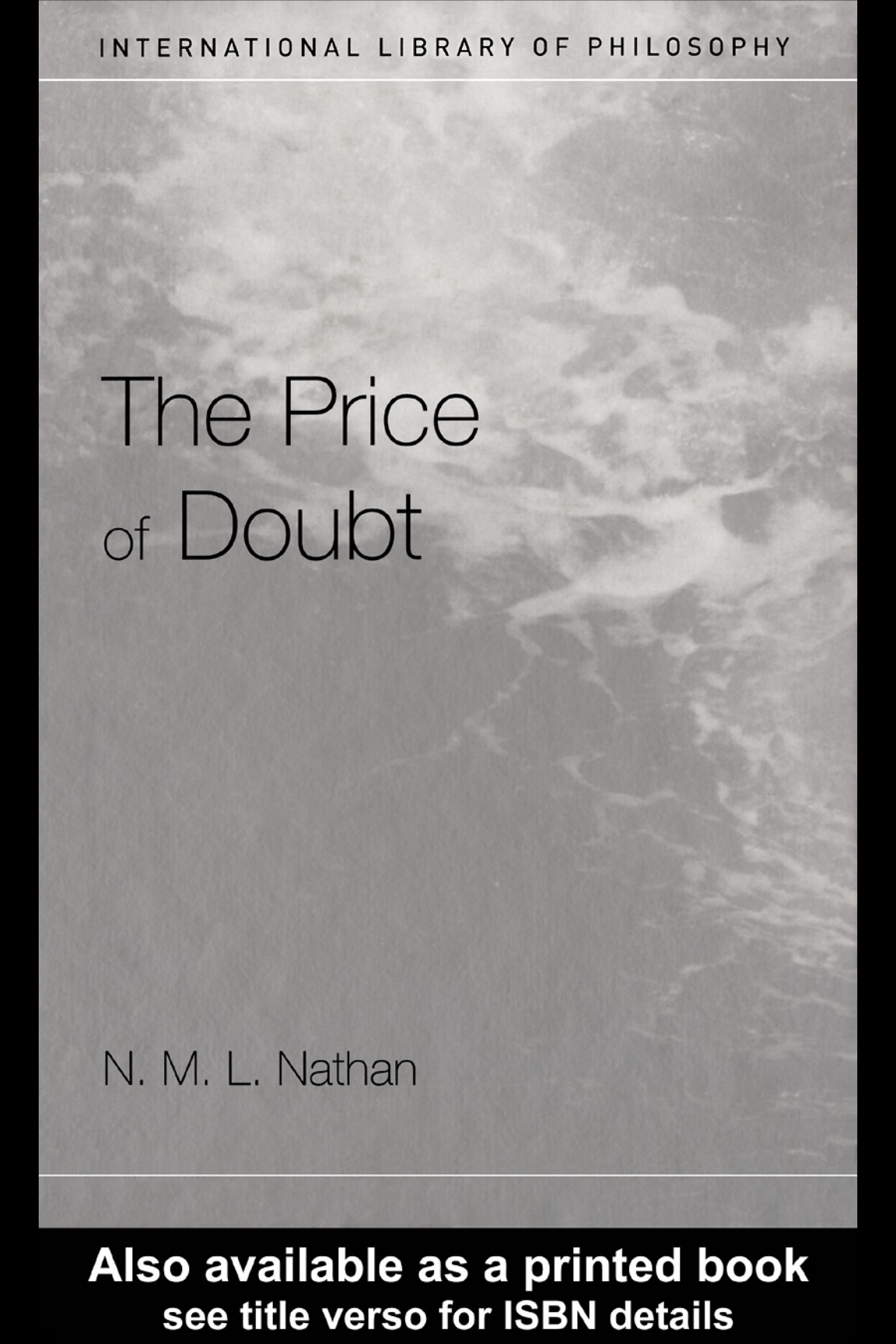


INTERNATIONAL LIBRARY OF PHILOSOPHY



The Price
of Doubt

N. M. L. Nathan

**Also available as a printed book
see title verso for ISBN details**

THE PRICE OF DOUBT

'N.M.L.Nathan's distinctive project is to investigate how philosophical conclusions bear on what we *want*. *The Price of Doubt* brings epistemology within the programme. None of us *wants* to suffer from the pervasive ignorance to which philosophical sceptics say we are doomed. Nathan argues that we needn't so suffer. The result is an intriguing study, which makes a novel contribution to epistemology'

Jennifer Hornsby, *Birkbeck College*

'...A highly original, provocative and thoroughly rigorous defence of why the "price" of sceptically-induced doubt is indifference. The arguments in this book are devilishly ingenious!'

Peter Klein, *Rutgers University*

An argument may fail because it is unsound. A trivial conclusion is another fault. Do all the sceptic's arguments have one or both of these deficiencies? This book conjectures that they do. Some sceptical arguments are entirely sound. Some have non-trivial conclusions. But none, it is conjectured, has both of these advantages.

To test this conjecture, Nathan explores both local and global forms of doubt, appraises arguments against justified and rational belief as well as arguments for ignorance. He assesses the Kierkegaardian desire for certain sceptical arguments to have true conclusions and examines our familiar ambition for belief to have a status which the sceptic tries to show it lacks. The chapters on arguments for ignorance discuss neo-Cartesian and infinite regress scepticism, together with more restricted doubts about inductive, evaluative and certain kinds of metaphysical knowledge. The treatment of arguments against justified belief is based on a critique of universalizing evaluation. Arguments against rational belief are measured against the requirements of a properly 'existential' attitude to one's own attitudes.

The Price of Doubt is an original and carefully constructed book which offers new remedies for the frustrations of traditional anti-sceptical epistemology.

Nicholas Nathan is University Research Fellow in the Department of Philosophy at the University of Liverpool.

INTERNATIONAL LIBRARY OF PHILOSOPHY

Edited by José Bermúdez, Tim Crane and Peter Sullivan

Advisory Board:

Jonathan Barnes

Fred Dretske

Frances Kamm

Brian Leiter

Huw Price

Sydney Shoemaker

Recent titles in the ILP:

PERSONAL IDENTITY AND SELF-CONSCIOUSNESS

Brian Garrett

PSYCHOLOGY FROM AN EMPIRICAL STANDPOINT

Franz Brentano

G.E.MOORE: SELECTED WRITINGS

G.E.Moore

Edited by Thomas Baldwin

THE FACTS OF CAUSATION

D.H.Mellor

THE CONCEPTUAL ROOTS OF MATHEMATICS

J.R.Lucas

STREAM OF CONSCIOUSNESS

Barry Dainton

KNOWLEDGE AND REFERENCE IN EMPIRICAL SCIENCE

Jody Azzouni

REASON WITHOUT FREEDOM

David Owen

THE PRICE OF DOUBT

N.M.L.Nathan

MATTERS OF MIND

Scott Sturgeon

LOGIC, FORM AND GRAMMAR

Peter Long

THE METAPHYSICIANS OF MEANING

Gideon Makin

THE PRICE OF DOUBT

N.M.L.Nathan



London and New York

First published 2001
by Routledge
11 New Fetter Lane, London EC4P 4EE

Simultaneously published in the USA and Canada
by Routledge
29 West 35th Street, New York, NY 10001

Routledge is an imprint of the Taylor & Francis Group

This edition published in the Taylor & Francis e-Library, 2003.

© 2001 N.M.L.Nathan

All rights reserved. No part of this book may be reprinted or reproduced or utilized in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data

Nicholas, Nathan, 1938–

The price of doubt/Nathan Nicholas.

p. cm.—(International library of philosophy)

Includes bibliographical references and index.

1. Belief and doubt. I. Title. II. Series.

BD215.N47 2000

121'.5—dc21 00-027028

ISBN 0-203-46821-X Master e-book ISBN

ISBN 0-203-77645-3 (Adobe eReader Format)

ISBN 0-415-23415-8 (Print Edition)

To A., M. and I.

CONTENTS

<i>Acknowledgements</i>	ix
Introduction	1
I Arguments for ignorance and the negation test	3
(1) Inductive ignorance	4
(2) The negation test	6
(3) Prospectus	16
II Demonic arguments	19
(1) The soundness of demonic arguments	20
(2) The negation test	24
(2.1) ENTAILMENT and consensus	26
(2.2) INCOHERENCE	28
(2.3) Excluding the conjunct	29
III Regressive arguments	31
(1) The soundness of regressive arguments	31
(1.1) Evidentialist arguments	32
(1.2) Semi-evidentialist arguments	34
(1.2.1) Deductivism	35
(1.2.2) Double regression	36
(2) The negation test	39
(2.1) Knowledge without evidence	40
(2.2) Consensual propositions	43

CONTENTS

IV Evidence-or-acquaintance arguments	50
(1) The objectivity of values	51
(1.1) Evaluative truth	51
(1.2) Dynamic unity	55
(2) Theism	63
(2.1) Evil	64
(2.2) Explanatory exclusion	67
V Anti-idealism	71
(1) Idealism and Anti-idealism	72
(2) Anti-idealism: an evidence-or-acquaintance argument	75
(3) The negation test	87
VI Arguments against justified belief	93
(1) Volitionally justified belief	95
(2) Arguments against volitionally justified belief	100
(3) The negation test	103
VII From rational action to rational belief	108
(1) Rational action	109
(2) Infinite regression	113
(3) Belief shortage	119
VIII Arguments against rational belief	124
(1) Continuing to believe	125
(2) Starting to believe	128
(3) Summary	134
IX The affirmation test	137
Appendix: Semantic objections to demonic arguments	142
<i>Notes</i>	149
<i>Works cited</i>	169
<i>Index</i>	174

ACKNOWLEDGEMENTS

I am greatly indebted to Professor Stephen Clark and the Academic Committee of the University of Liverpool who by allowing my transfer to a Research Fellowship gave me time to write this book. I thank them for that opportunity, and for their patience. For criticisms and encouragement I should like to thank Robert Audi, Bernard Beatty, Tim Crane, Barry Dainton, Kati Farkas, Robert Kane, Peter Klein, Ruth Weintraub, Linda Zagzebski and Routledge's anonymous referee.

INTRODUCTION

A good argument needs more than true premises and a properly derived conclusion. It needs a good conclusion. Q is true if 'P, so Q' is sound. What though if Q is trivial? What if its truth value is a matter of indifference? If Q is trivial, so is 'P, so Q'.

In what follows I apply this standard to a single, central group of philosophical arguments. The sceptic argues for your ignorance, or lack of justified or rational belief. Some of his arguments are, I think, entirely sound. But, I conjecture, no sound sceptical argument has a conclusion whose truth value is anything but a matter of indifference. All sceptical arguments are therefore either trivial or unsound. I pit this conjecture against a varied collection of ostensible counterexamples. Sound sceptical arguments I do find: some for ignorance, others against justified and against rational belief. But none have conclusions which are more than trivial.

