we aren't brains in vats. According to the particularist methodology, accounts of justification according to which it turns out that we aren't justified in believing such Moorean truths are, thereby, disqualified. The question of *how* it is that such epistemologists know these Moorean truths isn't very often addressed. But it seems that any answer given will be something along the lines suggested by Reid's account sketched above.

Furthermore, a defense of a Reidian account could be developed along the following lines. First one could point out that sense perceptual beliefs based on sense experiences can be justified despite the fact that we lack compelling noncircular inductive, abductive, or deductive arguments from the existence of such experiences to the truth of the beliefs they ground.⁴⁵ Then one could point out that a priori knowledge also involves belief processes in which a belief is based on a certain sort of seeming—a seeming which is an experience of some sort.⁴⁶ In this case too it looks like there is no deductive, abductive, or inductive argument from the existence of such an experiential ground to the truth of the a priori belief based on it. But this needn't lead us to seriously question or doubt the justification of our a priori beliefs any more than a similar concern forces us to seriously question or doubt the justification of our sense perceptual beliefs. In each case (sense perceptual and a priori) the belief is noninferentially justified as a result of its being based on the experiential ground in question.

Upon recognizing that the above conclusions are plausible, one could then argue that commonsense belief in contingent truths is very much like a priori belief insofar as they each have the same sort of experiential ground (i.e. a certain sort of seeming). Because it is plausible to take seriously both the existence of justified a priori beliefs as well as the account of them as experience-based, it is also plausible to take seriously both the existence of justified commonsense beliefs in contingent truths as well as an account of them as experience-based. Our starting point is the fact that we *do* seem to have justified noninferential beliefs of each kind despite the fact that in each case the belief in question is based on a ground from whose existence we can't deductively, abductively, or inductively infer the

⁴⁵ See Alston (1993b) for an extended critique of various attempts to show that there is some such connection between sense experiences and the beliefs they ground.

⁴⁶ See Plantinga (1993a: ch. 6). Because the experiential ground of a priori beliefs is nonempirical the beliefs are still properly called 'a priori' rather than 'a posteriori'.

truth of the belief it grounds. And my suggestion is that since many philosophers are inclined to express very little resistance to an account of justified sense perceptual belief according to which the justifying grounds don't entail the truth of the belief, there should also be very little resistance to similar accounts of justified *a priori* or commonsense belief—especially when such accounts fit so nicely with our introspective understanding of what is going on in typical cases of what seem like justified *a priori* or commonsense beliefs.

The above presentation of the Reidian position on how we know our faculties are reliable and on where epistemic circularity enters into our noetic structure isn't intended to address in a satisfactory manner all possible objections. Instead, it is intended only to describe the position, to briefly highlight its advantages over argument-based accounts of how we know our faculties are reliable, and to say something about how we can sensibly take seriously the core ingredients of the Reidian position. Since neither externalism, nor J_{PF}, nor my defense of epistemic circularity depends on the truth of this Reidian account, I don't need to defend its truth here. Nevertheless, I thought it would be helpful at least to lay this Reidian position on the table as a potential asset for those who feel uncomfortable with the suggestion that we rely on arguments to know our faculties are reliable.

This page intentionally left blank

FOR EDUCATIONAL PURPOSES ONLY

8

Responding to Skepticism

The most natural and obvious objection to externalism is that it doesn't seem to offer an adequate response to skepticism. It may not be easy to say exactly what this inadequacy consists in. But as soon as undergraduates in their first philosophy class hear what the externalist response to skepticism is, the more philosophically inclined among them immediately view it as unsatisfying. In this final chapter, I will be defending externalism against that very natural reaction. I'll begin, in Section 1, by laying out both the externalist response to skepticism and the charges made against that response by those who find it unsatisfying. In Section 2 I will argue that opponents of externalism are saddled with the very same problems they attribute to externalism (or with worse problems). In the final section, I'll reply to the complaints lodged against externalist responses to skepticism, arguing that those complaints don't give us a good reason to reject externalism.

1 THE EXTERNALIST RESPONSE TO SKEPTICISM

The skeptic's worry is most easily motivated by considering the fact that a brain in a vat or someone hooked up to the Matrix or someone victimized by a Cartesian evil demon can have sensory experience that exactly matches our own. These considerations give rise to the following question: How can we be justified in believing as we do about the world around us on the basis of our sensory experience unless we have good reason to think our sensory experience isn't artificially induced as it is in these skeptical scenarios? By stipulation, it seems, we aren't able to tell the difference between a situation in which our sensory experience is caused in what we think of as the normal way and a situation in which one of the above skeptical scenarios is true. So how could we possibly have a good reason for thinking our sensory experience isn't artificially induced in a misleading way? How could we possibly be justified in believing as we do about the world around us?

Let's begin by focusing on the simple reliabilist's reply to this sort of skeptical worry. With this reply, it is especially easy to appreciate the exasperation of those who object to the externalist's handling of skepticism. Here is the sort of

‘maddening’ response the simple reliabilist gives to skeptical concerns sketched in the previous paragraph:

The Simple Reliabilist’s Response to Skepticism: A person’s beliefs are justified if and only if they are produced in a reliable way. So the demon victim’s beliefs are unjustified because they aren’t produced in a reliable way whereas the person whose beliefs *are* produced in a reliable way has justified beliefs. Thus, even if the sensory experiences of the demon victim match the experiences of the person with beliefs produced in a reliable way, there remains that difference in reliability. And that’s what explains the difference in justification.

Thus, in answer to the question ‘How can we possibly be justified in thinking as we do about the external world (on the basis of sensory experience) despite the fact that our sensory experience matches exactly the sensory experience of a demon victim?’ the simple reliabilist says: ‘That’s easy. Beliefs are justified if and only if they’re reliably formed. And, although the demon victim’s beliefs aren’t reliably formed, ours are.’

As I said, it is very natural to react to this way of handling skepticism by viewing it as inadequate. Some might think this is an understatement. They think the externalist response to skepticism is absolutely outrageous. In any case, those who object to this externalist response will be apt to say that it seems philosophically unsatisfying. And in explaining why this is so, they will be likely to mention one or more of the following reasons:

1. *Uncomfortable Moving up a Level:* Externalist responses to skepticism naturally give rise to questions about higher-level beliefs such as ‘how do you know (or justifiably believe) that your beliefs *are* reliably formed?’ Yet when the time comes to answer these questions, externalists seem to lose confidence in their externalist intuitions. This is manifested in their discomfort at accepting a conclusion to which their position commits them, namely, that they can know (or justifiably believe) via track-record arguments that their faculties are reliable. This discomfort with accepting the implications of their externalist intuitions when applied to higher-level beliefs shows that externalists themselves have an implicit awareness of the implausibility of their position.¹
2. *Conditional Answer:* This externalist response to skepticism says, in effect, that *if* one’s belief is reliably formed, *then* it counts as justified. But we want more than this sort of conditional response; we are after an unconditional way of responding to the skeptical worry. And the externalist cannot give us that.²
3. *Anything Goes:* If we allow that this externalist response to external-world skepticism is a permissible move in a philosophical exchange, then we’ll be forced to admit that an externalist response to skepticism about paranoid or

¹ See Fumerton (1995: 173–80) for a discussion of this sort of worry.

² See Stroud (1994: 302–5) for a discussion of this worry.

insane or fanatical religious beliefs is also a permissible move in a philosophical exchange. We'll have to allow those who seriously hold such views to say in response to our skeptical doubts that the difference between us and them is that their beliefs are reliably formed and ours aren't. And it would be inappropriate to allow that.³

4. *Philosophically Irresponsible*: When pressed to explain how they know their situation isn't a skeptical one, externalists ignore the question and reassert that their situation isn't a skeptical one. In response to the question 'How do you know it doesn't just seem like your belief is reliably formed even though it isn't?' the externalist says 'My belief is justified because it *is* reliably formed whereas the beliefs of demon victims are unjustified because they aren't reliably formed'. But it is philosophically irresponsible simply to ignore a perfectly legitimate philosophical challenge in that way.⁴

In the final section of this chapter, I'll be discussing these four complaints at greater length. But at this point, I simply want to lay them out in order to capture the dissatisfaction people feel toward the externalist response to skepticism. It is a dissatisfaction that I think is very easy to appreciate.

So far, the only externalist response to skepticism we've looked at has been the simple reliabilist's response. And although that response makes it easy to appreciate the exasperation objectors feel when considering the externalist response to skepticism, it also contains some distracting elements. The main distraction is the fact that the simple reliabilist position has other flaws that aren't shared by all versions of externalism—flaws on which I *don't* want to focus in this chapter. One of these other flaws is its failure to handle the New Evil Demon Problem. According to that problem, reliabilism is implausible insofar as it claims that reliability is *necessary* for justification. This is implausible because people think that if we are justified in our ordinary beliefs, then surely demon victims with sensory experience matching our own are also justified in theirs, despite the fact that their beliefs are unreliably formed. This intuition conflicts with the simple reliabilist position described above according to which the reliable beliefs of ordinary believers are justified and the beliefs of demon victims with matching sensory

³ Richard Fumerton (2001: 343–4) argues that externalist responses to skepticism leave them without any cogent objections to Plantinga's 'Reformed Epistemology' defense of the rationality of Christian belief. And a well-known objection to that defense of the rationality of Christian belief is that if we allow it, we have no good objection to a similar defense offered on behalf of crazy views like belief in the Great Pumpkin or voodoo belief. (See DeRose (unpublished) and Plantinga (1983: 74–8, 2000: 342–52).) Putting those two together we get the 'anything goes' objection.

A similar concern is discussed by Sosa (1997: 424–5) concerning superstitious crystal ball gazers. It seems they can use the externalist response to skepticism in replying to skeptical challenges to their crystal ball beliefs (and it seems that externalists should feel uncomfortable about this).

⁴ This complaint is more often just under the surface than explicit perhaps because philosophers don't like to charge their fellow philosophers with being philosophically irresponsible. For example, Bonjour (1985: 36) notes that, until recently 'no serious philosopher of knowledge would have dreamed of proposing an externalist response to skepticism.'

experience are unjustified. Another flaw of simple reliabilism is that it implausibly claims that reliability is *sufficient* for justification. But as Plantinga has made clear, there can be beliefs due to malfunction that are reliably formed and yet clearly unjustified (such as the brain lesion, caused by a random burst of radiation, which causes you to believe you have a brain lesion).⁵

But neither of these defects of simple reliabilism is the source of the charge that the externalist response to skepticism is philosophically unsatisfying. It's true that that charge has to do with the insufficiency of the externalist's proposed supervenience base for justification. But switching from reliability to my proper function account won't alleviate the objector's concerns, even though doing so will give us a version of externalism without the two flaws noted in the previous paragraph. My proper function account of justification, J_{PF} , doesn't require reliability, so it doesn't fall prey to the New Evil Demon Problem.⁶ Nor does it say that reliability is sufficient for justification, so it won't be subject to the sorts of malfunction counterexamples that plague reliabilists. Nevertheless, it is just as vulnerable to the charge that it seems philosophically unsatisfying, in the ways noted above.

To see this, it will be helpful to introduce a more complicated skeptical scenario. In this scenario, the demon victim is a human whose cognitive faculties have been intentionally damaged by the demon so that her automatic belief responses to sensory experience are very different from our own. For example, the sorts of beliefs we automatically form in response to tactile experiences, she automatically forms in response to olfactory experiences, and vice versa. Moreover, the demon is the one producing the sensory experiences and it does so in such a way that the victim never notices any incoherence in her belief system. The seeming coherence of the victim's beliefs and their seeming appropriateness as responses to the experiences on which they're based matches the seeming coherence and appropriateness of our beliefs. That is, just as our beliefs seem to us to be mostly coherent and appropriate to our experience, so also the victim's beliefs seem to her to be mostly coherent and appropriate to her experience. Given that this demon victim has inappropriate belief responses to her sensory experience, internalists who are, like me, interested in objective justification will be inclined to say that the demon victim doesn't have justified beliefs (since her beliefs certainly aren't *in fact* fitting responses to her evidence).⁷ Let's call this sort of scenario the 'malfunctioning demon-victim scenario'.

The skeptical questions that arise in connection with this example are similar to those that arise in connection with the usual demon victim cases: How can

⁵ See Plantinga (1993b: ch. 9). Plantinga is discussing warrant, not justification, but this particular point applies just as well to objective justification of the sort we're focusing on in this book.

⁶ See Chapter 5 (especially pp. 111, 134, and 141–3) for a discussion of this point.