Is the object just to damn a set of arguments? There is another side. As I will take it, Q's truth value is a matter of indifference just when reflection has a certain want-destructive power. Suppose that were a normally intelligent person to want Q's truth, reflection would destroy his attitude, and again that were a normally intelligent person to want Q's falsity, reflection would destroy his attitude. Then, and only then, is Q's truth value a matter of indifference. 'No sound sceptical argument has a conclusion whose truth value is anything but a matter of indifference.' This concerns our wants and their fragility. Wanting what sceptics say we do not have, we try to prove their arguments unsound. Grant my conjecture, and it does not matter that we sometimes fail. When we thus fail, reflection stops us wanting what the sceptic says we lack. There is a stoic remedy for the chronic frustrations of traditional anti-sceptical epistemology. What of pro-sceptical epistemology? What of the fideist, who wants to be ignorant in order to make room for faith? Here, I admit, the conjecture is less comforting. The fideist tries to prove his ignorance. Suppose he fails; his sceptical argument is unsound. It is consistent with my conjecture that some unsound sceptical arguments have conclusions whose truth values are

more than a matter of indifference. Reflection may not stop the fideist wanting to be ignorant.

Let me now put my proposal in a rather handier form. As I defined it, Q's truth value is more than a matter of indifference if and only if either (A) Were a normally intelligent person to want Q's truth, reflection would leave his attitude unscathed, or (B) Were a normally intelligent person to want Q's falsity, reflection would leave his attitude unscathed. Let us then say that 'P, so Q' passes the *affirmation test* if and only if (A), and that 'P, so Q' passes the *negation test* if and only if (B). Two tests. The conjecture, now, is that all sound sceptical arguments fail both.

There are nine chapters. In eight of these, I look for but fail to find sound sceptical arguments which pass the negation test. In I–VI try out arguments for ignorance, in VII try out arguments against justified belief. VII is preliminary to VIII, in which I try out arguments against rational belief. IX takes up the affirmation test.

How more exactly are these falsificatory efforts organized? How have I picked the sceptical arguments to try out? These questions will be answered at the end of chapter I. First let me give an actual instance of a sound and sceptical argument which fails the negation test.

I

ARGUMENTS FOR IGNORANCE AND THE NEGATION TEST

‘All sound sceptical arguments fail both the negation and the affirmation tests.’ I start with arguments for ignorance and the negation test. Does some sound and sceptical argument for ignorance get through that test? If so, my conjecture falls.

An argument for ignorance is any argument whose conclusion has this form: ‘N does not know that P’. ‘Occasionalism is false, so Mary does not know that Occasionalism is true.’ ‘John does not understand what “Occasionalism” means, so John does not know that Occasionalism is true.’ ‘The Magdeburg records were destroyed. Only if they hadn’t been could John know that his grandfather was born in Magdeburg. So John does not know that his grandfather was born in Magdeburg.’ These are all arguments for ignorance. None, however, is a *sceptical* argument for ignorance. An argument for ignorance is sceptical if and only if it is philosophical, and it is philosophical only if it does not rely on any premise about N’s personal deficiencies or special historical circumstances, and it is an argument in which philosophy does some actual work. Neither of the two arguments about John is by this standard philosophical, even though one of them concerns the philosophical doctrine of Occasionalism. John doesn’t understand what ‘Occasionalism’ means, but quite a lot of other people do. It is a personal deficiency. The Magdeburg records were destroyed: that is the historical situation in which John happens to find himself. The argument about Mary is not philosophical because even though it does not appeal to her purely personal deficiencies or special historical circumstances still it is an argument in which philosophy is idle. We do not need philosophy to tell us that not-p entails that p is not known.

Sound arguments for ignorance are as common as blackberries. But are there sound sceptical arguments for ignorance? They aren’t, in fact, too difficult to construct. Several varieties will appear in the next few chapters. Let me however sketch one straightaway. After that I can look more closely at the negation test.

(1) Inductive ignorance

Some arguments for ignorance go like this:

- (1) No proposition is evidence for P;
- (2) P is not entailed by any true proposition of the form 'N is acquainted with x';

So: (3) N does not know that P.

Suppose that P, in an argument of this form, stands for the proposition that Induction is reliable, that most of the time Induction yields true conclusions from true premises. This, we take it, is a contingent proposition. And for present purposes we can suppose that Induction is simple enumerative induction, the practice of inferring that n per cent of Fs are Gs from the premise that n per cent of the many Fs so far examined have been Gs, where it is observable whether something is an F and observable whether something is a G.¹ When P meets these conditions, 'knowledge' and 'evidence' can be so interpreted that the argument is sound. 'Knowledge' can be taken so that for any person S and any contingent proposition p, S knows that p only if either (i) he believes a proposition which is evidence for p, or (ii) p is entailed by a true proposition of the form 'S is acquainted with x'. And 'evidence' can be taken so that 'q is evidence for p' only if the argument 'q, so p' is not question-begging. We can then show that (1) is true because there is no non-question-begging argument whose conclusion is that Induction is reliable, and that (2) is true because 'Induction is reliable' is not entailed by any true proposition of the form 'N is acquainted with x'. Given the way in which we are taking 'knowledge', and given that 'Induction is reliable' is a contingent proposition, (3) will follow from (1) and (2). This gives us a perfectly sound argument for ignorance. The argument is moreover sceptical. It contains no premise about N's purely personal deficiencies or special historical circumstances. Nor is it an argument in which philosophy is idle: its validity is not evident until 'knowledge' and 'evidence' have been to some extent explicated.

Is it so sure that (1) is true when P stands for 'Induction is reliable'? Is there really no non-question-begging argument whose conclusion is that Induction is reliable? Let us assume that 'q, so p' is question-begging if any normally intelligent person who considered the argument, and who was doubtful about p's truth, would thereby be made equally doubtful about either q's truth or the truth of the conditional 'if q then p'. Anything on the lines of 'Nature is uniform, so Induction is reliable' is by this standard question-begging. So is 'So far Induction has worked well, so Induction is reliable'. It is hard to imagine any other plausible candidates.

Is it so sure that (2) is true? Is it so sure that 'Induction is reliable' is not entailed by any true proposition of the form 'N is acquainted with x'?

If there are internal relations and we can be directly acquainted with the intrinsic character of the relata of such relations, we might also be

acquainted with the fact that the relation obtains. If propositions are the sorts of things that we can hold directly before our minds, and if making probable is an internal relation holding between propositions, it might not be that hard...to claim that one can hold directly before one's mind the kind of fact that makes propositions of the form 'E makes probable P' true.²

Suppose that N is acquainted with the fact that a Keynesian internal relation of probabilification holds between propositions of the form 'All Fs have been Gs so far' and propositions of the form 'All Fs are Gs'. Wouldn't 'Induction is reliable' be entailed by the true proposition that N is acquainted with this fact? The difficulty is that we are taking 'Induction is reliable' to be a contingent proposition. 'Induction is reliable' would also have to be entailed by the true proposition that this logical relation of probabilification holds. But this last proposition would be necessary, if true at all, and no necessary proposition entails a contingent proposition.³ Though we seem not to be acquainted with anything other than the contents of our own experiences, it is a disputed question whether any of our experiences have contents other than sense data or internal objects which could not exist if we were not acquainted with them. It may perhaps be that some of our experiences have as contents external objects which could still exist or obtain even if no one were acquainted with them. But even on that view it is hard to see how, given its contingency, 'Induction is reliable' can be entailed by any true proposition of the form 'N is acquainted with x'. It is a fact that Induction is reliable. But this is a fact partly about the future, and even if we have experiences whose contents are external objects, we do not seem to have experiences whose contents are facts partly about the future.

A natural worry about the arguments whose soundness I am defending is that they employ special, unusual, carefully tailored senses of 'knowledge' and 'evidence'. But what exactly is the worry here? Not, surely, that, as interpreted, form (1)–(3) arguments are actually unsound. Is the point rather that since such arguments are sound only when 'knowledge' and 'evidence' are taken in highly eccentric ways, their soundness is purchased at the price of triviality? As it happens, the charge of eccentricity seems misplaced. To require that knowledge conforms to the evidence-or-acquaintance condition is not to use 'knowledge' in a wildly eccentric way. Nor does there seem to be any great novelty in the requirement that q is evidence for p only if the argument 'q, so p' is non-question-begging. Often, we cite evidence in order to induce or change belief: 'You think he's never been in Philadelphia. That can't be so: Joanna said she met him there in 1989.' Citing evidence could not have this function if q could be evidence for p even when 'q, so p' is question-begging. It is because we build this function into the concept of evidence that it seems so odd to say that a proposition is evidence for itself, or is evidence for a conjunction of which it is itself a conjunct. But the essential point is that these linguistic matters are in

any case irrelevant. Maybe the soundness of our arguments for inductive ignorance is indeed purchased at the price of triviality, but no purely linguistic considerations can show that this is so. Terminological eccentricity is neither a bar to soundness nor a guarantee of triviality.

There are, then, sound sceptical arguments for ignorance. But are there sound sceptical arguments for ignorance which pass the negation test?

(2) The negation test

'P, so Q' passes the negation test if and only if it meets this condition: were a normally intelligent person to want Q's falsity, reflection would leave his desire unscathed. That was how I put it in the Introduction. Arguments for ignorance have conclusions of the form 'N does not know that P'. Shall we, when we apply the negation test, think only of N's own hypothetical attitude to his knowing that P, suppose him normally intelligent, and ask whether reflection would on that assumption destroy his desire to know that P? Or shall we also consider the attitudes which other people might have to N's knowing that P? It is enough, I assume, just to consider N's hypothetical attitude to his own knowledge. I will take it, then, that an argument for N's ignorance passes the negation test if and only if were N normally intelligent, reflection would leave unscathed any desire he might have to enjoy the knowledge which the argument says he lacks.

Are there sound sceptical arguments for ignorance which pass the negation test? In the next four chapters I look for and fail to find them. But why look further than our arguments for inductive ignorance? Why don't they get through? This section provides the answer. As we shall see in section (3), it also helps to fix the directions of our future search.

The arguments for inductive ignorance went like this:

- (1) No proposition is evidence for P;
- (2) P is not entailed by any true proposition of the form 'N is acquainted with x';

So: (3) N does not know that P.

'Know', in (3), is governed by the condition that for any person S and any contingent proposition p, S knows that p only if either (i) he believes a proposition which is evidence for p, or (ii) p is entailed by a true proposition of the form 'S is acquainted with x'. And P stands for the contingent proposition that Induction is reliable. Take some such argument and suppose that the person N to whom the argument refers does initially want to have the knowledge of Induction's reliability which the argument says he lacks. How by reflection might he lose this initial desire?

'Nothing is more certain', Hume assures us, 'than that despair has almost the same effect upon us with enjoyment, and that we are no sooner acquainted with the impossibility of satisfying any desire, than the desire itself vanishes.'⁴ It is true, I think, that no want will survive its owner's belief that it

has a logically incoherent content. And it may well be that some sound sceptical arguments for ignorance have logically true conclusions. Hume would in that case point us to a way in which certain sceptical arguments will fail the negation test. N is convinced by some sceptical argument that he does not know that P. He wants this knowledge. But, as he sees, it is logically impossible for him to have it. He sees that his desire for knowledge has an incoherent content. His want does not survive this thought. Perhaps. But this does not help us in the present case. Our arguments for inductive ignorance have contingent conclusions. Even in our evidence-or-acquaintance sense of 'knowledge' it is not logically impossible for N to know that Induction is reliable. It isn't logically impossible for him to be acquainted with the future. N's desire to know that Induction is reliable does not have an incoherent content. Does Hume's principle hold for weaker than logical impossibility? Surely not. No want will survive its owner's belief that it has a logically incoherent content. But it is perfectly easy to want what you think it is in a weaker sense impossible for you to have. It isn't difficult to want the non-occurrence of the physically inevitable. We need a new idea.