⁷ See Section 1.2 of Chapter 5 for further discussion.

we be justified in believing as we do about the world around us on the basis of our sensory experience without good reason to think our sensory experience isn't artificially induced and made to seem coherent and appropriate to experience as it is in the example of the malfunctioning demon-victim? By stipulation, it seems, we aren't able to tell the difference between a situation in which our sensory experience is caused in a fitting way and a situation in which we are like the malfunctioning demon-victim. So how could we possibly have a good reason for thinking our sensory experience isn't artificially induced in that sort of misleading way? How could we possibly be justified in believing as we do about the world around us?

Here is the proper functionalist's response to these skeptical worries:

The Proper Functionalist's Response to Skepticism: A person's beliefs are justified if and only if they satisfy the proper function and no-defeater conditions as these are laid out in J_{PF}. So the malfunctioning demon-victim's beliefs are unjustified because they aren't produced in the way J_{PF} requires whereas the person whose beliefs *are* produced in the way J_{PF} requires has justified beliefs. Thus, even if the beliefs of the malfunctioning demon-victim seem as coherent and appropriate to her as ours do to us, there remains the difference that our beliefs are produced in accord with proper function and the malfunctioning demon-victim's aren't. And that's what explains the difference in justification.

Presumably, the same exasperated reactions mentioned earlier in connection with the simple reliabilist's response to skepticism will be directed with equal force at the proper functionalist response to skepticism that I just presented. Objectors will be tempted to say things like the following:

- The proper functionalist response to skepticism says only that a belief is justified *if* it satisfies the proper function and no-defeater conditions; it doesn't give us the unconditional response we want.
- Lunatics or religious fanatics can give the very same response to skepticism about their views, namely, 'the difference is that our beliefs are formed in accord with proper function and yours aren't'.
- The proper functionalist response is philosophically irresponsible. For she ignores a legitimate skeptical question and reasserts her claim that her belief is justified because it satisfies the proper function condition whereas the belief of the person in the skeptical scenario doesn't.⁸

Thus, by expanding our focus to include—in addition to ordinary skeptical scenarios for simple reliabilism—the malfunctioning demon-victim example used

⁸ For reasons to be discussed in Section 3 below, there isn't much of a temptation to raise the first of the four specific complaints noted above (i.e. that the proper functionalist is in trouble when it comes to discussing the justification of higher-level beliefs because she is uncomfortable with relying on track-record arguments for a faculty's reliability).

against proper functionalism, we are able to clarify the nature of the objection to externalism under consideration. It is different from the New Evil Demon Problem and from complaints about some shortcoming of a particular externalist analysis. Instead, it is a more general complaint that will apply to any externalist method of responding to skepticism. It's the complaint that the externalist response to skepticism seems to be philosophically unsatisfying for reasons similar to the above three objections to the proper functionalist's response to skepticism.

Now that we have before us the externalist response to skepticism and a basic understanding of why people find it unacceptable, let's turn to my argument that objectors to externalism are saddled with problems of the same sort as (or worse than) those they attribute to externalism.

2 NONEXTERNALIST RESPONSES TO SKEPTICISM

The first clue that there is something seriously wrong with the very natural negative reaction to the externalist response to skepticism is that it seems to be relying on the assumption that there is a strong awareness requirement on justification and knowledge. (As I explained in Chapter 1, a strong awareness requirement demands, for the justification of S's belief B, that S be aware on reflection alone of something contributing to B's justification, and that she be aware of it in such a way that she conceives of that contributor as in some way relevant to B's justification.) The simple reliabilist and the proper functionalist both insist that what makes a belief justified is the obtaining of some fact of which the subject needn't be even potentially aware (the fact in question is the satisfaction of a condition supposed to be sufficient for justification). But the objectors to externalism seem to think it's important to note that the subject, to whom the externalist attributes a justified belief, is unable to tell the difference between the obtaining of that fact and the obtaining of some other incompatible fact (i.e. the skeptical scenario in which the justification conditions aren't satisfied). These complaints at the very least *suggest* that the objectors to externalism think the subject must be aware of the obtaining of certain facts contributing to the justification of her beliefs and that, in doing so, she must conceive (or be able to conceive) of these facts as different from certain incompatible facts that wouldn't contribute to the justification of her beliefs (this is what the ability to tell the difference requires).

But of course, if the objector insists on a strong awareness requirement on justification, then, as I argued in Section 3.2 of Chapter 1, she cannot avoid vicious regress problems which lead quickly and fairly obviously to radical skepticism. And, as I argued in Section 3.4 of Chapter 1, those regress problems are ones everyone should want to avoid, even those interested in taking skepticism seriously. So no one should be willing to accept the vicious regress problems that follow from insisting on a strong awareness requirement on justification. Consequently no one should be interested in pushing the objection to externalism

described in Section 1 of this chapter according to which the externalist response to skepticism is philosophically unsatisfying. For, as I suggested in the previous paragraph, that objection depends on the assumption that there is a strong awareness requirement on justification. That is the conclusion I want to defend in this section.

In order to establish that conclusion, it will be helpful to think of nonexternalist views of justification as falling into the following two categories:

Strong Awareness Internalism: the view that there is a strong awareness requirement on justification.

Moderate Nonexternalism: nonexternalist views of justification that require for justification things not required by externalism (such as weak awareness or evidence) but which don't impose strong awareness requirements on justification.

The most controversial part of my argument is the claim that:

P: Moderate nonexternalists fare no better than externalists when faced with the objection to externalism described in Section 1 (i.e. the charge that their response to skepticism is philosophically unsatisfying).

If I can establish P, then we will have a dilemma for all epistemologists: either we face the vicious regress problems associated with imposing a strong awareness requirement on justification or we are saddled with the consequence of holding a view that seems as philosophically unsatisfying as the externalist response to skepticism. My response to that dilemma is to take the second horn, which isn't so bad given that, as I argue in Section 3, that horn isn't very sharp.

In order to defend P, I'll divide up moderate nonexternalist views (as defined above) into the following three groups:

Low Standard: Moderate nonexternalist views according to which it is sufficient, for the justification of S's beliefs, that S thinks that all is well with her beliefs epistemically speaking.

Medium Standard: Moderate nonexternalist views that reject the Low Standard position but say that beliefs can be noninferentially justified even if they are not about facts that are directly before one's mind (perceptual beliefs are typically viewed as beliefs of this kind).

High Standard: Moderate nonexternalist views that reject the Low Standard position and say that beliefs can be noninferentially justified only if they are about facts directly before one's mind (introspective and a priori beliefs are often viewed as beliefs of this kind).

Let's focus first on moderate nonexternalist views endorsing the Low Standard position. To be honest, I don't know of anyone who holds such a position. And there seems to be a good reason for this. This sort of view attributes justified belief far too liberally. It suggests that the key to having justified beliefs is

optimism about the epistemic quality of one's beliefs. That response to skepticism will hardly seem like an improvement over externalism in the minds of those concerned that externalists aren't demanding enough about what is required for justification. We can, I think, safely ignore the Low Standard position in seeking for a version of moderate nonexternalism that does better than externalism at avoiding the objection considered in Section 1. Our focus should instead be on those views that admit the possibility that, though all seems well to the subject epistemically speaking, things aren't as they seem.

Let's turn, then, to the Medium Standard versions of moderate nonexternalism. According to views of this sort, noninferential justification for perceptual beliefs depends on their being based on sensory experience of the right kind. What makes the sensory experience of the *right kind*? Perhaps it is that the experience is such that the beliefs based on it *fit* that experience; or perhaps the experience must be *better explained* by the truth of the belief based on it than by any skeptical alternative. Whatever it is, this much is clear: according to the Medium Standard moderate nonexternalists, (a) it isn't enough for justification that the experience just seems to the subject to be of the right sort (since that is to take the Low Standards position); and (b) it isn't required for justification that the subject think of the belief as being based on an experience that the belief fits or that is best explained by the truth of the belief or whatever (since that is to impose a strong awareness requirement).

But given (a) and (b), it is easy to construct a skeptical scenario to which Medium Standard moderate nonexternalists seem unable to respond in a way that is any more philosophically satisfying than the externalist's way of responding to skepticism. We need only imagine a scenario in which this nonexternalist's proposed supervenience base for justification is absent for the subject's noninferential perceptual beliefs and yet everything seems to the subject to be going well epistemically speaking.⁹ Suppose, for example, that the problem is that the beliefs *don't* in fact fit the subject's sensory experience (as in the case of the malfunctioning demon-victim described in Section 1); or suppose the problem is that the sensory experiences in question *aren't* in fact best explained by the truth of the beliefs in question. It seems possible for a clever demon to arrange for those problems to occur while at the same time guaranteeing that everything seems to the subject to be epistemically hunky-dory: her perceptual beliefs *seem* to her to fit the evidence (even though they don't); her perceptual beliefs *seem* to her to be the best explanation of the sensory experiences on which they are based (even though they aren't).

Now the skeptic will ask the Medium Standard moderate nonexternalist how she knows she isn't in such a skeptical scenario. After all, the fact that she *seems*

⁹ Given that we know that these moderate nonexternalists don't think it sufficient for justification that everything seems to be going well epistemically speaking (since they reject the Low Standards position), we know that this sort of scenario is possible.

not to be doesn't help since that's how things would seem if she were in such a skeptical scenario. And, given that she can't tell whether she's in such a skeptical scenario, it seems that her beliefs cannot be justified. To this, the only response available to the Medium Standard moderate nonexternalist is that her beliefs *are* justified and the beliefs of the person in the skeptical scenario *aren't* because her beliefs really do fit her evidence (or they really are the best explanation of the experiences on which they're based) whereas this is not so for those in the skeptical scenarios just described. But of course this response is precisely as unsatisfying philosophically as the externalist's response to skepticism. The following sorts of things said against the externalist response to skepticism could be said against this response too:

- The Medium Standard moderate nonexternalist response to skepticism says only that a belief is justified *if* it satisfies the 'fitting the evidence' condition (or the 'best explanation of the evidence' condition); it doesn't give us the unconditional response we want.
- Lunatics or religious fanatics can give the very same response to skepticism about their views, namely, 'the difference is that our beliefs really do fit the evidence (or really are the best explanation of the evidence) and yours don't (and aren't)'.
- The Medium Standard moderate nonexternalist response is philosophically irresponsible. For she basically ignores a legitimate skeptical question and reasserts her claim that her belief is justified because it really does satisfy the 'fitting the evidence' condition (or the 'best explanation of the evidence' condition) whereas the belief of the person in the skeptical scenario only appears to do so.

Thus, Medium Standard moderate nonexternalism fares no better than externalism does when faced with the sort of objection considered in Section 1.

Let's turn, finally, to High Standard versions of moderate nonexternalism. These views differ from others insofar as they restrict noninferential justification to beliefs about facts directly before one's mind. As a result, they don't allow for cases where one has a justified belief about some fact F1 (which is not directly before one's mind) in virtue of having directly before one's mind some distinct fact F2 (which is indicative of F1 without entailing it). Skeptical scenarios often capitalize on the gap created by having before your mind something which merely indicates without being or entailing the fact your belief is about. By avoiding this gap, the High Standard versions of moderate nonexternalism make themselves invulnerable to certain sorts of skeptical hypotheses.

But invulnerability to skeptical hypotheses relying on such a gap isn't sufficient for invulnerability to all skeptical hypotheses. For even supporters of High Standard versions of moderate nonexternalism (such as Laurence Bonjour and Richard Fumerton) acknowledge that one can *think* one has some fact F directly

before one's mind when one doesn't.¹⁰ As a result, one can believe that *p*, thinking one has the fact that *p* directly before one's mind, when in fact one doesn't. This suggests the following skeptical scenario: a demon arranges for all of one's introspective beliefs to be mistaken because each seems to be a belief about a fact directly before the subject's mind when in fact it isn't. That is, all the introspective beliefs of this demon victim are like the wildly inaccurate introspective beliefs that a malfunctioning person can have about the number of spots in her visual field: they seem to her to be about facts directly before her mind, but in fact they aren't.¹¹

Now how does the supporter of High Standard moderate nonexternalism know that she isn't a victim of such a demon? She can't point out that many of her introspective beliefs *seem* to be about facts directly before her mind because that is exactly how things would seem if she were in the skeptical scenario in question. And since she can't tell that she's not in such a skeptical scenario, it seems that none of her introspective beliefs are justified. To this, the High Standard moderate nonexternalist can only reply that she differs from the person in the skeptical scenario as follows: her beliefs are about things that are *in fact* directly before her mind whereas it merely misleadingly *seems* this way to the person in the skeptical scenario. But once again, this response is just as unsatisfying philosophically as the externalist's response to skepticism. For here too, remarks like the following apply:

- This High Standard moderate nonexternalist response to skepticism says only that a belief is justified *if* what the belief is about is genuinely (and not merely apparently) directly before one's mind; it doesn't give us the unconditional response we want.
- Lunatics or religious fanatics can give the very same response to skepticism about their views, namely, 'the difference is that our beliefs are about things that are genuinely (and not merely apparently) directly before our minds and yours aren't'.
- This High Standard moderate nonexternalist response is philosophically irresponsible. For she basically ignores a legitimate skeptical question and reasserts her claim that her belief is justified because what the belief is about really

¹⁰ See Fumerton (1995: 77). Bonjour acknowledges (Bonjour and Sosa 2003: 192–4) that you can correctly think you have a fact directly before your mind but not be justified in thinking so (since the fact can be a complex one such as *this visual image contains 51 spots* and you might not be sufficiently competent at detecting whether that description applies). Since he allows for that possibility, he would surely also allow that one could (unjustifiedly) think one has a certain fact directly before one's mind when one doesn't.

¹¹ We can add, for good measure, that all of this demon victim's *a priori* beliefs are beliefs like Frege's belief, before Russell helped him, that for every property there is a set of just those things that have that property. These *a priori* beliefs are not in fact about facts directly before the demon victim's mind; nevertheless, the demon victim thinks they are.

is directly before her mind whereas the belief of the person in the skeptical scenario only appears to be about what is directly before her mind.

Thus, these High Standard moderate nonexternalists fare no better than externalists when faced with the sort of objection considered in Section 1.¹²

But perhaps the High Standard moderate nonexternalist could push for an even higher standard. Perhaps she could insist that although there are some introspective beliefs one could have that are apparently (but not actually) about facts directly before the subject's mind, there are other types of introspective beliefs which are such that it is impossible to even have the belief unless the belief is true. An example might be the indexical belief 'I'm experiencing that' (accompanied by an inner act of 'pointing' to a mental state): you might think that in order for S to have that belief, there must exist the experience to which 'that' refers (McGrew 1999: 228). Suppose that's right. And suppose the High Standard moderate nonexternalist insists that it is only introspective beliefs of that sort—infallible introspective beliefs—that are justified. Will that enable her to avoid the sort of objection faced by externalists in Section 1 of this chapter?

No. Consider the following skeptical scenario. Through the work of a clever demon, a subject is unable to differentiate infallible introspective beliefs from fallible ones. Many introspective beliefs which are admitted by High Standard moderate nonexternalists to be fallible seem to this subject to be obviously infallible. That is, just as it seems obvious to infallibilists that a belief such as 'I'm experiencing that' couldn't possibly be false, so also it seems utterly obvious to the subject in my skeptical scenario that a belief like 'there are 47 speckles on the visual image before my mind's eye' couldn't possibly be false. Now the skeptic asks the proponent of the infallibilist High Standard position how she knows that she isn't like the person in the skeptical scenario just described. How does she know that her favored introspective beliefs really are infallible? The fact that they seem to be infallible is just what she should expect if she were in the skeptical scenario in question. This suggests that none of her introspective beliefs is justified. To this, the infallibilist must reply by saying that the difference between her and the person in the skeptical scenario is that her introspective beliefs really *are* infallible whereas the introspective beliefs of the person in the skeptical scenario

¹² Thus Fumerton (2006) is mistaken when he says that *having a fact directly before one's mind* is philosophically satisfying in a way that *satisfying externalist conditions on justification* isn't. The reason he thinks *satisfying externalist justification conditions* isn't philosophically satisfying is that, as we all realize, it can seem to us as if our beliefs satisfy externalist conditions when in fact they don't. According to Fumerton, philosophical curiosity demands *assurance* that they really are satisfied; the mere fact that they *are* satisfied isn't enough. But the very same points apply to *having a fact directly before one's mind*. As I've just noted in the text, it can seem to you that the thing your belief is about is directly before your mind when in fact it isn't. To be consistent, then, Fumerton should insist that philosophical curiosity demands *assurance* that the thing the belief is about really is before one's mind; he should insist that the mere fact that it *is* before one's mind isn't enough. The problem is that this demand for assurance seems to be an endorsement of a strong awareness requirement on justification.

only *seem* to be infallible. But this response to the skepticism generated by this skeptical scenario is at least as philosophically unsatisfying as the externalist's response to the skeptical scenario designed to undermine the externalist's position. For (at the risk of belaboring the obvious) each of the following remarks apply:

- This infallibilist High Standard moderate nonexternalist response to skepticism says only that a belief is justified *if* it is genuinely (and not merely apparently) infallible; it doesn't give us the unconditional response we want.
- Lunatics or religious fanatics can give the very same response to skepticism about their views, namely, 'the difference is that our beliefs are genuinely (and not merely apparently) infallible and yours aren't'.
- The infallibilist High Standard moderate nonexternalist response is philosophically irresponsible. For she basically ignores a legitimate skeptical question and reasserts her claim that her belief is justified because it is genuinely infallible whereas the belief of the person in the skeptical scenario only appears to be infallible.

Thus, these infallibilist High Standard moderate nonexternalists, like all the other moderate nonexternalists, fare no better than externalists when faced with the sort of objection considered in Section 1.

We can summarize my argument against moderate nonexternalists as follows. In order to avoid being clearly unsatisfying philosophically, they will reject the Low Standard position according to which a belief is justified so long as things seem to the subject to be OK epistemically speaking. But from this it follows that, no matter what is selected as a supervenience base for noninferential justification—whether it is the belief's fitting the evidence or its being the best explanation of the evidence, whether it is having the fact the belief is about directly before one's mind or the belief's being infallible—it will be possible for it to seem to the subject as if all is well epistemically speaking, even though the supervenience base for justification is absent. And that possibility gives rise to a troubling skeptical scenario. It's troubling for the moderate nonexternalist because her only answer to the question 'how do you know you aren't in such a skeptical scenario?' is an answer that is just as philosophically unsatisfying as the answer the externalist gave to the skeptic.

I conclude, therefore, that nonexternalists are no better off than externalists when it comes to responding to the objection discussed in Section 1. They will either be moderate nonexternalists or they'll impose a strong awareness requirement on justification. If they are moderate nonexternalists, their response to the sort of skeptical challenge that causes trouble for their view will be at least as philosophically unsatisfying as the externalist's response to skepticism. If, on the other hand, they impose a strong awareness requirement on justification, they are in even worse trouble because they will face the vicious regress problems discussed

in Sections 3.2 and 3.4 of Chapter 1. Either way, they are no better off than externalists.

This exposes the failure of yet another motivation for internalism. In Chapters 1 and 2, I argued that the main motivation for internalism (i.e. that it enabled one to avoid the Subject's Perspective Objection) failed. In Chapter 4, I argued that there was no adequate motive for internalism to be found in deontology. And in Chapter 7 I argued that avoiding epistemic circularity wasn't an adequate motivation for internalism given that the costs of avoiding epistemic circularity are worse than the costs of epistemic circularity itself. In this chapter, the motivation for internalism we've been considering is that endorsing it supposedly enables one to avoid the externalist's problem of giving a response to skepticism that is philosophically unsatisfying. But I've argued that *everyone* is forced to do either the same or worse. So this motivation for internalism fails too.

3 DEFENDING EXTERNALISM

In Section 1, I presented what I think is the most natural objection to externalism, namely, that its response to skepticism is philosophically unsatisfying and, therefore, inadequate. In Section 2, I argued that all responses to skepticism are either at least as philosophically unsatisfying as the externalist's or else they suffer from a worse problem (i.e. the vicious regress that afflicts positions endorsing a strong awareness requirement on justification). But so far that doesn't really say anything positive about the externalist response to skepticism; it merely places all others in the externalist's supposedly bad company. In this section I will remedy this by arguing first, in Sections 3.1–3.4, that being at least as philosophically unsatisfying as externalism isn't a reason to reject a position and second, in Section 3.5, that externalism is preferable to the moderate nonexternalist positions that are, as I argued in Section 2, at least as philosophically unsatisfying as externalism.

In Section 1, I mentioned four specific complaints intended to show that externalism is philosophically unsatisfying: the 'uncomfortable moving up a level' objection, the 'conditional answer' objection, the 'anything goes' objection, and the 'philosophically irresponsible' objection. In Section 2, I argued that the last three of these complaints applied as much to moderate nonexternalism as to externalism itself. Below, in Sections 3.2–3.4, I'll argue that these three complaints don't give us a good reason to reject either externalism or moderate nonexternalism. But first, in Section 3.1, I will argue that the first complaint doesn't even apply to all versions of externalism.

3.1 First Complaint: 'Uncomfortable Moving Up a Level'

The first of the four complaints in Section 1 about the externalist response to skepticism went as follows:

Uncomfortable Moving up a Level: Externalist responses to skepticism naturally give rise to questions about higher-level beliefs such as ‘how do you know (or justifiably believe) that your beliefs *are* reliably formed?’ Yet when it comes time to answer these questions, externalists seem to lose confidence in their externalist intuitions. This is manifested in their discomfort with accepting a conclusion their position commits them to, namely, that they know (or justifiably believe) via track-record arguments that their faculties are reliable. This discomfort with accepting the implications of their externalist intuitions when applied to higher-level beliefs shows that externalists themselves have an implicit awareness of the implausibility of their position.

This complaint is a summary of an objection by Richard Fumerton in a recent paper of his.¹³ We can think of this objection as making two points. First, it says that the externalist response to skepticism naturally gives rise to questions about how we know our beliefs are reliably formed. Second, it says the externalist feels uncomfortable answering that question in the way she is committed to answering it (i.e. permitting reliance on a track record argument), thereby revealing her own tacit uneasiness with her position.

Now why think, as the first of those two points says, that the externalist’s response to skepticism naturally gives rise to the question about how we know our beliefs are reliably formed? The answer seems to be that the externalist has identified *reliability* as the supervenience base for justified belief and claimed that its presence is what makes the difference between our beliefs and the demon victim’s. This naturally makes us curious about how we know that supervenience base—i.e. reliability—is really present. But notice that this only applies to reliabilist versions of externalism since only they propose reliability as the supervenience base for justification. The question that would naturally arise for proper functionalist versions of externalism such as J_{PF} is how we know the proper function and no-defeater conditions are satisfied, not how we know the beliefs are reliably formed. And that difference is important. For although a track-record argument can be used to argue for the reliability of a belief’s formation, it can’t be used to argue for the satisfaction of the proper function and no-defeater conditions mentioned in J_{PF}. So the second point mentioned at the end of the previous paragraph doesn’t apply to proper functionalism; the proper functionalist *isn’t* committed to permitting reliance on track record arguments arguing that the proper function and no-defeater conditions are satisfied.

But suppose we ignore this concern and simply note that the proper functionalist, like the externalist, will feel uncomfortable saying that we can know, on the basis of track record arguments, that our faculties are reliable. Doesn’t this reveal that the proper functionalist is uncomfortable with her own

¹³ See Fumerton (2006). Fumerton is there responding to my defense of externalism (Bergmann 2000b) against his earlier objection to it (Fumerton 1995: ch. 6).

externalist principles? It needn't. For there are two very good reasons for someone to be uncomfortable with the suggestion that we can know our faculties are reliable by relying on epistemically circular track-record arguments—reasons that apply even if one is completely comfortable with endorsing externalist principles. I mentioned the first reason in Section 4.2.1 of Chapter 7. There I explained that people are uncomfortable with epistemically circular track record arguments because, as I explained in Section 4.1 of Chapter 7, such arguments are useless if relied on to persuade doubters; they can't be used to arrive at a justified belief in a claim which is doubted. Given that arguments are typically evaluated by how effective they will be at convincing doubters, it's no wonder that externalists as well as internalists are uncomfortable with epistemically circular track-record arguments.¹⁴ I alluded to the second reason for discomfort with such arguments in the opening paragraph of Section 5 of Chapter 7. There I pointed out that we don't in fact seem to rely on epistemically circular track-record arguments in coming to believe our faculties are reliable. Instead we seem to form such beliefs noninferentially, without the use of arguments. This too contributes to our discomfort with the suggestion that it is acceptable to rely on such arguments to form beliefs about the reliability of our faculties.

In short, because we think both that (a) it is implausible to think we actually do rely on epistemically circular track-record arguments and that (b) doing so is typically a bad thing to do, it is unsurprising that there is widespread discomfort with the proposal that belief in the reliability of our faculties can rely on such arguments. Moreover, as I explained in Chapter 7, externalists too can sensibly endorse (a) and (b). It seems, therefore, that one can be uncomfortable with reliance on track-record arguments without being uncomfortable with externalist epistemic principles.

3.2 Second Complaint: 'Conditional Answer'

The second of the four complaints mentioned in Section 1 said:

Conditional Answer: This externalist response to skepticism says, in effect, that *if* one's belief is reliably formed, *then* it counts as justified. But we want more than this sort of conditional response; we are after an unconditional way of responding to the skeptical worry. And the externalist cannot give us that.