P, in our arguments for inductive ignorance, stands for 'Induction is reliable'. Might N lose his desire to know that Induction is reliable by losing his desire for Induction actually to *be* reliable? If he wants Induction to be reliable, but reflection stops him wanting this, then the chances are that reflection will also stop him wanting to know that Induction is reliable. For it will hardly escape him that, however exactly 'knowledge' is taken in the sceptical argument for his ignorance, the proposition that Induction is reliable is entailed by the proposition that he knows that Induction is reliable. And if he does indeed believe that this entailment holds, then the way is clear for an application of the following general principle:

ENTAILMENT: For any person S and any propositions p and q, if S believes that p entails q, and if S does not want that q, then he does not want that p.

Grant for the sake of argument that this principle is true. How is reflection supposed to stop N wanting Induction to be reliable? Two not totally implausible trains of thought suggest themselves.

The first is this. Suppose that initially N does not believe that Induction is reliable but that reflection leads him to start believing that it is. Then reflection will also prevent him from wanting Induction to be reliable. For it seems to be a general truth that you cannot want what you already believe to be the case. So, given ENTAILMENT, and given that N believes that the relevant entailment holds, the reflection that leads him to start believing that Induction is reliable will destroy his desire to know that Induction is reliable. This is of course too good to be true. And the fault must lie with the principle that you cannot want what you already believe to be the case. Why exactly we should reject this principle, I am not entirely sure. Certainly it is difficult not to regard being glad that something is so as a species of wanting it to be so, and being glad that

something is so positively requires you to believe that it is so.⁵ I will, anyway, renounce the principle.

How else might reflection stop N wanting Induction to be reliable? His reflections might take a Popperian turn. He might stop wanting Induction to be reliable because he starts to think that we do not need to know or even to believe the empirical generalizations which Induction is supposed to be able to support. According to Popper, Hume proved that we do not know or are not justified in believing that Induction is reliable. Accordingly, we do not know the truth of any contingent empirical generalization which is partly about the future. The moral, for Popper, is that we should avoid the dogmatism of believing such generalizations. But this is no great sacrifice. The scientist can still ‘falsify’ or ‘refute’ his conjectured generalizations, in the sense of showing that they stand in certain logical relations to basic statements decided to be satisfactory or conventionally agreed to be true. He can ‘accept’ an empirical generalization in the sense that he can decide that attempts, or further attempts, should be made to ‘falsify’ it. And if an empirical generalization has survived more attempts to ‘falsify’ it than its ‘unfalsified’ rivals, then it is rational for us to act as if it is true. Knowledge of such generalizations is, for Popper, something we can learn to live without. So, likewise, is the reliability of Induction.

It is true that Popper does often talk as if he believes that, Hume notwithstanding, we can indeed gain knowledge of contingent empirical generalizations partly about the future. He talks of science as ‘conjectural knowledge’ and says that scientific knowledge grows. This lays him open to the charge of disguising his own scepticism. For David Stove, ‘conjectural knowledge’ is a ‘shocking phrase’, a glaring example of the more or less devious ‘neutralisation of success-words’ which is characteristic of the irrationalist philosophies of science spawned by Hume’s deductivism.⁶ This isn’t however a charge which can be made against all Popper’s followers. In his recent restatement and defence of Critical Rationalism, David Miller says that he is prepared to concede if he has to that science is not knowledge.⁷

Nor does there seem to be anything inconsistent about the rigorous programme of belief-abstention which Popper recommends. I do however doubt that most people who find themselves believing rather than in the Popperian sense accepting unknown and unfalsified empirical generalizations will feel themselves in the wrong. But if you do happily believe such generalizations why should you not want the repeated failure of falsificatory efforts to be evidence for their truth? And how could it be such evidence unless Induction is reliable?⁸

I take it then that reflection will not destroy N’s desire for Induction actually to be reliable. If sceptical arguments for inductive ignorance fail the negation test, it won’t be because reflection would destroy N’s desire for the actual truth of what these arguments say he does not know. We need another new idea.

You want what you think you cannot get. Naturally, you think of substitutes. Will something more easily obtainable do just as well? Believe as much, and your first desire will evaporate. You want five million marks purely and simply for the sake of buying a flat in Berlin. You come to realize that flats can be got there for just a million marks. This stops you wanting the five million. One million is an adequate substitute. By the sceptical argument, N lacks evidence-or-acquaintance knowledge that Induction is reliable. Can he find an adequate substitute for this missing knowledge?

One seeming substitute is justified belief. Crispin Wright maintains that there is

not necessarily any lasting discomfort in the claim that, contrary to our preconceptions, we have no genuine knowledge in some broad area of our thought—say in the area of theoretical science. We can live with the concession that we do not, strictly, *know* some of the things we believed ourselves to know, provided we can retain the thought that we are fully justified in accepting them.⁹

There is a problem here. Knowledge that *p* may well seem preferable to justified belief that *p* when *p* is a proposition which you want actually to be true. Justified belief is on most interpretations non-factive: your being justified in believing that *p* is compatible with *p*'s falsity. Knowledge is by contrast factive: necessarily if you know that *p* then *p* is true. If you want *p*'s truth, the factiveness of knowledge may make it preferable to mere justified belief. Does N want it to be true that Induction is reliable? We couldn't see why not. It looks then as if his substitute for the missing knowledge won't be justified belief.

N's missing knowledge is of the evidence-or-acquaintance kind. It is governed by this condition: for any person *S* and any contingent proposition *p*, *S* knows that *p* only if either (i) he believes a proposition which is evidence for *p*, or (ii) *p* is entailed by a true proposition of the form 'S is acquainted with *x*'. If justified belief won't serve N as a substitute for this missing knowledge, maybe some other kind of knowledge will. Why for example should he not be content just to have the kind of knowledge of Induction's reliability which consists in true belief non-accidentally generated by a reliable process? Given that Induction *is* reliable, reliabilist knowledge of its reliability is of course available. As numerous writers have insisted, Induction itself could then be the reliable process which non-accidentally generates the true belief in the reliability of Induction.¹⁰ From

(4) So far, most inductive arguments with true premises have had true conclusions

he could infer

(5) Most inductive arguments with true premises have true conclusions.

The argument '(4), so (5)' is doubtless question-begging. Anyone who considered this argument and who was doubtful about the truth of (5) would thereby be made equally doubtful about the truth of the conditional 'If (4) then (5)'. But this is no problem. Evidence-or-acquaintance knowledge requires the knower to have propositional evidence, when his knowledge is not by acquaintance. There is no such requirement, with reliabilist knowledge. Nor is it necessary for N to consider the argument in order for the inference to form his belief. Though question-begging, '(4), so (5)' is not viciously circular in the sense that 'its conclusion is contained among its premises',¹¹ or in the sense that 'a necessary condition of using it to gain knowledge of (or justified belief in) its conclusion is that one *already have* knowledge of (or justified belief in) its conclusion'.¹²

N may indeed think that, as knowledge goes, reliabilist knowledge is not inferior to knowledge of the evidence-or-acquaintance kind. But it seems to me that once he asks what reason he has to want the missing evidence-or-acquaintance knowledge that Induction is reliable, he will embark on a train of thought which will leave him with no desire for *any* kind of knowledge of that proposition. A fortiori, the argument for inductive ignorance will fail the negation test. We can perhaps say that this train of thought will leave him with a non-epistemic substitute for the missing knowledge. But the argument for his ignorance will fail the negation test for reasons which do not depend on his contentment with a different kind of knowledge of the same proposition.

N asks what reason he has to want the missing evidence-or-acquaintance knowledge that Induction is reliable. This leads him to pose a more general question. Why should we ever prefer knowledge to true belief? Stock answers are available. One, derived from Plato's *Meno*, is that knowledge is less likely to 'run away'.¹³ When true belief comes under pressure it is less likely to disintegrate if it is also knowledge. If it is knowledge then there is evidence to prop it up, or at least there is the chance that the reliable process can be repeated by which the belief was first generated. Another stock answer concerns the public good. When your own true belief is also knowledge, this increases the chances that others will grasp the same truth. Your evidence is perhaps publicly available: others can get hold of it too. The reliable process by which your belief was first generated can also do its work for others. It may even be that you will be able to function as a source of information about the truth you know: you have a property, detectable to persons to whom it is not yet detectable that p, which correlates in a law-like, non-accidental way with your being right that p.¹⁴ It seems to me that once it has occurred to N that there are these two general reasons for wanting knowledge, and he has started to apply them to his missing knowledge of Induction's reliability, he will find that he has certain beliefs which oblige him to conclude that neither reason really can be applied to this missing knowledge. The propositions by believing which he will be led

to this conclusion, are as follows: (A) Induction *is* reliable; (B) Almost anyone who entertains the proposition that Induction is reliable will then believe it; and (C) It is highly likely that he will believe that Induction is reliable whenever he entertains that proposition.

Let me take this bit by bit. Why, to start with, should N find himself believing these three propositions? First, N will *entertain* these three propositions, when he starts to wonder whether the stock reasons for wanting knowledge apply to the particular case of knowing that Induction is reliable. Obviously he will have (A) in mind. And (B) and (C) will be present to his mind as possible obstacles to the applicability of the stock reasons. But, once he does entertain these three propositions, he will in all probability believe them. Why so? Because that is the kind of proposition that 'Induction is reliable' is. It is what we can call a *consensual proposition*. A proposition is consensual if and only if (i) almost anyone who entertains it firmly believes it; (ii) almost anyone who entertains (i) firmly believes (i); (iii) almost anyone who entertains (ii) firmly believes (ii); and so on, ad infinitum. Other likely examples of consensual propositions include 'No proposition is both true and false', 'There are other minds' and 'Most people have feelings of the same quality when they spill boiling water on their hands'. Grant that 'Induction is reliable' is consensual. Now suppose that N entertains

(A) Induction is reliable.

Then by (i), it is highly probable that he will believe (A). Next, suppose that N entertains

(B) Almost anyone who entertains the proposition that Induction is reliable will then believe it.

By (ii) it is highly probable that he will believe (B). Suppose finally that he entertains

(C) It is highly likely that he will believe that Induction is reliable whenever he entertains that proposition.

Since he entertains (B) and in all probability believes it, and since (C) follows from (B), it is highly likely that he believes (C).

Now I must show that if N does thus believe the three propositions (A), (B) and (C), this will convince him that the two stock general reasons for wanting knowledge do not apply to his missing knowledge of Induction's reliability.

Consider to begin with the Platonic 'running away' idea. As we are supposing, N believes that

(C) It is highly likely that he will believe that Induction is reliable whenever he entertains that proposition.

As N will realize, (C) completely undermines the applicability of the 'running away' idea to his knowledge that Induction is reliable. Doubtless there are *some* propositions, which he has 'running away' reasons for wanting to know. Suppose he believes he has the true belief that Mary has no musical talent. He wishes he were wrong, and knows how prone he is to wishful thinking. It would

be better if he knew she had no musical talent, rather than just truly believed it, because that would make it easier for him to resist the pressure he constantly experiences to believe that she is a future concert pianist. Or again, he believes he has the true belief that his country will lose the war. But defeatism has become dangerous. Better then for him to know that his country will be defeated. This will make it less likely that fear will lead him to accept the popular illusion of inevitable victory. If however N believes (C), he will believe that his true belief that Induction is reliable is quite stable, and hence that the ‘running away’ reason does not apply. Nor in any case would he see *this* belief as threatened by his propensity for wishful thinking. Unlike ‘Mary has no musical talent’, ‘Induction is reliable’ is not something that he wants to be false. Nor will he see ‘Induction is reliable’ as something that it is dangerous to believe. There may be a widespread illusion that his country will win the war. But there isn’t a widespread illusion that Induction is unreliable.