And, as I argued in Section 2, similar complaints apply to various versions of moderate nonexternalism:

¹⁴ However, as I explained in Section 4.2.1 of Chapter 7, one can sensibly use such an argument for a different purpose, namely, to arrive at a belief in a conclusion you've never doubted. This shows that even though it is understandable that people are uncomfortable with track-record arguments, such arguments aren't in themselves bad things.

- The Medium Standard moderate nonexternalist response to skepticism says only that a belief is justified *if* it satisfies the ‘fitting the evidence’ condition (or the ‘best explanation of the evidence’ condition); it doesn’t give us the unconditional response we want.
- This [fallibilist] High Standard moderate nonexternalist response to skepticism says only that a belief is justified *if* what the belief is about is genuinely (and not merely apparently) directly before one’s mind; it doesn’t give us the unconditional response we want.
- This infallibilist High Standard moderate nonexternalist response to skepticism says only that a belief is justified *if* it is genuinely (and not merely apparently) infallible; it doesn’t give us the unconditional response we want.

By way of response, the first thing to notice is that we have good reason to think there can be noninferential justification. Recall the following argument from Section 2.1 in Chapter 7:

1. A belief can be justified only if it is inferentially justified.
2. A belief can be inferentially justified only if the belief from which it is inferred is a justified belief.
3. Therefore, a belief is justified only if it is justified via circular reasoning or it is justified via an infinite chain of reasoning. [from 1 and 2]
4. No beliefs can be justified via circular reasoning.
5. None of our beliefs are justified via infinite chains of reasoning.
6. Therefore, none of our beliefs are justified. [from 3, 4 and 5]

As I explained there, given the plausibility of premises 2, 4, and 5, we are forced to choose between denying premise 1 and accepting this easy demonstration of extreme skepticism. Even those sympathetic to skepticism don’t think it can be established this easily. As a result, there is widespread acceptance of the foundationalist view that premise 1 is false—i.e. that there *can* be noninferential justification.

The reason I bring this up is that philosophers who think there can be noninferential justification will think that a noninferential belief is justified *if it satisfies some condition C*. And just in virtue of holding this foundationalist view they are making the kind of conditional assertion the second complaint from Section 1 is targeting. Moreover, these foundationalists will think a belief can be justified in virtue of satisfying C even if the person holding the belief doesn’t in any way support that belief by appealing to the fact that it satisfies C. For again, the belief is *noninferentially* justified simply in virtue of satisfying C. So the foundationalist will resist any demands to give a reason for thinking that the belief satisfies C. She will insist that satisfying C is sufficient for justification—that is, she will insist that *if* the belief satisfies C, it is justified. But in saying this, she will be giving what the proponent of this second complaint says is a merely conditional answer.

It seems, therefore, that those pressing this second complaint should have the same complaint about anyone who says there can be noninferential justification (whether they're externalists or moderate nonexternalists). For if there is noninferential justification, then the sensible thing to say—when someone asks for a reason to think that some allegedly noninferentially justified belief satisfies the condition C on which noninferential justification is supposed to supervene—is that you don't need to give such a reason in order for the belief to be justified; it's enough for justification that the belief does in fact satisfy C. This will, of course, be a conditional claim saying that a belief is justified if it satisfies C. But the reason this claim is given is to make it clear that the justification is noninferential in nature—that satisfying C is enough. Thus, on the assumption that those who propose the second complaint don't want to object to the foundationalist view that there can be noninferential justification, it seems that their complaint is based on confusion. It takes the foundationalist's sensible insistence that there can be noninferentially justified belief (an insistence which is expressed by asserting a conditional like 'if a belief B satisfies C, B is justified') and it wrongly concludes that there is something unsatisfyingly conditional about her position.

Moreover, it's not as if the foundationalist (whether externalist or internalist) can't add that the subject also has a further justified belief—albeit one that isn't required for the justification of B—in the antecedent of the conditional (i.e. that belief B *does* satisfy C). A belief in that antecedent will be justified so long as it too satisfies C.¹⁵ To complain that this gives us only another conditional claim is just to complain again about the foundationalist view that there can be noninferential justified belief. But in light of the argument for foundationalism given above (and defended in Chapter 7), it doesn't seem there will be many who are tempted to take up the alternative position that a belief can be justified only if it is inferentially justified.

3.3 Third Complaint: 'Anything Goes'

The third of the four complaints against externalism mentioned in Section 1 runs as follows:

Anything Goes: If we allow that this externalist response to external-world skepticism is a permissible move in a philosophical exchange, then we'll be forced to admit that an externalist response to skepticism about paranoid or insane or fanatical religious beliefs is also a permissible move in a philosophical exchange. We'll have to allow those who seriously hold such views to say in response to our skeptical doubts that the difference between us and them is that their beliefs are reliably formed and ours aren't. And it would be ridiculous to allow that.

¹⁵ See Sosa (1997: 428–9) for this sort of response.

As I argued in Section 2, here too there are similar complaints that can be lodged against various versions of moderate nonexternalism:

- Lunatics or religious fanatics can give the very same response [as Medium Standard moderate nonexternalism gives to the skeptical challenge it faces] to skepticism about their views, namely, ‘the difference is that our beliefs really do fit the evidence (or really are the best explanation of the evidence) and yours don’t (and aren’t)’.
- Lunatics or religious fanatics can give the very same response [as fallibilist High Standard moderate nonexternalism gives to the skeptical challenge it faces] to skepticism about their views, namely, ‘the difference is that our beliefs are about things that are genuinely (and not merely apparently) directly before our minds and yours aren’t’.
- Lunatics or religious fanatics can give the very same response [as infallibilist High Standard moderate nonexternalism gives to the skeptical challenge it faces] to skepticism about their views, namely, ‘the difference is that our beliefs are genuinely (and not merely apparently) infallible and yours aren’t’.

In each case, the complaint depends on the assumption that we can’t allow someone who seriously holds some crazy view (such as a fanatical religious view¹⁶) to respond to our skeptical challenges by saying ‘the difference between my view and yours is that my view genuinely satisfies condition C whereas yours only appears to satisfy C’; instead, we must be able to say that such a response is not a permissible move in a proper philosophical exchange.

But is this right? Consider for example direct acquaintance theorists such as Fumerton and Bonjour who endorse versions of High Standard moderate nonexternalism according to which a belief is justified only if it is genuinely (not merely apparently) about what is directly before one’s mind. Now suppose they realize that no matter what they propose as a supervenience base for noninferential justification, it will be possible for a skeptic to design a skeptical scenario in which the supervenience base seems to the subject to be present though it in fact is not. And suppose also that because they realize this, their response when they are faced with such a skeptical scenario is to insist that the difference between the demon victims in the skeptical scenario and them is that their beliefs are genuinely about what is directly before their minds and the demon victim’s beliefs aren’t. Is this a permissible move in a philosophical exchange?

¹⁶ It’s probably worth mentioning that there are differences of opinion about what counts as a ‘fanatical’ religious view. I suspect my own religious position—which includes at its core most of what Thomas Aquinas, John Calvin, and C. S. Lewis agreed about theologically—would, in virtue of including those Christian doctrines, be viewed as ‘fanatical’ by some people. I disagree of course. But I do think there are fanatical religious views: take for example Jim Jones, David Koresh, Osama bin Laden, and their devoted followers. I have views like those in mind when I speak of religious fanatics.

In Section 3.4, I will argue that it is. Suppose for now that I'm right about that. (It certainly doesn't seem *obvious* that it isn't permissible. After all, if anyone is concerned to make moves that are permissible in a proper philosophical exchange, it is High Standard moderate nonexternalists.) What follows? It follows, I think, that the religious fanatic who makes a similar reply in defense of her position is making a permissible move. And this is so even if the religious fanatic says the supervenience base on which justification supervenes is proper function rather than *being about a fact directly before one's mind*. What matters is that the structure of the reply is the same.

But how can this be? How can we say that the religious fanatic, who claims that the difference between her belief and ours is that hers is formed in accord with proper function and ours isn't, is making a permissible move in a proper philosophical exchange? These questions arise, I believe, out of some important misunderstandings. One misunderstanding is the thought that radical disagreement (about such things as fanatical religious views) can be resolved if we follow the rules for permissible moves in a proper philosophical exchange. This thought is a pipe dream, a philosopher's false hope. The disagreement between clever religious fanatics and those skeptical of their claims, like the disagreement between High Standard moderate nonexternalists and those skeptical of their claims, can 'bottom out' in the sort of exchange we've been imagining. The High Standard moderate nonexternalist can insist that genuine direct acquaintance with certain facts is sufficient for justification and that the demon victim with merely apparent direct acquaintance is out of luck justification-wise. The skeptic will find that unsatisfying. But the High Standard moderate nonexternalist won't be moved by the skeptic's dissatisfaction. The same sort of thing will happen in the case of religious fanatics: they won't be moved by the skeptic's dissatisfaction with their externalist response (nor, of course, will the skeptic be moved by their externalist response).

Suppose then that we give up the false hope that 'playing by the rules' of proper philosophical exchange will enable us to resolve radical disagreements. What then? Must we agree that the religious fanatics have respectable views? To think we must (because we've allowed that the religious fanatic is playing by the rules of proper philosophical exchange) is another misunderstanding. Both the religious fanatic and the person skeptical of her views will consider the other to be seriously out of touch with the way things are—each will think the other is either 'blinded' (i.e. lacking the ability to see some important truth) or 'hallucinating' (i.e. having an experience of something seeming to be true when it isn't). The skeptic will think the religious fanatic is hallucinating. The religious fanatic will think the religious skeptic is blinded to some important truths. We can agree that the externalist response to skepticism, even when employed by the religious fanatic, is a permissible philosophical move. But we can sensibly insist that such a fanatic is hallucinating, though we aren't sure how to deliver her from that error (given that she's philosophically astute and making no mistakes in her reasoning). Likewise,

we can sensibly say of those skeptical about introspection that their skepticism is due to some sort of blindness preventing them from seeing that their doubts aren't epistemically appropriate, though we aren't sure how to deliver them from that skepticism (since they aren't making any philosophical mistakes). We can say that what those skeptics need is somehow to see that the conditions we identify as sufficient for the noninferential justification of such beliefs *are* sufficient and to stop doubting that they are satisfied.

I say that we can insist the religious fanatic is hallucinating and that those skeptical of introspection are subject to some sort of blindness. But can't the religious fanatic just respond by saying that those of us who reject her view are subject to some sort of blindness? And can't those skeptical of introspection respond by saying that we who rely on it are or might be hallucinating (where it seems to us that our introspective beliefs are genuinely infallible or that they are about facts genuinely before our minds, even though they aren't)? Yes, the religious fanatic and the skeptic about introspection can respond in those ways. But in response to the religious fanatic, we can say: 'the difference between our claim that you're hallucinating and your claim that we're blind is that our claim satisfies the conditions necessary and sufficient for justification and yours does not'. Likewise, in response to the skeptic about introspection, we can say: 'the difference between our claim that you're blind and your claim that we're hallucinating is that our claim satisfies the conditions necessary and sufficient for justification and yours does not'. It's true that this is unlikely to satisfy the religious fanatic or the skeptic about introspection and that they will likely have similar things to say about us. But the point here is just this: The fact that those with whom we disagree (e.g. the religious fanatic or the skeptic about introspection) can respond with philosophical moves similar in form to our own might keep us from complaining that they aren't following the proper rules for philosophical exchange. It may even prevent us from resolving our dispute using the methods of philosophy. However, it doesn't commit us to thinking that their views are sensible or respectable.¹⁷

The response, then, to the 'anything goes' objection is to point out first that what's supposedly objectionable about the externalist response to external-world skepticism is also a feature of every moderate nonexternalist response to skepticism (including the direct acquaintance theorist's response to skepticism). This should help us see that this type of response is a permissible move in proper philosophical exchange. (As I said earlier, I'll say more in defense of the permissibility of this move in Section 3.4.) Once we see that, we should then be prepared to say that the externalist response by religious fanatics to skepticism about their views is also a permissible move in a proper philosophical exchange. But this doesn't mean we are forced to accept their religious fanaticism, or even to think of it as a respectable position; we can still think such fanatics are hallucinating. What it

¹⁷ Nor, of course, does it *prevent* us from thinking their views are sensible or respectable (which we can think even if we also think their views are mistaken).

does mean is that we must give up on the false hope that playing by the rules of proper philosophical exchange will enable us to resolve all serious disagreements about matters such as fanatical religious views.

But doesn't this reply to the 'anything goes' objection simply encourage us to give up on rational discourse as a way to resolve serious differences of opinion? It needn't. I'm certainly not recommending that we give up on philosophical exchange as a way to settle disagreements. I'm merely recommending a recognition of the fact that it has its limits, even when used properly. Nevertheless, I would strongly encourage attempts to employ philosophical methods, even when agreement seems difficult to achieve. If those contributing to the exchange are honest, willing to listen, interested in discovering the truth, and capable of philosophical give and take, then it's worth trying. There's always hope that such people can be moved by the reasons others present to them. Or at the very least, such people may be able to appreciate and even be moved by their interlocutor's careful attempt to explain how her overall position is a reasonable one in light of what she thinks of as the sensible insights included in the foundation of her perspective on the world. My point in the preceding paragraphs is just that we shouldn't conclude that our interlocutor is failing to follow the rules of proper philosophical exchange simply because she persists in her beliefs while being unable to satisfy our questions or challenges. Such persistence isn't always a sign of that sort of failure.

3.4 Fourth Complaint: 'Philosophically Irresponsible'

The fourth complaint is, I think, the most challenging:

Philosophically Irresponsible: When pressed to explain how they know their situation isn't a skeptical one, externalists ignore the question and reassert that their situation isn't a skeptical one. In response to the question 'How do you know it doesn't just seem like your belief is reliably formed when it isn't?' the externalist says 'My belief is justified because it *is* reliably formed whereas the beliefs of demon victims are unjustified because they aren't reliably formed'. But it is philosophically irresponsible simply to ignore a perfectly legitimate philosophical challenge in that way.

But it's important to remember that it is a challenge to us all, not just to externalists. As I noted above, moderate nonexternalists face similar complaints:

- The Medium Standard moderate nonexternalist response is philosophically irresponsible. For she ignores a legitimate skeptical question and reasserts her claim that her belief is justified because it really does satisfy the 'fitting the evidence' condition (or the 'best explanation of the evidence' condition) whereas the belief of the person in the skeptical scenario only appears to do so.

- This [fallibilist] High Standard moderate nonexternalist response is philosophically irresponsible. For she ignores a legitimate skeptical question and reasserts her claim that her belief is justified because what the belief is about really is directly before her mind whereas the belief of the person in the skeptical scenario only appears to be about what is directly before her mind.
- The infallibilist High Standard moderate nonexternalist response is philosophically irresponsible. For she ignores a legitimate skeptical question and reasserts her claim that her belief is justified because it is genuinely infallible whereas the belief of the person in the skeptical scenario only appears to be infallible.

If we're going to think seriously about justification, we have to choose between (i) imposing a strong awareness requirement on justification (in effect, saying that all justification is inferential) and (ii) facing this fourth complaint. I've already explained why I think the consequences of (i) are worse than the consequences of (ii): the former lead to vicious regresses making justification for mere mortals obviously hopeless. Most philosophers agree, it seems, and take the externalist or moderate nonexternalist line instead. But, as I argued in Section 2, they are thereby forced to face the fourth complaint.

Like the other complaints, the fourth complaint takes place in a context of discussing skepticism. A foundationalist (whether an externalist or a moderate nonexternalist) has proposed some condition C as necessary and sufficient for justification. The skeptic notes that there are skeptical scenarios in which it seems to a person that C is satisfied when in fact it isn't. And she asks the foundationalist in question how she can justifiedly believe she isn't in a scenario like that. The foundationalist's first response will be that it isn't required for the justification of S's belief B that S has the further justified belief that *B satisfies condition C*. The problem is that the skeptic can reply as follows:

Skeptic's Defense of the Irresponsibility Complaint: But it *is* required for the justification of S's belief B that S not have any defeaters for B. And it seems that when S is pressed about whether B is justified, she has only four options:

- (a) to believe that B is justified;
- (b) to believe that B isn't justified;
- (c) to explicitly withhold judgment on whether B is justified;
- (d) to do none of the above.

If S takes option (a), the question will just be raised again about the higher-level belief that *B is justified*. And again S will have the same four options. S can't take option (a) at every level because eventually the higher-level beliefs will have contents so complex that S can't even grasp them.¹⁸ Thus, S will eventually have

¹⁸ See Chapter 1 Section 3.2 for more on this problem.

to take some option other than (a). But S can't take option (b) or (c). For if S takes option (b), she will have a defeater for B. Likewise, if she takes option (c), she will also have a defeater for B. (To consider whether B is justified and explicitly withhold judgment on the matter indicates significant uncertainty about the credentials of B; the correct response to that uncertainty is to withhold B.) This leaves option (d). But to take option (d) will involve ignoring the questions 'Is your belief B justified? Does it satisfy condition C?' For if you consider these questions seriously at all, you must at least withhold judgment on whether B is justified; the only way to do none of (a)–(c) is not even to consider these questions. Thus, S's only remaining option—i.e. (d)—involves refraining from even thinking about the proposition that B is justified (or that B satisfies condition C). But this option involves sticking your head in the sand and refusing to face a legitimate philosophical challenge. And that is philosophically irresponsible.¹⁹

It's important to remind ourselves that it isn't just the externalist who faces this argument from the skeptic. Everyone who avoids the vicious regresses of Strong Awareness Internalism (rejecting the 'all justification is inferential' approach) will face this same complaint.

What is the correct response to the skeptical argument in the previous paragraph? As I see it, there are only three available responses:

1. Find some flaw in the assumptions or reasoning just laid out in the *Skeptic's Defense of the Irresponsibility Complaint*—a flaw which makes it a mistake to conclude that, when pressed about whether B is justified, S must eventually take option (d).
2. Accept the conclusion that, when pressed about whether B is justified, S must eventually take option (d) but deny that doing so is philosophically irresponsible.
3. Accept the conclusion that, when pressed about whether B is justified, S must eventually take option (d) and acknowledge that, in doing so, she is being philosophically irresponsible.

Response 3 is to be avoided because it pushes one toward Strong Awareness Internalism (recall that the 'Philosophical Irresponsibility' complaint applies to all who reject Strong Awareness Internalism). I'd be happy to take response 1 but I think the reasoning in question is quite persuasive and the relevant assumptions are quite plausible. Hence, I'm inclined to take response 2 by arguing that one can take option (d) without being philosophically irresponsible. What follows are two accounts of how this can be so.²⁰

¹⁹ For a more detailed presentation of this sort of skeptical argument, see Bergmann (2005: sections IIIa and IIIb).

²⁰ See Bergmann (2005: section IIIc) for further discussion of how one might respond to the skeptical argument under consideration.

First, suppose that S takes option (a) for the first few levels, believing that her beliefs at the previous levels are justified. But suppose that eventually S reaches a higher-level claim about a belief of hers being justified which is so complex that she isn't sure she can even grasp it. Because she isn't sure she can even grasp it, she neither believes it nor disbelieves it nor explicitly withholds it. This is a case where she takes option (d) without being philosophically irresponsible. For surely it isn't philosophically irresponsible to refrain from those attitudes (belief, disbelief, withholding) toward a proposition if you aren't sure you can even grasp the proposition.

The second account of how S can take option (d) without being philosophically irresponsible allows that S can responsibly refrain from considering the question the skeptic asks, even when she can clearly grasp the proposition she is being asked to consider (i.e. that her belief B is justified). Suppose that S has thought through this sort of question before. She realizes that it is a bad idea to endorse Strong Awareness Internalism. And she is aware of the line of reasoning given in the *Skeptic's Defense of the Irresponsibility Complaint*. Moreover, she is aware of the first reply above and realizes that if she persists in considering the higher-level propositions, she will justifiably believe each new claim until she reaches one she is unable to grasp and that when she reaches that point she won't be philosophically irresponsible in taking option (d). Based on all of this, she concludes that it would be fruitless to make an effort to believe as many of these higher-level propositions as she is able to grasp; for she knows that in the end she will unproblematically take option (d). So she takes option (d) up front without moving up even one level.

Moreover, she thinks that the skeptic's interest in pursuing these skeptical questions about ever higher-level claims is an indication of a philosophical temperament gone awry; it is not the legitimate pursuit of worthwhile philosophical questions. She thinks that, in the end, the foundationalist position—according to which satisfying some condition C is necessary and sufficient for justification—is the sensible and philosophically responsible approach to take. And she thinks this even though she is aware of the skeptical scenarios that can be proposed in which one thinks one's beliefs have satisfied C when they haven't. For, in light of what I've said in the previous two paragraphs, she thinks it can be philosophically responsible not to consider the skeptic's question about how she knows she isn't in one of those skeptical scenarios. She knows the sort of answer she *would* give to such a question and she knows that the conversation would continue as follows:

She would point out that she believes noninferentially (perhaps via common sense—see Section 5 of Chapter 7) that her beliefs satisfy C or that they are justified. The credentials of this higher-level belief would be challenged and, if she thought about the matter, she would noninferentially believe that these higher-level beliefs too satisfy C. The epistemic credentials of that still higher-level belief

would then be challenged and she'd respond in the same way. Things would continue until she is unable to grasp the content of the skeptical question, at which point she'd neither believe, disbelieve, nor withhold the relevant higher-level claim. She would thereby be taking option (d) in a philosophically responsible way.

Given that she knows all this, she is being philosophically responsible in not pursuing the skeptic's question even to the first higher-level belief. This is the second way to take option (d) without being philosophically irresponsible.

But isn't this demanding too much of those responding to the skeptic? Surely your average person won't be able to produce a reply of the sort I just laid out. Does that mean that your average person will be forced to succumb to the skeptic's challenges? No. The skeptic's charge is that those able to follow her complaints are being philosophically irresponsible if they simply ignore them. But those who aren't able to follow the skeptic's complaints aren't being philosophically irresponsible. They don't even see the problem. Some philosophical novices might follow the skeptic for a bit but then conclude that they are out of their depth (or that they're bored) and quite reasonably decide to leave these discussions to the experts. That doesn't seem to be philosophically irresponsible either. The charge of philosophical irresponsibility has a chance of sticking to its target only when applied to trained epistemologists capable of pursuing these matters in detail and at length. If in response to the skeptic's questions *they* simply bolt and run, despite the fact that they are able to think through these matters, then *perhaps* they are being philosophically irresponsible. What I've said in the previous paragraph is my attempt to explain how even capable philosophers can refrain from considering the skeptic's questions without being philosophically irresponsible.

Before moving on, let me return, briefly, to the argument of Section 3.3 in order to tie up a loose end. There I focused on the direct acquaintance theorist who thinks it is necessary for a belief's justification that the fact the belief is about is genuinely (and not merely apparently) before the subject's mind. I considered what such a direct acquaintance theorist might say when presented with a skeptical scenario in which a demon victim's introspective beliefs were only apparently (never genuinely) about things directly before the subject's mind. In particular, I considered what this direct acquaintance theorist might say when asked how she knows she isn't in such a skeptical scenario. The response I imagined this direct acquaintance theorist giving to such a question is this: 'The difference between my introspective beliefs and the introspective beliefs of the person in the skeptical scenario is that mine are genuinely about what is directly before my mind and the demon victim's introspective beliefs are only apparently about things directly before her mind'. My claim was that this was a permissible move in a philosophical exchange. And I said I would defend this assessment in Section 3.4. In the preceding paragraphs I've done just that by explaining how

we can responsibly refrain from considering the skeptic's questions when we're asked.

3.5 The Superiority of Externalism

In Section 2, I argued that moderate nonexternalism is at least as philosophically unsatisfying as externalism and that Strong Awareness Internalism is worse. That showed that the opponents of externalism had no advantage over externalists on this score. But it didn't show that externalism had a clean bill of health. In Sections 3.1–3.4, I argued that none of the four complaints against externalism (the last three of which also applied to moderate nonexternalism) gives us a good reason to reject either externalism or moderate nonexternalism. In fact, none of those four complaints gives us a good reason to think either externalism or moderate nonexternalism is philosophically unsatisfying. (Thus, when I said that moderate nonexternalism is at least as philosophically unsatisfying as externalism, I wasn't thereby saying that externalism was philosophically unsatisfying. Nor was I suggesting that moderate nonexternalism was as philosophically satisfying as externalism.)

I have, therefore, addressed the worry that externalism is philosophically unsatisfying. But why think externalism is to be preferred to moderate nonexternalism? After all, what I've said in Sections 3.2–3.4 helps the moderate nonexternalist as much as it helps the externalist. We can begin answering this question by noting briefly the following three virtues of externalism. First, externalism fits very well with our commonsense attributions of knowledge and justified belief to ordinary folk, including children. Second, externalism is consistent with the very plausible Assumption first introduced in Chapter 3:

Assumption: A belief is justified only if it is held in the right way. This requires, at the very least, that the belief is nondeviantly caused by an appropriate input to the belief-forming process.

Third, in Chapters 5 and 6 we have a plausible defense of an externalist account of justification.