The other stock reason for wanting your true belief also to be knowledge was that this would increase the chances that other people will grasp the truth that you are lucky enough to believe yourself. Once again, there will be numerous propositions which N does have this reason for wanting to know. But they won’t include ‘Induction is reliable’. As we are supposing, N believes that

(B) Almost anyone who entertains the proposition that Induction is reliable will then believe it.

And he will realize that if (B) is true then his own knowledge that Induction is reliable will do nothing to increase the chances that other people will grasp this truth. To grasp it, they need only entertain it. They do not need the evidence or generally reliable process of belief formation by virtue of which he would know. Nor do they need him as a source of information.

With the two stock reasons gone, what further reason can N have for wanting knowledge that Induction is reliable? There are perhaps *some* consensual propositions which N will have a reflectively indestructible desire to know, even though he recognizes that neither stock reason applies to that knowledge. There is always the possibility that a consensual proposition is entailed by a non-consensual proposition which N has a reflectively indestructible desire to know. In the next chapter we shall be looking at the neo-Cartesian hypothesis that nothing exists but N and a demon who makes N believe whatever it is that in the actual world N does believe. Clearly, the negation of that hypothesis is a consensual proposition. The negation of the demonic hypothesis is however entailed by the proposition that N’s sister is happy. If N believes that he knows that this entailment holds, and if he has a reflectively indestructible desire to know that his sister is happy, then he may want to know the negation of the demonic hypothesis purely and simply for the sake of knowing that his sister is happy. For the knowledge he wants that his sister is happy may be knowledge of a kind which is governed by the principle that if S knows that p, and knows that p entails q, then S knows that

q. Will N have *this* kind of reason for wanting to know that Induction is reliable? Not so far as I can see. ‘Induction is reliable’ seems not to be entailed by any other proposition which N would have a reflectively indestructible desire to know.

Are there yet further general reasons for wanting knowledge, which also apply to N’s missing knowledge that Induction is reliable? In his *Descartes*, Bernard Williams gives this explanation of why we want knowledge. Imagine a ‘primitive truth-gatherer’, A, who has ‘no elaborate or reflective demands’ but just wants true answers to various questions of the form ‘whether p’. He sees that once he has formed a belief that p it is no good trying to check it to see if it is really true. ‘Since to believe something is to believe that it is true, to acquire a belief is already to assume an answer to the question of whether it is true.’¹⁵ He will see that what he must do is to acquire answers to his questions by a method which generally yields true beliefs. In effect, he wants reliabilist knowledge for the sake of getting true answers to his questions. This is inapplicable to the present case. As we are supposing, N will believe that Induction is reliable as soon as he entertains it. If p were ‘Induction is reliable’ the primitive truth-gatherer would have his answer just as soon as he posed his question. He couldn’t wonder about p’s truth value without entertaining p, and that alone would be enough to ensure that he believed that p was true.

According to William Alston, we want to have non-question-begging arguments for the reliability of our doxastic or belief-forming practices because ‘we are interested in *discriminating* those that can reasonably be trusted from those that cannot’. For this reason we are not satisfied with ‘track-record’ arguments for the reliability of a doxastic practice. I earlier mentioned an argument in which the reliability of Induction is inferred from its past successes. This is a track-record argument, and it is sound only if Induction is indeed reliable. Alston thinks that by showing merely that *if* a given belief-source is reliable then it can be shown by its track-record to be reliable we do ‘nothing to indicate that the source belongs with the sheep rather than with the goats’.¹⁶ Could it then be that N has a discriminatory reason for wanting evidence-or-acquaintance knowledge that Induction is reliable? Not if, as we are supposing, N believes that

(B) Almost anyone who entertains the proposition that Induction is reliable will then believe it.

There may well be certain actual or imaginable belief-forming practices which N is uncertain about whether to trust. Those based on mystical experience may be an example. But Induction won’t be one of them, if N believes (B). Since he will then be sure that Induction is not among the goats, non-question-begging arguments for its reliability will strike him as superfluous.

Don’t we on reflection want just any knowledge for its own sake, even knowledge of Induction’s reliability? That again seems highly doubtful. It is much more likely that when you do indeed want knowledge, you want it for the

sake of true belief, either your own true belief (as in the running away story), or the true belief of others (as in the other stock story). Better perhaps, you want it for the sake of not having false belief: even if it is true that ten thousand years ago an ant was standing just where my foot is now, I do not in the least mind not believing this. In the rhetoric of the academy, things are of course otherwise. The intrinsic value of knowledge is proclaimed. But when in such contexts ‘knowledge’ does not function just as an imposing stand-in for true belief, the real thought seems to be that there is an intrinsic value in the intellectual virtue which is required for the pursuit of knowledge, or in the formal beauty of certain organized bodies of true propositions whose organization is contingent on their evidential relations.

I suggest then that thanks to the consensual nature of ‘Induction is reliable’, and to the absence of any reason for wanting to know that proposition for the sake of knowing something else, our arguments for inductive ignorance fail the negation test. N will have no reflectively indestructible desire for his missing knowledge. His reflections will make it plain to him that he has no reason for wanting either evidence-or-acquaintance or any other kind of knowledge of that proposition.

Two worries before I leave these arguments for inductive ignorance. The first is this. I have suggested that part of what saves N from wanting knowledge of Induction’s reliability is that he will believe the three propositions (A), (B) and (C). But may he not wonder if these saving beliefs of his are really true? I reply that if N does indeed have these saving beliefs, and if he does raise the question of whether they are true, then the question will answer itself. Let S stand for the three propositions (A), (B) and (C). N believes that because he believes S the argument for his inductive ignorance fails the negation test. He wonders however whether S is true. But if he does indeed believe S, then ‘Yes’ will be his answer to the self-addressed question, ‘Is S true?’. And if he does indeed believe that S, and also believes that he believes that S then ‘Yes it is’ will be his answer to the self-addressed question, ‘Is my belief that S true?’. Why so? Well, his belief that S is either dispositional or occurrent, and if it is occurrent then it will cause him to have a dispositional belief that S. But a dispositional belief that p is in part a disposition to say ‘Yes’ in answer to the self-addressed question, ‘Is p true?’. And by the same token, if he does indeed believe that S, and also believes that he believes that S, then ‘Yes it is’ will be his answer to the self-addressed question, ‘Is my belief that S true?’. A dispositional belief that you believe that p is in part a disposition to answer ‘Yes’ to the self-addressed question, ‘Do I believe that p?’. Put a dispositional belief that p together with a dispositional belief that you believe that p, and ask yourself the question ‘Is my belief that p true?’ and by the very nature of the two dispositions your answer will be ‘Yes it is’. Maybe you can stifle that answer, by somehow treating yourself as if you were someone else, like Sartre’s girl who when the man takes her hand tries to observe the event as if it isn’t really happening to

her at all, as if the hand he has taken is not really her own hand which she can herself decide brusquely to snatch away, but an alien object whose movements she cannot control. If however you have a properly existential attitude to your own attitudes then ‘Yes it is’ is the answer that will emerge.

At this point someone will say that really *N* *should* stand back and view his own beliefs in a more objective way. Suppose he believes that *S*, and believes he believes it. Then *I* won’t be surprised, if *he* answers ‘Yes it is’ when he asks himself whether he has a true belief that *S*. But there is nothing here which guarantees that *I* believe that *S*, and nothing therefore which prevents *me* from believing that if he did know that *S*, then his belief would be more likely to be true. Shouldn’t he then be prepared to put himself in *my* place, when considering whether *S* is true? Shouldn’t he be prepared to look at things in a way which would be equally acceptable to me? These scruples are inapplicable in the present case. Since ‘Induction is reliable’ is consensual, I too will believe that *S* if I entertain it.

But what about ‘I believe that *p* but of course I may be wrong’? This kind of remark is often made, with seeming sincerity, even by people who do both believe that *p* and believe that they believe it. But doesn’t it follow from the line we have taken that if you do believe that *p* you are disposed to believe that you are definitely right? A.M.McIver once set an Analysis competition in which the task was to explain how such remarks are to be taken. As McIver reported, the problem had no solvers and in fact there were only two competitors. Attempting his own solution, McIver suggested that when I say ‘I am aware that I may be wrong’ I may perhaps just be saying in an oddly misleading way ‘that, even if I have no doubt that I am right, yet, if something were to happen which I must think will not happen (namely, that I should be proved wrong), then I should alter my opinion’.¹⁷ But he was not happy with this solution:

what is required is not only that I should declare myself disposed to change my mind for good reason, but that I should actually be (or at least think myself) so disposed, and it is a fact of experience that to be so disposed, if it is possible at all, requires continuous effort and resistance to temptation; but why should I ever make this effort unless I recognise that, if I do not, then I am in danger of being left with opinions which are false? But this seems to bring us right back to our original difficulty, because it seems to imply that I must recognise that my present opinions perhaps are false.¹⁸

If I had entered McIver’s competition, I think I would have said just that ‘I believe that *p* but I am aware that I may be wrong’ is a piece of innocent diplomatic insincerity. (Cf. Wittgenstein: “‘*x* is in error” has no real point for *x*=myself”¹⁹, and ‘If there were a verb meaning “to believe falsely”, it would not have a significant first person present indicative.’²⁰)

There is one more worry. What of arguments for ignorance in which the proposition which N does not know is that 'Induction is reliable' is a consensual proposition? Given the use we made of the consensual status of 'Induction is reliable', won't N have rather a strong reason to want to know that it does indeed have this status? But how can he know this? Isn't there scope at this higher level for a sound sceptical argument for ignorance which does pass the negation test? To this I reply that just as the consensual nature of 'Induction is reliable' ensures that arguments for inductive ignorance fail the negation test, so the consensual nature of "'Induction is reliable" is consensual' ensures the failure of arguments for ignorance of the consensual nature of 'Induction is reliable'. If 'Induction is reliable' is indeed consensual, it can't be denied that it is consensual that it is consensual. A proposition is consensual if and only if (i) almost anyone who entertains it firmly believes it; (ii) almost anyone who entertains (i) firmly believes (i); (iii) almost anyone who entertains (ii) firmly believes (ii); and so on, ad infinitum. By this definition 'p is consensual' is itself consensual. This is because each conjunct i of the infinite conjunction which is equivalent to 'p is consensual' will itself be the first conjunct of an infinite conjunction which is equivalent to 'i is consensual'. Thus (i) is the first member of an infinite conjunction whose next conjunct is (ii) and which is equivalent to '(i) is consensual', (ii) is the first conjunct of an infinite conjunction whose next conjunct is (iii) and which is equivalent to '(ii) is consensual', and so on.

(3) Prospectus

'All sound sceptical arguments fail both the negation and the affirmation tests.' This is the conjecture we are trying to falsify. We are looking for counterexamples among arguments for ignorance, and we are looking more particularly for sound and sceptical arguments for ignorance which will pass the negation test. Our arguments for inductive ignorance would not serve. Though sound and sceptical, they failed the negation test. Where should we look next?