But the most important reason for thinking externalism is preferable to moderate nonexternalism is this: *unlike externalists, moderate nonexternalists add unmotivated requirements on justification that externalists rightly think aren't necessary.* In Chapters 1 and 2 I argued that the main motivation for adding an awareness requirement, the Subject's Perspective Objection, is inadequate. In Chapter 4 I argued that the deontological motivation is inadequate. And in Chapter 3 I argued that mentalist requirements were also inadequately motivated. In Chapter 7, I argued that the desire to avoid permitting epistemic circularity didn't adequately motivate a rejection of externalism since all sensible views, whether externalist or not, had to permit it. Finally, in this chapter I argued that the complaints about the externalist response to skepticism didn't

motivate a rejection of externalism either. Not only do moderate nonexternalists face the same complaints, the complaints don't stand up under careful scrutiny. It's true that moderate nonexternalists sometimes complain that externalism requires more than it should for justification when it requires that the belief be reliably formed. But this charge applies only to reliabilist theories, not to all version of externalism (certainly not to the position defended in Chapters 5 and 6). By contrast, my charge that nonexternalist positions require more than they should applies to all versions of internalism and to mentalism as well. And I don't know of a version of moderate nonexternalism (with a following) that isn't either internalist or mentalist. For all of these reasons externalism stands out as the best position to take on epistemic justification.

FOR EDUCATIONAL PURPOSES ONLY

This page intentionally left blank

FOR EDUCATIONAL PURPOSES ONLY

References

- ALLEN, COLIN, BEKOFF, MARK, AND LAUDER, GEORGE (eds.) (1998), *Nature's Purposes: Analyses of Function and Design in Biology* (Cambridge, MA: MIT Press).
- ALSTON, WILLIAM (1985), 'Concepts of Epistemic Justification', *Monist*, **68**: 57–89, in Alston (1989: 81–114) [All page references are to reprint].
- (1986), 'Internalism and Externalism in Epistemology', *Philosophical Topics*, **14**: 179–221, in Alston (1989: 185–226) [All page references are to reprint].
- (1988), 'The Deontological Conception of Epistemic Justification', in James Tomberlin (eds.), *Philosophical Perspectives*, **2: Epistemology** (Atascadero, CA: Ridgeview), 257–99; in Alston (1989: 115–52).
- (1989), *Epistemic Justification: Essays in the Theory of Knowledge* (Ithaca, NY: Cornell University Press).
- (1993a), 'Epistemic Desiderata', *Philosophy and Phenomenological Research*, **53**: 527–51.
- (1993b), *The Reliability of Sense Perception* (Ithaca, NY: Cornell University Press).
- (1995), 'How to Think About Reliability', *Philosophical Topics*, **23**: 1–29.
- (2002), 'Plantinga, Naturalism, and Defeat', in Beilby (2002: 176–203).
- (2005a), *Beyond "Justification": Dimensions of Epistemic Evaluation* (Ithaca, NY: Cornell University Press).
- (2005b), 'Perception and Representation', *Philosophy and Phenomenological Research*, **70**: 253–89.
- ANNIS, DAVID (1973), 'Knowledge and Defeasibility', *Philosophical Studies*, **24**: 199–203.
- AUDI, ROBERT (2001), 'Doxastic Voluntarism and the Ethics of Belief', in Steup (2001b: 93–111).
- BEDAU, MARK (1993), 'Naturalism and Teleology', in Steven Wagner and Richard Warner (eds.), *Naturalism: A Critical Appraisal* (Notre Dame, IN: University of Notre Dame Press), 23–51.
- BEILBY, JAMES (ed.) (2002), *Naturalism Defeated?* (Ithaca, NY: Cornell University Press).
- BERGMANN, MICHAEL (1997a), 'Internalism, Externalism, and Epistemic Defeat' (PhD Dissertation: University of Notre Dame).
- (1997b), 'Internalism, Externalism and the No-Defeater Condition', *Synthese*, **110**: 399–417.
- (2000a), 'Deontology and Defeat', *Philosophy and Phenomenological Research*, **60**: 87–102.
- (2000b), 'Externalism and Skepticism', *The Philosophical Review*, **109**: 159–94.
- (2002), 'Commonsense Naturalism', in Beilby (2002: 61–90).
- (2004a), 'Epistemic Circularity: Malignant and Benign', *Philosophy and Phenomenological Research*, **69**: 709–27.
- (2004b), 'Externalist Justification Without Reliability', *Philosophical Issues, Epistemology*, **14**: 35–60.

- BERGMANN, MICHAEL (2005), 'Defeaters and Higher-Level Requirements', *Philosophical Quarterly*, 55: 419–36.
- (2006), 'A Dilemma for Internalism', in Crisp, Davidson, and Vanderlaan (2006): 134–74.
- (forthcoming a), 'BonJour's Dilemma', *Philosophical Studies*.
- (forthcoming b), 'Is Klein an Infitist about Doxastic Justification?', *Philosophical Studies*.
- BLOCK, NED (1990), 'Inverted Earth', in Tomberlin (1990: 51–79).
- (1994), 'Qualia', in S. Guttenplan (ed.), *A Companion to Philosophy of Mind* (Oxford: Blackwell), 514–20.
- (1996), 'Mental Paint and Mental Latex', in Villanueva (1996: 19–49).
- (1998), 'Is Experiencing Just Representing?', *Philosophy and Phenomenological Research*, 58: 663–70.
- (1999), 'Sexism, Racism, Ageism and the Nature of Consciousness', in Richard Moran, Jennifer Whiting, and Alan Sidelle (eds.), *The Philosophy of Sydney Shoemaker, Philosophical Topics* 26 (Fayetteville, NC: University of Arkansas Press), 39–70.
- (2003) "Mental Paint", in Martin Hahn and Bjorn Ramberg (eds.), *Reflections and Replies: Essays on the Philosophy of Tyler Burge* (Cambridge, MA: MIT Press), 165–200.
- BONJOUR, LAURENCE (1985), *The Structure of Empirical Knowledge* (Cambridge, MA: Harvard University Press).
- (2001), 'Towards a Defense of Empirical Foundationalism', in Michael DePaul (ed.), *Resurrecting Old-Fashioned Foundationalism* (Lanham, MD: Rowman and Littlefield), 21–38.
- (forthcoming), 'Replies to Bergmann, Kelly, and Feldman', *Philosophical Studies*.
- AND SOSA, ERNEST (2003), *Epistemic Justification* (Malden, MA: Blackwell).
- BOORSE, CHRISTOPHER (unpublished), 'Functions: A Current Scorecard'.
- BRAITHWAITE, RICHARD (1953), *Scientific Explanation* (Cambridge: Cambridge University Press).
- BRUECKNER, ANTHONY (1996), 'Deontology and Internalism in Epistemology', *Noûs*, 30: 527–36.
- BURGE, TYLER (1997), 'Two Kinds of Consciousness', in Ned Block, Owen Flanagan, and Guven Güzeldere (eds.), *The Nature of Consciousness: Philosophical Debates* (Cambridge, MA: MIT Press), 427–33.
- BYRNE, ALEX (2001), 'Intentionalism Defended', *The Philosophical Review*, 110: 199–240.
- CHALMERS, DAVID (1996), *The Conscious Mind* (Oxford: Oxford University Press).
- CHISHOLM, RICHARD (1977), *Theory of Knowledge, Second Edition* (Englewood Cliffs, NJ: Prentice Hall).
- (1982), *The Foundations of Knowing* (Minneapolis, MN: University of Minnesota Press).
- (1986a), 'Self-Profile', in Radu Bogdan (ed.), *Roderick M. Chisholm* (Dordrecht: D. Reidel), 3–77.
- (1986b), 'The Place of Epistemic Justification', *Philosophical Topics*, 14: 85–92.
- (1991), 'Firth and the Ethics of Belief', *Philosophy and Phenomenological Research*, 51: 119–28.

- COHEN, STEWART (1984), 'Justification and Truth', *Philosophical Studies*, 46: 279–95.
- (2002), 'Basic Knowledge and the Problem of Easy Knowledge', *Philosophy and Phenomenological Research*, 65: 309–29.
- COMESAÑA, JUAN (2002), 'The Diagonal and the Demon', *Philosophical Studies*, 110: 249–66.
- CONEE, EARL AND FELDMAN, RICHARD (1998), 'The Generality Problem for Reliabilism', *Philosophical Studies*, 89: 1–29.
- (2001), 'Internalism Defended', in Hilary Kornblith (ed.), *Epistemology: Internalism and Externalism* (Malden, MA: Blackwell), 230–60.
- (2004), *Evidentialism* (New York: Oxford University Press).
- CRISP, THOMAS, DAVIDSON, MATTHEW, AND VANDERLAAN, DAVID (eds.) (2006), *Knowledge and Reality: Essays in Honor of Alvin Plantinga* (Dordrecht: Kluwer Academic Publishers).
- DEROSE, KEITH (unpublished), 'Voodoo Epistemology', available online at: <http://pantheon.yale.edu/%7Ekd47/voodoo.htm>.
- DONAGAN, ALAN (1977), *The Theory of Morality* (Chicago, IL: University of Chicago Press).
- DRETSKE, FRED (1994), 'If You Can't Make One, You Don't Know How it Works', in Peter French, Theodore Uehling, and Howard Wettstein (eds.), *Midwest Studies in Philosophy*, 19, *Philosophical Naturalism* (Notre Dame, IN: University of Notre Dame Press), 468–82.
- (1995) *Naturalizing the Mind* (Cambridge, MA: MIT Press).
- (1996), 'Phenomenal Externalism or If Meanings Ain't in the Head, Where are Qualia?', in Villanueva (1996: 143–58).
- (2000), 'Norms, History, and the Constitution of the Mental', in Dretske, *Perception, Knowledge and Belief: Selected Essays* (New York: Cambridge University Press), 242–58.
- FALES, EVAN (1996), *A Defense of the Given* (Lanham, MD: Rowman and Littlefield).
- FELDMAN, RICHARD (1988a), 'Having Evidence', in D. F. Austin (ed.), *Philosophical Analysis* (Dordrecht: Kluwer Academic Publishers), 83–104.
- (1988b), 'Subjective and Objective Justification in Ethics and Epistemology', *Monist*, 71: 405–19.
- (2000), 'The Ethics of Belief', *Philosophy and Phenomenological Research*, 60: 667–95.
- (2001), 'Voluntary Belief and Epistemic Evaluation', in Steup (2001b: 77–92).
- (2003), *Epistemology* (Upper Saddle River, NJ: Prentice Hall).
- (2004), 'In Search of Internalism and Externalism', in Richard Schantz (ed.), *The Externalist Challenge. New Studies on Cognition and Intentionality* (New York: de Gruyter), 143–56.
- (2005), 'Internalist Epistemic Evaluations', in Matthias Steup and Ernest Sosa (eds.), *Contemporary Debates in Epistemology* (Malden, MA: Blackwell), 270–84.
- FELDMAN, RICHARD AND CONEE, EARL (1985), 'Evidentialism', *Philosophical Studies*, 48: 15–34.
- FIRTH, RODERICK (1978), 'Are Epistemic Concepts Reducible to Ethical Concepts?', in Alvin Goldman and Jaegwon Kim (eds.), *Values and Morals* (Dordrecht: D. Reidel), 215–29.