Our arguments for inductive ignorance relied on the assumption that

- (A) S knows a contingent proposition p only if either (i) he believes a proposition which is evidence for p, or (ii) p is entailed by a true proposition of the form 'S is acquainted with x'.

Maybe the sceptic can construct other sound arguments which rely on (A). Or, if not on (A), then at least on the rather stronger

- (B) S knows a contingent proposition p only if either (ia) he believes a proposition which is evidence for p, and no proposition is stronger evidence for not-p, or (ii) p is entailed by a true proposition of the form 'S is acquainted with x'.

Among these further arguments there are perhaps some which pass the negation test. An argument for ignorance has this form of conclusion: 'N does not

know that P'. P, in our arguments for inductive ignorance, was a consensual proposition. That did at least ensure that two stock reasons for wanting knowledge did not apply. If however there are any sound sceptical arguments which rely on (A) or (B), then there ought to be some in which P is non-consensual. We could look for these. An argument for ignorance in which P is non-consensual may admittedly still fail the negation test. Even in this case, reflection may still destroy N's initial desire for the knowledge which the argument says he lacks. If P is the proposition which the argument concludes that N does not know, then even if P is non-consensual N's desire to know that P will it seems still be at the mercy of his desire for P's truth. If Popper had stopped N wanting Induction to be reliable, that by itself would have stopped N wanting to know that Induction is reliable. But if there are any sound sceptical arguments which rely on (A) or (B), some of them ought also to be ones in which P's actual truth is an object of reflectively indestructible desire. This marks an area to search.

But not the only area. (A) and (B) give what we can call optional conditions for knowledge. In this they contrast with 'S knows that p only if p is true'. It is not totally eccentric to say that S knows that p only under the conditions specified by (A), and it is not totally eccentric to say that even though the conditions specified by (A) are not satisfied, still S knows that p. Likewise with (B). But 'S knows that p only if p is true' gives a non-optional condition for knowledge: it is totally eccentric to say that though p is false still N knows that p. (A) and (B) are not the only optional conditions for knowledge. There is also the condition that knowledge is closed under known entailment: if S knows that p, and knows that p entails q, then he knows that q. Also optional is the regressive condition that if S knows that p only because he believes another proposition which is evidence for p, then he already knows that other proposition. And then there are various infallibilist conditions: 'S knows that p only if "S believes that p" entails that p'; 'S knows that p only if S believes a proposition which is evidence for p and which entails p'; 'S knows that p, iff p holds in every possibility uneliminated by S's evidence, except for those possibilities which we are ignoring, and properly ignoring'.²¹ An evidence-or-acquaintance argument can be defined as any sceptical argument for ignorance which invokes no optional condition for knowledge other than either (A) or (B). Our arguments for inductive ignorance were on this definition evidence-or-acquaintance arguments. With due precautions in his choice of P, the sceptic may be able to construct a sound evidence-or-acquaintance argument which will pass the negation test. But we should also consider the possibility that sound sceptical arguments which pass the negation test can be found among arguments for ignorance which invoke optional conditions other than (A) or (B).

I shall not be considering sceptical arguments which employ senses of 'knowledge' governed by infallibilist conditions. For, I assume, such arguments will fail the negation test: infallibilist knowledge has less stringent

substitutes. But in the next two chapters I will look at arguments which employ closure-governed and regressive senses of 'knowledge'. In the next chapter I consider those neo-Cartesian arguments in which the sceptic invokes closure to derive your wider ignorance from your inability to know that you aren't the victim of a demonic system of deception. In chapter III I look at sceptical arguments for ignorance which invoke the condition that if you know that p only because you believe another proposition which is evidence for p, then you already know that other proposition. In neither chapter do I find a sound sceptical argument for ignorance which gets through the negation test. After these chapters, I return to evidence-or-acquaintance arguments. They are the topic of chapters IV and V.

Sceptics have a hostile interest not just in knowledge but in every kind of status with which we try to dignify belief. As well as arguments for ignorance, they have arguments against justified and against rational belief. Among them, perhaps, are arguments which get through the negation test. That we can consider after chapter V. For the moment, let us concentrate on arguments for ignorance.

II

DEMONIC ARGUMENTS

Some arguments for ignorance go like this:

- (1) N knows that P entails Q;
- (2) For any person S and any propositions p and q, if S knows that p, and knows that p entails q, then he knows that q;
- (3) N does not know that Q;

So: (4) N does not know that P.

This chapter is about (1)–(4) arguments in which Q stands for the negation of (D) Nothing exists but N and a demon who makes him believe whatever it is that in the actual world he does believe,

and P stands for some contingent proposition which entails the negation of (D). Such arguments I call demonic.

There are of course other sceptical arguments of the form (1)–(4). Some may be sound. Take ‘freedom’ and ‘determinism’ so that freedom is incompatible with determinism. Let Q stand for the negation of this determinism and P for the proposition that we enjoy this freedom. With luck, N will then know that P entails Q: (1) may be true. It may also be that, however lucky or clever N is, he will not know that the relevant deterministic hypothesis is false: (3) may be true. Given (2), we may then have a sound (1)–(4) argument. But demonic arguments are of more general interest. In demonic arguments P can stand for any of a vast range of contingent propositions: that I am sitting by the fire, that Mary is in Manchester, that Manchester is larger than Sheffield, and so on. There are few contingent propositions that P cannot stand for, other than that N exists, and that he thinks. And when it comes to (3), demonic arguments have an advantage. Q stands for not-(D): a cunning choice, which seems to ensure that with minimal assumptions about how we take ‘knowledge’ (3) has to be true. It is not surprising that in the latest phase of their age-old struggle to refute the sceptic, epistemologists have been so heavily engaged with (1)–(4) arguments in which Q stands for not-(D) or for the functionally similar negation of some brain-in-a-vat hypothesis.¹

‘No sound sceptical argument passes either the negation or the affirmation test.’ Counterexamples could be arguments for ignorance which pass the

negation test. Are there sound demonic arguments which pass the negation test? In section (1) I maintain that some demonic arguments are entirely sound. In section (2) I show that no sound demonic argument will pass the negation test.²

(1) The soundness of demonic arguments

Demonic arguments go like this:

- (1) N knows that P entails not-(D);
- (2) For any person S and any propositions p and q, if S knows that p, and knows that p entails q, then he knows that q;
- (3) N does not know that not-(D);

So: (4) N does not know that P.

(D) stands for

Nothing exists but N and a demon who makes him believe whatever it is that in the actual world he does believe.

It looks as if the sceptic can find or construct a sense of ‘knowledge’ on which the second and third premises of a demonic argument will be necessary truths, and on which there is nothing to prevent the first premise from being contingently true. If this is right, then given the obvious validity of (1)–(4) arguments, some demonic arguments are sound.³

What sense of ‘knowledge’ will actually do this work? Any sense governed by these three conditions:

- (CP1) For any person S, and any propositions p and q, if S knows that p, and knows that p entails q, then S knows that q;
- (K1) For any person S and any proposition p, S knows that p only if S believes that p;
- (K2) For any person S and any contingent proposition p, S knows that p only if were p false S would not believe that p.

When ‘knowledge’ is governed by these three conditions, there is nothing to prevent the first premise of a demonic argument from being contingently true. And the second and third premises will be necessary truths. (CP1), (K1) and (K2) will be necessary truths, if that is how we are taking ‘knowledge’. (CP1) itself appears as the second premise. And the third premise will be a necessary truth if it follows from ((K1) and (K2)). Which it does. By (K2), it is a necessary condition for N to know that not-(D) that, were not-(D) false, he would not believe that not-(D). In other words, it is a necessary condition that were (D) true, N would not believe that not-(D). Now either (i) N believes that not-(D), or (ii) it is not the case that N believes that not-(D). Suppose that (i). Necessarily if (i) is true then were (D) true, the demon would make N believe that not-(D). But then N would not know that not-(D). For by (K2) it is a necessary condition for him to know that not-(D) that were (D) true he would *not* believe that not-(D). Suppose, alternatively, that (ii): it is *not* the case that N believes that not-(D). Then again it follows that he does not know that not-(D).

For by (K1) it is a necessary condition for him to know that not-(D) that he does indeed believe that not-(D). So whether or not N believes that not-(D), 'N does not know that not-(D)' follows from ((K1) and (K2)). So 'N does not know that not-(D)' follows from ((K1) and (K2)).

It is easy to see that for any sense of 'knowledge' governed by (CP1), (K1) and (K2), there are sound demonic arguments. But can we be more specific about such senses? Can we give necessary and sufficient conditions?

Consider to begin with, this definition, which I derive from Nozick's truth-tracking analysis of 'the' concept of knowledge.⁴ For any person S, and any contingent proposition p, S truth-trackingly knows that p if and only if (i) p is true; (ii) S believes that p; (iii) were p false then S would not believe that p; and (iv) were p true but under circumstances slightly different from those actually obtaining then S would still believe that p. And for any person S, and any non-contingent proposition p, S truth-trackingly knows that p if and only if p satisfies (i), (ii) and some condition (iva) which requires S to arrive at his belief that p via some appropriate method such as mathematical proof.⁵ This truth-tracking sense of 'knowledge' is governed by (K1), and so far as contingent propositions are concerned it is governed by (K2). Nor does it seem to prevent the first premise of a demonic argument from being contingently true. But what about (CP1), the principle that knowledge is closed under known entailment? Anthony Brueckner has a convincing if complex argument to show that if 'knowledge' is taken in the truth-tracking way then for some values of p and q there is a possible world in which N knows that p and knows that p entails q, but still does not know that q.⁶

Let us then modify the truth-tracking sense of 'knowledge' in such a way that (CP1) no longer clashes with the other governing conditions. Let us say that S knows that p, in the modified truth-tracking sense, if and only if (a) he knows that p in the unmodified truth-tracking sense, and (b) there is no proposition q such that (i) S knows in the unmodified truth-tracking sense that p entails q, and (ii) in the unmodified truth-tracking sense, S does not know that q. Since S knows that p in this modified truth-tracking sense only if he knows that p in the unmodified truth-tracking sense, modified truth-tracking knowledge continues to be governed by (K1), and, so far as contingent propositions are concerned, continues to be governed by (K2). But, also, modified truth-tracking knowledge conforms to closure under known entailment.⁷

We don't have to suppose that 'knowledge' has a single, central ordinary sense in order to judge that this manoeuvre has generated a pretty eccentric sense of the term. But to say that an argument for ignorance employs an eccentric sense of 'knowledge' is not to say that this argument is unsound. Terminological eccentricity is quite compatible with soundness. Nor can we conclude from the terminological eccentricity of a sound demonic argument that N will not want the knowledge which it shows that he does not have, or that if he does want this knowledge then

reflection will abolish his desire. Terminological eccentricity is also compatible with success in the negation test.⁸

Suppose then that ‘knowledge’ in a demonic argument is taken in the modified truth-tracking sense. Is that really enough to secure the soundness of the argument? There is an objection about counterparts. Demonic arguments go like this:

- (1) N knows that P entails not-(D);
- (2) For any person S and any propositions p and q, if S knows that p, and knows that p entails q, then he knows that q;
- (3) N does not know that not-(D);

So: (4) N does not know that P.

They have counterparts which go like this:

- (5) N knows that P entails not-(D);
- (6) For any person S and any propositions p and q, if S knows that p, and knows that p entails q, then he knows that q;
- (7) N knows that P;

So: (8) N knows that not-(D).

Take any demonic argument, and pair it with a (5)–(8) counterpart in which N stands for the same person, and P for the same proposition. Suppose also that in both these arguments ‘knowledge’ is used in the same sense. It cannot then be that both the arguments are sound. Sound arguments have both true premises and true conclusions. If (4) is true then (7) is false, and if (8) is true then (3) is false. But why claim soundness for the demonic argument rather than for its counterpart? Presumably, because (3) is more plausible than (7). But is (3) really more plausible than (7), when ‘knowledge’ is used in our modified truth-tracking sense? Consider (3). It does indeed seem that when ‘knowledge’ is taken in a modified truth-tracking way N does not know that not-(D). Were not-(D) false, N would believe that not-(D). Of those possible worlds in which (D) is true, the ones closest to the actual world are ones in which N believes that not-(D). But now consider (7). Suppose that P stands for ‘N is sitting by the fire’. Were it false that N is sitting by the fire, then N would not believe he is sitting by the fire: of those possible worlds in which N is not sitting by the fire, the ones closest to the actual world are ones in which N does not believe that he is sitting by the fire, but rather ones in which he believes such things as that he has gone down to the cellar to get more coal, or that he has gone upstairs to bed. If the actual world contained nothing but N and the demon then possible worlds in which N is not sitting by the fire but believes he is might indeed be closer to the actual world than ones in which he is not sitting by the fire but believes he has gone down to the cellar. And you can’t claim to *know* that the actual world contains more than yourself and the demon without begging the question against a demonic argument. But that you can’t make this claim to knowledge will not prevent you from believing and agreeing that the actual world is not like this. So why reject (7)? Why isn’t the counterpart argument sound?

The answer is that if P is a contingent proposition which entails not-(D), and if 'knowledge' is taken in the modified truth-tracking sense, then either (5) or (7) is false. So the counterpart argument is unsound. (5) and (7) might both be true if 'knowledge' were taken in the unmodified truth-tracking sense. But if 'knowledge' is taken in a modified truth-tracking sense then N knows that P only if there is no proposition q such that (i) he knows in the unmodified truth-tracking sense that P entails q, and (ii) in the unmodified truth-tracking sense, N does not know that q. But there is such a proposition, namely not-(D). For if (5) is true N knows that P entails not-(D) in the unmodified as well as in the modified truth-tracking sense. And N does not know that not-(D) in the unmodified truth-tracking sense. So if (5) is true then (7) is false.

Can demonic arguments be refuted with the help of a Wittgensteinian or externalist theory of meaning? There is an Appendix at the end of the book which contains some critical remarks about this line of attack. These criticisms are of course quite inconclusive. But they are not, perhaps, entirely pointless. We are looking for sound sceptical arguments for ignorance which pass the negation test: such arguments would be counterexamples to the conjecture that all sound sceptical arguments for ignorance have conclusions whose truth value is a matter of indifference. Suppose that semantic considerations do prove all demonic arguments to be unsound. Would that make it unnecessary to consider whether demonic arguments pass the negation test? Would it be sensible for convinced externalists to skip the next section, and move on to chapter III? There is another way to look at it. Wanting what sceptics say you do not have, you struggle to refute their arguments. This you do in order to restore your equanimity. Sometimes, however, there is a stoic or want-destructive way to achieve the same end: sometimes reflection will destroy your desire for what the sceptic says you do not have. And sometimes this stoic method is less arduous than a head-on refutatory assault. This could be how things stand with demonic arguments. To show that such arguments fail the negation test is to show that there is a stoic way to overcome the disquiet that they initially induce. Even if they can be refuted with the help of an externalist theory of meaning it may be more arduous thus to refute them than to show how they fail the negation test. The Appendix may do something to support that suggestion.

A word finally on the contextualist approach to demonic arguments. According to the contextualist, one of the conventions which currently governs the use of sentences of the form 'S knows that p' is that any particular sentence of this form will express different propositions in different contexts of utterance. When the context is quotidian and commonsensical the proposition expressed will not be the same as the proposition expressed when the context is one in which some sceptical hypothesis has been mentioned. And the proposition which the sentence expresses when uttered in a sceptical context will be true only under conditions which are much more exacting than

those necessary for the truth of the proposition which the sentence expresses when uttered in an everyday context. If the sceptic advances a demonic argument then he creates a context which ensures that his argument has a true conclusion. But if the 'N does not know that P' sentence which the sceptic uses to express the conclusion of his argument had been used in a quotidian context it might well have expressed a false proposition.⁹

As a purely descriptive thesis about the conventions currently governing the use of knowledge sentences contextualism may or may not be true.¹⁰ But for present purposes, the question is immaterial. Suppose that contextualism is true, and the propounding of a demonic argument does create a context in which a sound argument will be expressed. We still want to know whether this argument passes the negation test, and more generally, whether or not the argument is trivialized by its conclusion. If on the other hand contextualism is false, and the sceptic does not create a context which ensures the soundness of the argument he advances, he still has another way to construct a sound argument for ignorance. He can innovate, take 'knowledge' in a more or less unconventional way. And we cannot assume that what he constructs won't pass the negation test.

Contextualism has been presented as a solution to The Sceptical Problem.¹¹ Needless to say, this isn't the sceptical problem that concerns us here. It isn't the problem of defending the hypothesis that no sound sceptical argument passes the negation or the affirmation test. The Sceptical Problem is a paradox. A paradox is a set of sentences which seems to express a set of mutually inconsistent propositions each of which is independently plausible. In demonic terms, the sentences would be of these forms: 'N knows that P entails not-(D)', 'For any person S and any propositions p and q, if S knows that p, and knows that p entails q, then he knows that q', 'N does not know that not-(D)', 'N knows that P'. The contextualist solution to the paradox is to insist that in the context created by our consideration of the entire set of sentences 'N knows that P' does not after all express a plausible proposition. On the contrary, it expresses an obviously false proposition. If this sentence seems to express a plausible proposition, it is because we are vaguely aware that if the sentence were used in an everyday context it would express a plausible proposition. This may solve the paradox, but it doesn't show that the sceptic can't construct a sound demonic argument which passes the negation test.

(2) The negation test

Arguments for ignorance have conclusions of this form: 'N does not know that P'. And as I explained in the last chapter, a sceptical argument for N's ignorance fails the negation test if and only if were N normally intelligent, and were he to want the knowledge which the argument concludes he lacks, then reflection would destroy his desire for that knowledge. A demonic argument

proves that when knowledge is taken in a modified truth-tracking sense N does not know that P. How could reflection destroy N's desire for this missing knowledge?

It could suggest a substitute. It could lead N to recognize that some other and more easily available kind of knowledge that P would serve him just as well. But what is the nature of the substitute? Just how does it differ from what it is supposed to replace? Among the conditions which govern modified truth-tracking knowledge there are these:

- (CP1) For any person S, and any propositions p and q, if S knows that p, and knows that p entails q, then S knows that q;
- (K1) For any person S and any proposition p, S knows that p only if S believes that p;
- (K2) For any person S and any contingent proposition p, S knows that p only if were p false S would not believe that p.

Our substitute kind of knowledge cannot be governed by all of these. If it were, it would be no more easily available than the modified truth-tracking knowledge which it is supposed to replace. For as we saw in the last section, demonic arguments work for *any* sense of 'knowledge' governed by (CP1), (K1) and (K2). And yet each of these three conditions does at first seem to be one which any satisfactory kind of knowledge must meet. In any proposed substitute one or more of these three conditions will be conspicuous by its absence. Just how is its absence to be explained away? (K1) seems essential. We want to know things because we want certain conditions to be satisfied by the beliefs upon which we act. That rationale collapses if we can know without even believing. Any knowledge which is more than a species of accidentally true belief will have to be governed by something like (K2). As Gettier showed, we can't rule out accidentally true belief that p just by requiring that S believes some proposition which is good evidence for p. And (CP1) allows us to extend our non-logical knowledge by means of our logical knowledge.

But fortunately there is another way to show that demonic arguments fail the negation test. Demonic arguments go like this:

- (1) N knows that P entails not-(D);
- (2) For any person S and any propositions p and q, if S knows that p, and knows that p entails q, then he knows that q;
- (3) N does not know that not-(D);

So: (4) N does not know that P.

(D) is the proposition that nothing exists but N and a demon who makes him believe whatever it is that in the actual world he does believe. To show that these arguments fail the negation test we first establish that (i) on reflection N won't want both to know that P and to know that P entails not-(D). Next we establish that (ii) he will on reflection want to know that P entails not-(D). Then from (i) and (ii) we infer that he won't on reflection want to know that P. How to establish (i)? Two strategies suggest themselves. One relies on the consensual nature of not-(D) and on

ENTAILMENT: For any person S and any propositions p and q, if S believes that p entails q, and if S does not want that q, then he does not want that p.

The other relies neither on ENTAILMENT nor on the consensual nature of not-(D) but rather on

INCOHERENCE: No want survives its owner's belief that it has an incoherent content.

The rest of this section falls into three parts. The first part is about the ENTAILMENT and consensus argument for (i). The second part is about the INCOHER-ENCE argument for (i). The third part is about the move from (i) via (ii) to the final conclusion that N will not on reflection want to know that P.

(2.1) *ENTAILMENT and consensus*

The ENTAILMENT and consensus argument is designed to show that on reflection N won't want it to be true both that

(A) He knows that P,

and that

(B) He knows that P entails not-(D).

It proceeds by first of all showing that reflection will ensure that N does not want it to be true that

(C) He knows that not-(D),

and then by claiming that, given what N can be expected to believe about the logical relations between (C) and ((A) and (B)), he won't want it to be true that ((A) and (B)).

Why does reflection ensure that N does not want that (C)? Part of the thought here is that not-(D) is a consensual proposition: (i) it will be firmly believed by almost anyone who entertains it; (ii) almost anyone who entertains (i) will firmly believe (i); and so on. By the argument of the last chapter this ensures the inapplicability of two stock reasons for wanting knowledge: N won't be afraid that his true belief that not-(D) will 'run away', and he won't think that his knowing that not-(D) would serve the public interest. Might he nevertheless want to know that not-(D) purely and simply for the sake of knowing some other proposition by which, as he believes he knows, not-(D) is entailed? He will indeed realize that not-(D) is entailed by numerous other propositions which he has a reflectively indestructible desire to know, such as for example that his sister is happy. It may also be that the knowledge he wants of these other propositions is knowledge in a sense which is governed by

(CP1) For any person S, and any propositions p and q, if S knows that p, and knows that p entails q, then S knows that q.

But we have to remember that the knowledge that not-(D), which according to a demonic argument N does not have, and which we aim to show that on reflection he will not want to have, is knowledge which, by that argument, it is

impossible for him to have. The sense of ‘knowledge’ employed in a demonic argument is governed not just by (CP1) but also by:

(K1) For any person S and any proposition p, S knows that p only if S believes that p,

and

(K2) For any person S and any contingent proposition p, S knows that p only if were p false S would not believe that p.