- FUMERTON, RICHARD (1988), 'Foundationalism, Conceptual Regress, and Reliabilism', *Analysis*, 48: 178–84.
- (1995), *Metaepistemology and Skepticism* (Lanham, MD: Rowman and Littlefield).
- (2001), 'Plantinga, Warrant, and Christian Belief', *Philosophi Christi*, 3: 341–51.
- (2006), 'Epistemic Internalism, Philosophical Assurance and the Skeptical Predicament', in Crisp, Davidson, and Vanderlaan (2006: 175–87).
- GETTIER, EDMUND (1963), 'Is True Justified Belief Knowledge?', *Analysis*, 23: 121–3.
- GINET, CARL (1975), *Knowledge, Perception, and Memory* (Dordrecht: D. Reidel).
- (2001), 'Deciding to Believe', in Steup (2001b: 63–76).
- GOLDMAN, ALVIN (1979), 'What is Justified Belief?' in Pappas (1979: 1–23).
- (1986), *Epistemology and Cognition* (Cambridge, MA: Harvard University Press).
- (1992), 'Epistemic Folkways and Scientific Epistemology', in Goldman, *Liaisons: Philosophy Meets the Cognitive and Social Sciences* (Cambridge, MA: MIT Press), 155–75.
- GOODMAN, NELSON (1983), *Fact, Fiction and Forecast*, 4th edn. (Cambridge, MA: Harvard University Press).
- GRECO, JOHN (1990), 'Internalism and Epistemically Responsible Belief', *Synthese*, 85: 245–77.
- (1999), 'Agent Reliabilism', in Tomberlin (1999: 273–96).
- (2000), *Putting Skeptics in their Place* (New York: Cambridge University Press).
- (unpublished), 'Holding Defeat to the Fire'.
- HAACK, SUSAN (1993), *Evidence and Inquiry: Towards Reconstruction in Epistemology* (Oxford: Blackwell).
- (1997), 'A Foundherentist Theory of Empirical Justification', in Pojman (1999: 289–93).
- HARMAN, GILBERT (1990), 'The Intrinsic Quality of Experience', in Tomberlin (1990: 31–52).
- (1996), 'Explaining Objective Color in Terms of Subjective Reactions', in Villanueva (1996: 1–17).
- HETHERINGTON, STEPHEN (1990), 'Epistemic Internalism's Dilemma', *American Philosophical Quarterly*, 27: 245–51.
- (1991) 'On Being Epistemically Internal', *Philosophy and Phenomenological Research*, 51: 855–71.
- HOWARD-SNYDER, DANIEL (1998), 'BonJour's "Basic Antifoundationalist Argument" and the Doctrine of the Given', *The Southern Journal of Philosophy*, 36: 163–77.
- JACKSON, FRANK (2000), *From Metaphysics to Ethics* (New York: Oxford University Press).
- KLEIN, PETER (1976), 'Knowledge, Causality, and Defeasibility', *The Journal of Philosophy*, 73: 792–812.
- (1981), *Certainty: A Refutation of Scepticism* (Minneapolis: University of Minnesota Press).
- (1999), 'Human Knowledge and the Infinite Regress of Reasons', in Tomberlin (1999: 297–325).
- (2000), 'The Failures of Dogmatism and A New Pyrrhonism', *Acta Analytica*, 15: 7–24.

- (forthcoming a), 'How to be an Infnitist about Doxastic Justification', *Philosophical Studies*.
- (forthcoming b), 'Human Knowledge and the Infinite Progress of Reasoning', *Philosophical Studies*.
- KOONS, ROBERT (2000), *Realism Regained: An Exact Theory of Causation, Teleology and the Mind* (New York: Oxford University Press).
- KORCZ, KEITH ALLEN (1997), 'Recent Work on the Basing Relation', *American Philosophical Quarterly*, 34: 171–91.
- (2000), 'The Causal-Doxastic Theory of the Basing Relation', *Canadian Journal of Philosophy*, 30: 525–50.
- KVANVIG, JONATHAN (1984), 'Subjective Justification', *Mind*, 93: 71–84.
- LEHRER, KEITH (1990), *Theory of Knowledge* (Boulder, CO: Westview Press).
- LEHRER, KEITH AND PAXSON, THOMAS (1969), 'Knowledge: Undefeated Justified True Belief', *The Journal of Philosophy*, 66: 225–37.
- LEVINE, JOSEPH (1997), 'Are Qualia Just Representations? A Critical Notice of Michael Tye's *Ten Problems of Consciousness*', *Mind and Language*, 12: 101–13.
- (2001), *Purple Haze* (Oxford: Oxford University Press).
- LEWIS, DAVID (1980), 'Mad Pain and Martian Pain', in Ned Block (ed.), *Readings in Philosophy of Psychology, Volume 1* (Cambridge, MA: Harvard University Press), 216–22.
- (1986), *On the Plurality of Worlds* (New York: Blackwell).
- LYCAN, WILLIAM (1996), *Consciousness and Experience* (Cambridge, MA: MIT Press).
- MARKIE, PETER (2004), 'Nondoxastic Perceptual Experience', *Philosophy and Phenomenological Research*, 68: 530–53.
- MERRICKS, TRENTON (1995), 'Warrant Entails Truth', *Philosophy and Phenomenological Research*, 55: 841–55.
- MCGREW, TIMOTHY (1995), *The Foundations of Knowledge* (Lanham, MD: Littlefield Adams Books).
- (1999), 'A Defense of Classical Foundationalism', in Pojman (1999: 224–35).
- McKITRICK, JENNIFER (2002), 'Reid's Foundation for the Primary/Secondary Quality Distinction', *The Philosophical Quarterly*, 52: 478–94.
- MITTAG, DANIEL (2002), 'On the Causal-Doxastic Theory of the Basing Relation', *Canadian Journal of Philosophy*, 32: 543–60.
- MOSER, PAUL (1985), *Empirical Justification* (Boston, MA: D. Reidel).
- (1989), *Knowledge and Evidence* (New York: Cambridge University Press).
- NEANDER, KAREN (1991a), 'Functions as Selected Effects: The Conceptual Analyst's Defense', *The Philosophy of Science*, 58: 168–84.
- (1991b), 'The Teleological Notion of Function', *Australasian Journal of Philosophy*, 69: 454–68.
- NICHOLS, RYAN (2003), 'Reid's Inheritance from Locke, and how he Overcomes It', *Journal of the History of Philosophy*, 41: 471–91.
- NISSEN, LOWELL (1997), *Teleological Language in the Life Sciences* (Lanham, MD: Rowman and Littlefield).
- NOZICK, ROBERT (1981), *Philosophical Explanations* (Cambridge, MA: The Belknap Press).

- PAPPAS, GEORGE (ed.) (1979), *Justification and Knowledge* (Dordrecht: D. Reidel).
- PEACOCKE, CHRISTOPHER (1983), *Sense and Content* (Oxford: Oxford University Press).
- (1984), 'Colour Concepts and Colour Experience', *Synthese*, 58: 365–82.
- PLANTINGA, ALVIN (1974), *The Nature of Necessity* (Oxford: Oxford University Press).
- (1983), 'Reason and Belief in God', in Alvin Plantinga and Nicholas Wolterstorff (eds.), *Faith and Rationality* (Notre Dame, IN: University of Notre Dame Press), 16–93.
- (1993a), *Warrant and Proper Function* (New York: Oxford University Press).
- (1993b), *Warrant: The Current Debate* (New York: Oxford University Press).
- (1993c), 'Why We Need Proper Function', *Notus*, 27: 66–82.
- (1995), 'Reliabilism, Analyses and Defeaters', *Philosophy and Phenomenological Research*, 55: 427–64.
- (1996), 'Respondeo', in Jonathan Kvanvig (ed.), *Warrant in Contemporary Epistemology* (Lanham, MD: Rowman and Littlefield), 307–78.
- (1997), 'Warrant and Accidentally True Belief', *Analysis*, 57: 140–45.
- (2000), *Warranted Christian Belief* (New York: Oxford University Press).
- (2002), 'Reply to Beilby's Cohorts', in Beilby (2002: 204–75).
- (unpublished), 'Naturalism Defeated', available on-line at <http://www.homestead.com/philofreligion/files/alspaper.htm>.
- POJMAN, LOUIS (ed.) (1999), *The Theory of Knowledge: Classical and Contemporary Readings*, 2nd edn. (Belmont, CA: Wadsworth Publishing Company).
- POLLOCK, JOHN (1979), 'A Plethora of Epistemological Theories', in Pappas (1979: 93–113).
- (1986), *Contemporary Theories of Knowledge* (Savage, MD: Rowman and Littlefield).
- (1990), *Nomic Probability and the Foundations of Induction* (New York: Oxford University Press).
- AND CRUZ, JOSEPH (1999), *Contemporary Theories of Knowledge*, 2nd edn. (Lanham, MD: Rowman and Littlefield).
- PRYOR, JAMES (2001), 'Highlights of Recent Epistemology', *British Journal for the Philosophy of Science*, 52: 95–124.
- (2004), 'What's Wrong with Moore's Argument?', *Philosophical Issues, Epistemology*, 14: 349–78.
- REID, THOMAS [1764] (1997), *An Inquiry into the Human Mind on the Principles of Common Sense*, ed. Derek Brookes (Edinburgh: Edinburgh University Press).
- [1785] (2002), *Essays on the Intellectual Powers of Man*, ed. Derek Brookes (Edinburgh: Edinburgh University Press).
- ROBINSON, WILLIAM (1998), 'Intrinsic Qualities of Experience: Surviving Harman's Critique', *Erkenntnis*, 47: 285–309.
- RUSSELL, BRUCE (2001), 'Epistemic and Moral Duty', in Steup (2001b: 34–48).
- SCHMITT, FREDERICK (1992), *Knowledge and Belief* (New York: Routledge).
- SEARLE, JOHN (1992), *The Rediscovery of the Mind* (Cambridge, MA: MIT Press).
- SELLARS, WILFRID [1956] (1963), 'Empiricism and the Philosophy of Mind', in H. Feigl and M. Scriven (eds.), *Minnesota Studies in the Philosophy of Science*, 1: *Foundations of Science and the Concepts of Psychology and Psychoanalysis* (Minneapolis: University of

- Minnesota Press), 253–329; repr in *Science, Perception and Reality* (London: Routledge and Kegan Paul, 1963), 127–96.
- SHOPE, ROBERT (1983), *The Analysis of Knowing: A Decade of Research* (Princeton, NJ: Princeton University Press).
- SMITH, A. D. (1990), 'Of Primary and Secondary Qualities', *The Philosophical Review*, 99: 221–54.
- SORENSEN, ROY (1991), '"P, Therefore P" Without Circularity', *The Journal of Philosophy*, 88: 245–66.
- SOSA, ERNEST (1991), *Knowledge in Perspective: Selected Essays in Epistemology* (New York: Cambridge University Press).
- (1993), 'Proper Function and Virtue Epistemology', *Noûs*, 27: 51–65.
- (1997), 'Reflective Knowledge in the Best Circles', *The Journal of Philosophy*, 94: 410–30.
- STEUP, MATTHIAS (1996), *An Introduction to Contemporary Epistemology* (Upper Saddle River, NJ: Prentice Hall).
- (1997), 'A Defense of Internalism', in Pojman (1999: 373–84).
- (2000), 'Doxastic Voluntarism and Epistemic Deontology', *Acta Analytica*, 15: 25–56.
- (2001a), 'Epistemic Duty, Evidence, and Internality', in Steup (2001b: 134–48).
- (ed.) (2001b), *Knowledge, Truth and Duty* (New York: Oxford University Press).
- (2001c), 'The Analysis of Knowledge', in Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Spring 2001 Edition)*, URL = <<http://plato.stanford.edu/archives/spr2001/entries/knowledge-analysis/>>.
- STROUD, BARRY (1994), 'Scepticism, "Externalism", and the Goal of Epistemology', *Proceedings of the Aristotelian Society*, Supplementary Volume 68: 291–307.
- TOMBERLIN, JAMES (ed.) (1990), *Philosophical Perspectives 4, Action Theory and Philosophy of Mind* (Atascadero, CA: Ridgeview).
- (ed.) (1999), *Philosophical Perspectives, 13, Epistemology* (Cambridge, MA: Blackwell).
- TYE, MICHAEL (1995), *Ten Problems of Consciousness* (Cambridge, MA: MIT Press).
- UNGER, PETER (1986), 'The Cone Model of Knowledge', *Philosophical Topics*, 14: 125–78.
- VAN CLEVE, JAMES (1984), 'Reliability, Justification, and the Problem of Induction', in Peter French, Theodore Uehling, and Howard Wettstein (eds.), *Midwest Studies in Philosophy, IX, Causation and Causal Theories* (Minneapolis: University of Minnesota Press), 555–67.
- VILLANUEVA, ENRIQUE (ed.) (1996), *Philosophical Issues 7: Perception* (Atascadero, CA: Ridgeview).
- VOGEL, JONATHAN (2000), 'Reliabilism Leveled', *The Journal of Philosophy*, 97: 602–23.
- WOLTERSTORFF, NICHOLAS (2001) *Thomas Reid and the Story of Epistemology* (New York: Cambridge University Press).
- WRIGHT, CRISPIN (2002) '(Anti-)Sceptics Simple and Subtle: G.E. Moore and John McDowell', *Philosophy and Phenomenological Research*, 65: 330–48.