And as we saw in the last section, it is logically impossible for N to know that not-(D), in any sense of ‘knowledge’ governed by these two conditions. If N is convinced by a demonic argument whose conclusion is that he does not know that P, he will on reflection realize that the soundness of the argument depends on its taking ‘knowledge’ in a way which makes it necessarily true that he does not know that not-(D). And he will hardly want knowledge that not-(D) for the sake of knowledge of some other proposition which he knows to entail not-(D) if he also thinks that knowledge that not-(D) is in any case logically impossible for him to obtain. I admit that if N took ‘knowledge’ in a (CP1)-governed sense which was less demanding than the one employed in demonic arguments he might want knowledge that not-(D) for the sake of having knowledge of some proposition which he knew to entail not-(D). N might have a reflectively indestructible desire to know that his sister was happy, in some (CP1)-governed sense of ‘knowledge’ on which it was a merely contingent matter whether he knew that not-(D). If he realized that unless (D) was false his sister wouldn’t even exist then he might well want to know that not-(D), in this relatively undemanding sense of ‘knowledge’, purely and simply for the sake of knowing that his sister was happy. But this kind of reason for wanting to know that not-(D) is inapplicable when, as in the present case, ‘knowledge’ is so taken that knowledge that not-(D) is logically impossible. It seems then that reflection will indeed ensure that N does not want that (C).

How can it be shown that if reflection ensures that N does not want that (C) it will also ensure that he does not want that ((A) and (B))? Here the argument invokes

ENTAILMENT: For any person S and any propositions p and q, if S believes that p entails q, and if S does not want that q, then he does not want that p.

To make his demonic argument convincing, the sceptic takes ‘knowledge’ in a (CP1)-governed sense. And it will hardly escape N that, if ‘knowledge’ is taken in a (CP1)-governed sense, then ((A) and (B)) does indeed entail (C). N believes that ((A) and (B)) entails (C), and, we are supposing, reflection has ensured that he does not want that (C). So, by **ENTAILMENT**, he does not want that ((A) and (B)).

ENTAILMENT is however doubtful. It has, I grant, a certain plausibility. Cases can be imagined in which your lack of desire that q would when combined with your belief that p entails q leave you wanting at best a part of

what you believe that p must involve. You want to be an otolaryngologist like Uncle Victor. But then you realize that this has to involve your treating diseases of the throat, which you don't in the slightest want to do. You may perhaps still want to treat diseases of the ear, but that won't be enough for you to go on wanting to be a complete otolaryngologist. Mightn't you still want to be one, just for the sake of being like Uncle Victor, or for the sake of being able to treat diseases of the ear? Perhaps, but then you would after all want to treat diseases of the throat, if only as a means to something else. But, plausible though ENTAILMENT is, there are various seeming counterexamples. You want that p . But, as you realize, whatever proposition r stands for, p entails (p or r). So by ENTAILMENT, you want that (p or r). But that is absurd. If you do not want that r , you will not want that (p or r). And if there are no limitations on what r stands for, you may well not want that r . Another, and perhaps more awkward seeming counterexample would be this. Although you realize, when someone tells you, that the bottle's containing brandy entails that it doesn't contain only methylated spirits, you waste no energy in wanting it not to contain only that. But contrary to ENTAILMENT this doesn't prevent you from wanting it to contain brandy.

Can such cases be explained away? Is there some modified version of ENTAILMENT which, though it excludes such cases, does nevertheless still allow us to move from N 's lack of desire that (C) to his lack of desire that ((A) and (B))? For present purposes, we can evade these questions. We do not need to invoke ENTAILMENT in order to show that on reflection N won't want that ((A) and (B)). As we will see in the next chapter ENTAILMENT, or something like it, is in fact needed in order to show that certain regressive arguments fail the negation test. But for present purposes ENTAILMENT is superfluous. INCOHERENCE is enough.

(2.2) INCOHERENCE

We want to show that on reflection N won't want that ((A) and (B)). We have seen that in demonic arguments 'knowledge' is so used that it is logically impossible for N to know that not- (D) . So (C) is necessarily false. But we have also seen that in demonic arguments the sceptic takes 'knowledge' in a (CP1)-governed sense, and that when 'knowledge' is taken in a (CP1)-governed sense, ((A) and (B)) entails (C). Now if (C) is necessarily false and entailed by ((A) and (B)) then ((A) and (B)) is also necessarily false. And if N is convinced by a demonic argument that he does not know that P then it will hardly escape him that ((A) and (B)) is for this reason necessarily false. This makes way for

INCOHERENCE: No want survives its owner's belief that it has an incoherent content.

If, as we are assuming, reflection will lead N to believe that ((A) and (B)) is necessarily false, then, by INCOHERENCE, it will also ensure that he does not

want that ((A) and (B)). Given INCOHERENCE, we do not need to appeal to ENTAILMENT and the consensual nature of not-(D).

A word now in defence of INCOHERENCE. When in the last chapter I mentioned the Humean thought that your desire will vanish when once you are acquainted with the impossibility of its satisfaction I distinguished between logical and weaker forms of impossibility. Just by drawing the distinction we disarm one likely objection to INCOHERENCE. But there are other objections. ‘Don’t we constantly want the past to be different? Don’t such desires have obviously incoherent contents?’ This confuses the desire now to change the past, which does have an incoherent content, with the all too coherent desire for this or that not to have happened. What then about wanting to have one’s cake and eat it, in full awareness that it is logically impossible to do both? How can INCOHERENCE possibly accommodate that? We must distinguish between, on the one hand, simultaneously wanting that p and wanting that q, and, on the other hand, wanting that p and q. In the first case there are two simultaneous wants. In the second case there is a single want with a conjunctive content. Wanting to have your cake and eat it is a case of the first kind, dressed up to look like one of the second. Really, you just want to have your cake and want to eat it, which is perfectly compatible with also thinking that the one logically excludes the other. Wanting to have your cake and eat it is indeed compatible with preferring to eat it. There would be no such compatibility if there were just a single desire with a conjunctive content, for one thing is preferred to another only when there are two desires of unequal strength.

(2.3) Excluding the conjunct

Now at last for the move from not wanting that ((A) and (B)) to not wanting that (A). If it is true that thanks to INCOHERENCE N does not want that ((A) and (B)), and if some separate reflection would make N want that (B), then reflection would ensure that he does not want that (A). Why so? Isn’t not having a want with a conjunctive content entirely compatible with separately wanting each conjunct of the relevant conjunction? That I don’t like whisky and ginger ale but do like neat ginger ale doesn’t entail that I don’t also like neat whisky. Similarly, ‘N does not want that ((A) and (B)) but does want that (B)’ does not entail ‘N does not want that (A)’. I agree that it is not a general truth that if you don’t want (x and y), but do want y, then you don’t want x. I think however that if there is nothing about x which makes it more desirable when not combined with y, then it is true that if you don’t want (x and y), but do want y, then you don’t want x. And this seems to be how things are in our present case. There is nothing about knowing that P which makes this more desirable when it is not combined with knowing that P entails not-(D). And so if, as we are supposing, N does not want both to know that P and to know that P entails not-(D), but does want that

(B) He knows that P entails not-(D),
then he does not want that

(A) He knows that P.

What separate reflection might make N want that (B)? Note first that if N is indeed convinced by a demonic argument that he does not know that P, then he will think that any desire he may have to know that P will be frustrated. This being so, he will want not to have such a desire. But now, if N wants *not* to want that

(A) He knows that P,

he will indeed on reflection want to have the desire that

(B) He knows that P entails not-(D).

For, by our previous argument, and given that N does not want that ((A) and (B)), his wanting that (B) will ensure that he does not want that (A). In other words, his wanting that (B) will appear to him as a means to the end of not wanting that (A). Now even if his wanting that (B) is something that he wants only as means to this end, still on reflection he will want to want that (B). And then, on further reflection, he will actually want that (B). For, as he will realize, it is only if (B) is true that his coveted desire that (B) will not be frustrated. I do not of course suppose it to be a general truth that whenever you want to want x, reflection will lead you also to want x. A romantic youth might want to experience hopeless longing, and have not the slightest desire to possess what he would then long for. A psychological researcher might want to experience at first hand what it is like to crave a dangerous drug, but feel nothing but terror at the prospect of actually taking the drug itself. It does however seem fair to assume that if N wants to want

(B) He knows that P entails not-(D),

this isn't a case of not actually wanting what you want to want.

Thanks to INCOHERENCE, N won't on reflection want that ((A) and (B)). Separate reflection will however make him want that (B). So reflection will ensure that he does not want to know that P. So demonic arguments fail the negation test. As I conjectured, all sound sceptical arguments fail both the affirmation and the negation test. This entails that no sound sceptical argument for ignorance passes the negation test. Demonic arguments are no exception.¹²

III

REGRESSIVE ARGUMENTS

Regressive arguments for ignorance are inspired by the old sceptical idea that an infinite regress of justification is both necessary for knowledge and at the same time impossible. Their distinguishing feature is that they employ a sense of ‘knowledge’ which is governed by this condition: For any person S and any proposition p, if S knows that p only because he believes another proposition which is evidence for p, then he already knows that other proposition. Any such sense of ‘knowledge’ I call regressive, and a regressive argument for ignorance I define as one which employs a regressive sense of ‘knowledge’. In this chapter I look for sound and sceptical regressive arguments for ignorance which pass the negation test.

In section (1) I describe some regressive arguments for ignorance which are at least sound and sceptical. Arguments for ignorance have this form of conclusion: ‘N does not know that P’. In some of the kinds of arguments surveyed in (1), P can stand for any proposition at all. Others place restrictions on what P can stand for. But such restrictions do not require us to take P as a consensual proposition or prevent P from being a proposition whose truth we strongly desire. In section (2) I show that despite these advantages, the arguments surveyed in (1) are no more successful in the negation test than the demonic arguments which I considered in the last chapter.

(1) The soundness of regressive arguments

Arguments for ignorance have conclusions of the form ‘N does not know that P’. And in a regressive argument for ignorance, ‘knowledge’ is so taken that for any S and any p, if S knows that p only because he believes another proposition which is evidence for p, then he already knows that other proposition. The first part of this section is about regressive arguments which employ an evidentialist sense of ‘knowledge’. By this I mean a sense governed by the condition that for any S and any p, S knows that p only if he believes a proposition which is evidence for p. In the second part I deal with arguments which employ what I will call semi-evidentialist senses of ‘knowledge’. I did not deny the

eccentricity of the modified truth-tracking sense of ‘knowledge’ used in the demonic arguments whose soundness I defended in the last chapter. Nor do I deny the eccentricity of regressive senses of ‘knowledge’. As before, eccentricity is neither a bar to soundness nor a guarantee of failure in the negation test.

(1.1) *Evidentialist arguments*

In order to construct sound regressive arguments of the evidentialist kind the sceptic needs to employ quite a demanding sense of ‘evidence’. In particular, he must take it that q is evidence for p only if the argument ‘ q , so p ’ is not question-begging: any normally intelligent person who considered the argument and was doubtful about p ’s truth, would thereby be made equally doubtful either about q ’s truth or about the truth of the conditional ‘if q then p ’. This condition was explained in the first section of Chapter I. The sceptic must also take it that q is evidence for p only if q is true, and (to put it very crudely) only if when a proposition like q is true so usually at least is a proposition like p . Once armed with a sense of ‘knowledge’ which is both regressive and in this way evidentialist, the sceptic can take his first step. He can show that N knows that P only if there is an infinite and non-repetitive series of propositions, whose first member is P , each member of which after P is evidence for its predecessor, and each member of which is believed by N .