This page intentionally left blank

FOR EDUCATIONAL PURPOSES ONLY

Index

- accessibilism 45–6 n. 1, 49, 53, 54
- acquaintance 29–31, 230–2, 237
- Alston, W. 3 n. 1, 5 n. 6, 10, 15 n. 24, 53 n. 17, 54–5, 60, 78, 83–8, 93 n. 23, 121 n. 28, 122 n. 30, 164–5, 181 n. 4, 201 n. 34, 207–8, 209 n. 44, 210 n. 45
- Annis, D. 154 n. 1
- Assumption 50–3, 61–2, 63, 71–2, 238
- Audi, R. 78 n. 1
- awareness
 - actual 14–15
 - conceptual 13 n. 20, 17, 19, 20–1, 27–9
 - doxastic 14–16
 - nonconceptual 13 n. 20, 19–20, 27, 30
 - nondoxastic 14, 17–18
 - potential 9 n. 12, 14, 15–16, 18, 31, 35, 39
 - requirement 7, 9–24, 25–39, 45–7, 49–50, 53–5, 58, 77, 80, 81–2, 94–9, 101, 104–5, 131, 175–7, 218–19, 238
 - strong 13, 14–19, 22–3, 27–9, 30–1, 34, 35–9, 58, 95, 97 n. 28, 218–20, 223 n. 12, 224–5, 234–6
 - weak 13, 19–21, 23–4, 27–9, 30–8, 97 n. 28
- basing
 - mentalism and 48, 52, 65–6, 69
 - relation 51 n. 13
 - requirement 4, 8, 61–2, 85 n. 11
- Bedau, M. 145 n. 66
- blamelessness,
 - duty and 79, 81–2
 - epistemic 5, 8–9, 89–90, 92–6, 98–9, 100–1, 104–5, 135–6
- Blamelessness Argument 89–99
- blameworthiness,
 - derived 91, 93, 98–9, 105
 - duty and 79, 81–82
 - epistemic 5, 92–3, 98–9, 104–5
 - generic 90–91
 - original 91, 98–9, 101
- Block, N. 122 n. 30
- BonJour, L. 10–13, 17 n. 29, 20 n. 35, 22, 25–6, 35–8, 102, 141, 175–7, 186 n. 12, 209, 215 n. 4, 221–2, 230
- Boorse, C. 144 n. 64, 145 n. 67
- Braithwaite, R. 181 n. 3
- Brueckner, A. 81
- Burge, T. 122 n. 30
- Byrne, A. 122–30
- Chalmers, D. 122 n. 30
- Chisholm, R. 5 n. 7, 22, 112 n. 5, 146
- Cohen, S. 141–3, 196 n. 27, 200 n. 33
- coherentism 13, 25, 186
- Comesaña, J. 143
- common sense 207–11, 236
- Conee, E. 4 n. 5, 5–6, 10 n. 15, 20 n. 34, 45–50, 53 n. 17, 54–6, 59–60, 64–6, 73 n. 32, 110 n. 1, 111–12, 116 n. 15, 139–41, 149–51
- connectors 114–16, 140
 - characterized 114–15
 - disconnectors and 114–16, 140
 - examples 120, 138 n. 54, 140
 - learned and unlearned 117–18, 120, 138 n. 54
 - Objectivity and 115–18
- Cruz, J. 10 n. 15, 45, 47, 62 n. 21, 65–75, 131 n. 43
- defeaters 154 n. 2, 159–60
- defeaters
 - actual 160–1, 168–9, 170–2, 174–5
 - believed 160–2, 163, 168–9, 170–2, 174–5, 188–9, 192, 199–200
 - continuing 155, 157–8
 - disconnectors and 115 n. 11
 - epistemic circularity and 199–200
 - full 155 n. 6
 - J_{PF} and 133–4
 - justification- 94, 155 n. 6
 - mental state 154–9, 163, 170
 - newly acquired 155–7
 - no-reason 159 n. 12
 - partial 155 n. 6
 - propositional 154, 162–3
 - reason-defeating 159 n. 12
 - rebutting 158–9, 163, 189 n. 23
 - skeptical argument and 234–5
 - undercutting 115 n. 11, 158–9, 163, 189 n. 23, 199
 - unjustified beliefs as 164–7
- deontologism
 - characterized 77
 - blamelessness and 5–6, 8–9, 89–99, 104–5
 - epistemic duty and 78–83

- deontology (*cont.*)
 evidentialism and 101–5
 internalism and 77–105
 ‘ought implies can’ and 83–9
 proper function and 96–7 n. 28
 DeRose, K. 215 n. 3
 design 132–7, 142, 144–6, 147–51, 171
 direct acquaintance, *see* acquaintance
 Donagan, A. 78 n. 2, 79 n. 3
 doxastic responses
 fitting and unfitting 110–14, 117–21,
 135–7, 147–8, 169–72
 learned and unlearned 116–21, 135,
 137–8, 147–8
 proper function and 130–1, 133–5
 doxastic voluntarism 77–8, 87 n. 15, 100
 Dretske, F. 122 n. 30, 145 n. 65
 duty, *see* epistemic duty
- epistemic circularity
 benign 198–200
 characterized 179–82
 defended 184–92, 193–7, 202–6
 malignant 198–200
 epistemic duty
 objective 78–81, 100
 subjective 78–9, 81–2, 100
 epistemic responsibility 102–5
 evidentialism
 characterized 110–13
 deontology and 101–5
 Objectivity and 116
 proper function and 130–2
 propositional justification and 6
 externalism
 basing requirement and 61–2
 defense outlined 109–10, 238–9
 epistemic circularity and 183–4, 192–3
 mentalism and 59–60, 71–2
 no-believed-defeater condition and 175–7
 scepticism and 213–18, 225–38
 Subject’s Perspective Objection and
 11–12
- Fales, E. 38–43
 Feldman, R. 4 n. 5, 5–6, 10 n. 15, 20 n. 34,
 45–50, 53 n. 17, 54–6, 59–61, 63–6,
 73 n. 32, 78 n. 1, 79 n. 4, 110 n. 1,
 111–12, 113 n. 9, 116, 139–41,
 149–51, 209
 Firth, R. 4 n. 2, 146 n. 69
 foundationalism 25, 184–7, 228–9
 Fumerton, R. 10–11, 12 n. 18, 13 n. 21, 15,
 19, 22, 25–6, 29–31, 112 n. 5, 116,
 172–4, 183–4, 201 n. 34, 214 n. 1, 215
 n. 3, 221–2, 223 n. 12, 226, 230
- generality problem 20 n. 34, 53 n. 17
 Gettier, E. 6–7, 133 n. 45, 154
 Ginet, C. 78, 83–9
 Goldman, A. 78 n. 2, 117 n. 17, 142 n. 62,
 173, 176 n. 35
 Goodman, N. 187 n. 18
 Greco, J. 93 n. 23, 120 n. 24, 167 n. 22, 175
 n. 34
- Haack, S. 112 n. 5, 186 n. 15
 Harman, G. 122 n. 30
 Hetherington, S. 13, n. 22
 Howard-Snyder, D. 186 n. 12
- infallibilism 32–5, 39, 41–2, 223–4
 infinitism 186–7
 intentionalism 121–30
 internalism
 characterized 3, 7, 9–11, 46–7, 49–57,
 70
 deontology and 77–105
 dilemma for 13–14, 26–9
 internal states and 66–7
 mental account of 45, 49–50, 54–7,
 70
 motivation 11–12, 19, 23–4, 77, 105,
 197, 225
 no-believed-defeater condition and 175–7
 scepticism and 218–25
- Jackson, F. 126 n. 37
 J_{PF}
 defense 133–51, 163–77
 defined 133
 epistemic circularity and 199
 scepticism and 216–17
 track record arguments and 226
 justification
 degrees of 133 n. 47
 deontological 5–6, 8–9, 77–105
 doxastic 4–5, 6, 8–9, 48, 50 n. 11, 51, 62
 n. 20, 66, 85, 110 n. 1, 186–7
 objective 5–6, 8–9, 92, 112, 136, 216
 ordinary 6–8, 68–9, 71–75, 92–3
 procedural 68–70, 71–75
 propositional 4, 6, 8, 48, 50 n. 11, 62 n. 20,
 66, 84–9, 110 n. 1, 186
 subjective 5–6, 8–9, 79 n. 4, 116, 135–6
- Klein, P. 154, 186–7
 knowledge 7, 9 n. 10, 154
 Koons, R. 145 n. 65
 Korcz, K. 4 n. 3, 51 n. 13
 Kvanvig, J. 78 n. 2,

- Lehrer, K. 12 n. 18, 154 n. 1, 176 n. 36
 Levine, J. 122 n. 30
 Lewis, D. 119 n. 22, 161 n. 14
 logical circularity 181–2, 185–6, 188 n. 19,
 201 n. 35
 Lycan, W. 122 n. 30

 Markie, P. 120 n. 24
 McGrew, T. 32–3, 223
 McKittrick, J. 126 n. 37, 127 n. 39
 mentalism
 defined 45, 48, 65–70
 Conee and Feldman on 48–65
 internalism and 45, 49–57, 70
 objections to 57–64, 71–5
 Pollock and Cruz on 65–75
 proper function and 131
 Merricks, T. 133 n. 45
 methodism 22–3,
 Mittag, D. 51 n. 13
 Moser, P. 12 n. 18, 13, 15 n. 24, 17, 19, 25 n.
 4, 112 n. 5, 116, 176 n. 36

 naturalism 144–6
 Neander, K. 145 n. 65
 Necessity
 characterized 112–13
 counterexample to 118–21
 defense of counterexample to 121–30
 no-believed-defeater condition and
 169–72
 Nichols, R. 127 n. 39
 Nissen, L. 144 n. 64
 no-defeater condition
 defended 133–4, 163–8
 defined 162–3
 externalism and 175–7
 internalism and 175–7
 no-actual-defeater condition 163
 no-believed-defeater condition 163–8,
 169
 no-propositional-defeater condition 162
 warrant and 133
 noninternalist nonexternalist theory 56–7
 Nonreliability 112–13, 145–6
 Nozick, R. 176 n. 35

 Objectivity 112, 114–18, 145–6
 ‘ought implies can’ principle 83–9

 particularism 22–3, 210
 Peacocke, C. 122 n. 30
 Plantinga, A. 5 n. 6, 51 n. 12, 78 n. 2, 79 n. 3,
 80–81, 120 n. 24, 121 n. 27, 132–4, 136
 n. 51, 137 n. 53, 143 n. 63, 145 n. 66,
 148 n. 72, 149 n. 73, 155 n. 5, 155 n. 7,
 156 n. 9, 159 n. 11, 164–5, 176 n. 35,
 186 n. 13, 186 n. 14, 187 n. 18, 210 n.
 46, 215 n. 3, 216
 Pollock, J. 10 n. 15, 45, 47, 62 n. 21,
 65–75, 78 n. 2, 131 n. 43, 159, 160,
 187 n. 18
 primary qualities 125–9
 proper function, *see also* J_{PF}
 defeaters and 133, 156 n. 9, 170–2, 174,
 189, 199
 deontologism and 80–1, 96–7 n. 28
 epistemic circularity and 199
 fittingness and 130–31
 justification and 133–51
 mentalism and 59–61
 naturalism and 144–6
 skepticism and 216–18
 warrant and 132–3
 Pryor, J. 45 n. 1, 197 n. 29

 Rea, M. 117 n. 18
 regress problems 9–10 n. 13, 14–19, 22–3,
 35–38, 38–43, 95–6 n. 27, 98 n. 30,
 103, 172–5, 185–7, 218–19
 Reid, T. 117 n. 17, 119–21, 125–7, 128 n.
 40, 177, 207–11
 reliabilism 20, 52–3, 55, 60, 68 n. 26, 73–4,
 111–12, 133–4, 139–42, 172–4,
 183–4, 213–16, 226
 representationalism, *see* intentionalism
 responsibility, *see* epistemic responsibility
 Robinson, W. 122 n. 20
 Russell, B. 5–6, 112 n. 5, 116, 121 n. 27

 Schmitt, F. 75 n. 33
 Searle, J. 145 n. 67
 secondary qualities 125–9
 Sellars, W. 13 n. 21, 25–6
 Shope, R. 154 n. 1
 skepticism 14–19, 21–3, 32–3, 40, 50 n. 11,
 167–8, 185, 203, 209 n. 44, 213–38
 Smith, A. 127 n. 39
 Sorenson, R. 182 n. 7
 Sosa, E. 33 n. 10, 142–3, 147, 186 n. 14, 215
 n. 3, 229 n. 15
 Steup, M. 78 n. 1, 99–105, 111 n. 4, 112 n.
 5, 146 n. 68
 Stroud, B. 214 n. 2
 Subject’s Perspective Objection (SPO)
 11–12, 19–21, 27, 29, 31–2, 34, 38, 46,
 57–9
 supervenience 146

 track record argument 183–4, 201–2, 206–7,
 226–7

transparency 39–43

Tye, M. 122 n. 30

Unger, P. 78 n. 2

Van Cleve, J. 181 n. 3

Vogel, J. 141, 183–4

warrant 5–6, 9 n. 10, 111 n. 2, 121 n. 27,
132–33, 134, 149 n. 73, 155 n. 6, 164,
176, 181, 216 n. 5

well-foundedness 4 n. 5, 6, 48

Wolterstorff, N. 119 n. 21

Wright, C. 188 n. 19

FOR EDUCATIONAL PURPOSES ONLY