How can the sceptic show that regressive and evidentialist knowledge requires the existence of such a series? Since the argument ‘ P , so P ’ is question-begging, P can’t be evidence for itself. But ‘knowledge’ is being taken by the sceptic so that it is governed by the evidentialist condition that for any S and any p , S knows that p only if he believes a proposition which is evidence for p . So N knows that P only if he believes another proposition Q which is evidence for P . The sceptic is also taking ‘knowledge’ so that it is governed by the regressive condition that for any S and any p , if S knows that p only because he believes another proposition which is evidence for p , then he already knows that other proposition. So N knows that P only if he knows this other proposition Q . But by the evidentialist condition he knows that Q only if he believes another proposition R which is evidence for Q . And by the regressive condition he knows that Q only if he knows that R . And so on ad infinitum.

Why can’t the series of evidentially related propositions be repetitive? Why can’t Q be evidence for P , and R evidence for Q , and P itself evidence for R ? The difficulty is that if N knows that P , in a regressive and evidentialist sense, then he *already* knows that Q . And if he knows that Q , in a regressive and evidentialist sense, then he already knows that R . So he knows that P only if he already knows that R . If however he knows that R only because he believes that P and because P is evidence for R , then he knows that R only if he already knows that P . But it can’t be the case both that he knows that R before he knows that P , and that he knows that P before he knows that R .

What does the sceptic think is wrong with an infinite and non-repetitive series of propositions, whose first member is P, each member of which after P is evidence for its predecessor, and each member of which is believed by N? Why does the need for such a series show that N does not know that P? As many writers have pointed out, it would be rash to assume that the human mind is incapable of harbouring an infinite number of beliefs. The beliefs could after all be dispositional. More suspicious is the idea that each of the infinitely many distinct propositions after P is evidence for its predecessor. The sceptic can claim that when 'evidence' is governed by the 'no question-begging' condition, then whatever P stands for it is contingently false that there is any such series. There could obviously be an infinite and non-repetitive series of true propositions each member of which after the first *entails* its predecessor. But given the 'no question-begging' condition, 'q is true, and q entails p' is not sufficient for 'q is evidence for p'. Sosa invites us to consider an infinite sequence of 'beliefs': (P1) There is at least one real number in the interval (0, 1), (P2) There are at least two real numbers in the interval (0, 1),...As he says, 'the whole infinite chain of beliefs might be held by some believers, perhaps by all of us, given that the beliefs involved can be dispositional'.¹ He then suggests that even if no belief in the series is justified unless its successor is justified, still it is 'not at all obvious' that the infinite series may not consist of nothing but justified beliefs. It does however seem clear enough that if the (P1), (P2),...sequence is a sequence of propositions, and if 'evidence' is governed by the 'no-question-begging' condition, then it is not the case that each post-(P1) element of the sequence is evidence for its predecessor. Nobody who is doubtful about whether there is at least one real number in the interval (0, 1) is going to be helped just by the assurance that there are at least two such numbers. Nor, to take a different case, is it easy to imagine how, when 'evidence' is governed by the no-question-begging condition, I am supposed to have regressive and evidentialist knowledge that 'There is a scarlet poppy before me'. What is my evidence for this proposition? It might be 'I have an experience as of a scarlet poppy before me'. But how does the series go on? It does not seem possible to continue it. Could the third member be 'There is a scarlet poppy before me'? There is no *general* objection to saying that propositions about 'the external world' can be evidence for propositions about our experiences. If however the series is continued in this way it will become repetitive, and, as we have seen, that is excluded by the sceptic's senses of 'knowledge' and 'evidence'. The sceptic may claim that whatever P stands for, there is no infinite and non-repetitive series of propositions whose first member is P, each member of which after P is evidence for its predecessor, and each member of which is believed by N. In the sceptic's regressive and evidentialist sense of 'knowledge', N knows that P only if there is such a series. So in this sense of 'knowledge' N does not know that P.

It is also possible for the sceptic to construct sound evidentialist and regressive arguments for ignorance without denying that there is an infinite and non-repetitive series of propositions, whose first member is P, each member of which after P is evidence for its predecessor, and each member of which is believed by N. Give the sceptic the 'no-question-begging' sense of 'evidence'. He can then show that if N has regressive and evidentialist knowledge that P then there is an infinite and non-repetitive series of propositions, whose first member is P, each member of which after P is evidence for its predecessor, and each member of which is known by N. Now allow the sceptic to take 'knowledge' in a sense which, as well as being regressive and evidentialist, is activist. Activist senses of 'knowledge' are governed by this condition: For any S and any p, S knows that p only if he has performed a conscious and deliberate act of assuring himself that p is true. The act could be one of searching out propositional evidence for p, as when you set out to deduce from your normal rate of coal consumption and the latest coal bill that there is still some coal left in the cellar. Or it could be an act by which you lay yourself open to an appropriate kind of experience, as when, wondering whether there is any coal left in the cellar, you go down to have a look. To know that P, in a regressive, evidentialist, and activist sense of 'knowledge', N must have performed an infinite number of distinct acts of assurance, each of which was a searching out of evidence. The first act will be a searching out of evidence for P, the second a searching out of evidence for the proposition which is found to be evidence for P, and so on. The sceptic can argue that N will not have lived long enough to meet this condition. No one has been living forever. No conscious and deliberate act is instantaneous. And it is hard to believe that there is no minimum finite period of time which it takes to perform such an act.

(1.2) Semi-evidentialist arguments

A sense of 'knowledge' is evidentialist if and only if it is governed by the condition that for any S and any p, S knows that p only if he believes a proposition which is evidence for p. Can the sceptic construct sound regressive arguments for ignorance without employing an evidentialist sense of knowledge? Given his freedom to use 'knowledge' in unusual and demanding ways, it would be surprising if he could not. Let us say that a sense of 'knowledge' is semi-evidentialist if it makes it necessarily true that for some but not all values of p S knows that p only if he believes a proposition which is evidence for p. So for example a sense of 'knowledge' would be semi-evidentialist if, though it allowed some propositions to be known purely by acquaintance, or just by virtue of the knower's standing in a causal relation to what he knows, it nevertheless ruled out the possibility that propositions about the future are known in either of these ways, and made it necessary that when p is a proposition about the future S knows that p only if

he believes a proposition which is evidence for p . The sceptic can build sound, regressive and semi-evidentialist arguments for ignorance. This he can do in either of two ways. He can place a deductivist constraint on ‘evidence’, and say that q is evidence for p only if q entails p . Or he can employ a sense of ‘knowledge’ which, as well as being regressive and semi-evidentialist, has a so-called ‘internalist’ character. I will now consider these two possibilities in turn. P can’t stand for just any proposition, in a sound semi-evidentialist argument. But P does not have to be consensual, and it need not be a proposition for the actual truth of which we have no firm desire: regressive and semi-evidentialist arguments for ignorance need not suffer from those handicaps, when it comes eventually to the negation test. I shall in fact be looking at semi-evidentialist arguments for ignorance in which P can stand for any contingent empirical generalization of the form ‘All F s are G s’, which is about the future in the sense that at the time when, according to the argument, N does not know that p , some F s are yet to occur.

(1.2.1) Deductivism

Consider this form of argument for ignorance, in which P stands for a contingent empirical generalization of the kind just specified:

- (1) N knows that P only if he believes a proposition which is deductive evidence for P ;
- (2) No proposition is deductive evidence for a contingent universal generalization some instances of which are unobserved unless one of the conjuncts of the evidential proposition is that nature is uniform;
- (3) Any deductively valid argument for the proposition that nature is uniform is question-begging;

So: (4) N does not know that P .

Here we have something reminiscent of Hume’s famous thoughts about the nugatory role of Reason in our inferences from past experience to the future.² But the reasoning is nevertheless unimpressive. The sceptic may indeed secure the necessary truth of the form (1) premise by taking ‘knowledge’ in an appropriate semi-evidentialist and deductivist sense. But he still runs into difficulties with the form (2) premise. Why shouldn’t deductive evidence for ‘All F s are G s’ have the form ‘All F s are F 1s and All F 1s are G s’?

If however the sceptic’s sense of ‘knowledge’ is not only semi-evidentialist and deductivist, but also regressive, then he can construct a better argument for the same conclusion, an argument which does not rely on the false premise about the evidential indispensability of the uniformity of nature. As before, N knows that P only if he believes a proposition Q which is deductive evidence for P . And we can assume that Q is of the form ‘All F s are F 1s and All F 1s are G s’; it is a conjunction of generalizations partly about the future. Since ‘knowledge’ is now being taken in a

regressive sense, it is also necessary for N to know that Q. And given the value assumed for Q, and the kind of semi-evidentialist sense in which 'knowledge' is being taken, N will know that Q only if he believes a proposition R which is deductive evidence for Q. But R will also be a conjunction of generalizations partly about the future, and it will again be necessary for N to know that R. And so on ad infinitum. So N will know that P only if there is an infinite and non-repetitive series of true propositions, of which P is the first member, and each subsequent member of which is a conjunction of generalizations partly about the future, and each subsequent member of which is deductive evidence for its predecessor. And it does not look as if there is any such series.

Why should there not be such a series? The difficulty is this. P is of the form 'All Fs are Gs'. Suppose the deductive evidence for P is a proposition of the form 'All Fs are F1s and all F1s are Gs', that part of the deductive evidence for the latter is a proposition of the form 'All F1s are F2s and all F2s are Gs', that part of the deductive evidence for the latter is 'All F2s are F3s and all F3s are Gs', and so on. It looks as if we will eventually reach a class of Fns for which there is no class of F_{n+1} s such that all Fns are F_{n+1} s and all F_{n+1} s are Gs, and such that a proposition of the form 'All Fns are F_{n+1} s and all F_{n+1} s are Gs' is evidence for one of the form 'All Fns are Gs'. It is true of course that if all Fns are Gs, and some non-Fn things are Gs, then there is a further class of F_{n+1} s to which all Fns belong, and all of whose members are Gs. Think of the class composed of Fns and those non-Fn things which are Gs. Suppose for example that Fns are toadstools with red caps and Gs are poisonous things. If all red-capped toadstools are poisonous we can construct a further class to which all red-capped toadstools belong, and all of whose members are poisonous, just by adding to the class of such toadstools the class of arsenic pills. Does this mean that there is deductive evidence for 'All red-capped toadstools are poisonous'? Not if we assume that q is evidence for p only if the argument 'q, so p' is not question-begging. For an argument 'q, so p' is question-begging if any normally intelligent person who was doubtful about its conclusion would thereby be made equally doubtful either about the truth of q or about the truth of the conditional 'if q then p'. And any normally intelligent person doubtful about whether it is true that all red-capped toadstools are poisonous would thereby be made equally doubtful about whether everything which is either a red-capped toadstool or an arsenic pill is poisonous.

(1.2.2) Double regression

What happens if the sceptic abandons the deductivist constraint on 'evidence'? Can he still construct sound regressive and semi-evidentialist arguments for ignorance? Suppose, as before, that P stands for some contingent empirical generalization of the form 'All Fs are Gs', which is

partly about the future. And suppose as before the sceptic takes 'knowledge' in a regressive and semi-evidentialist sense on which it is necessarily the case that when P is taken in this way N knows that P only if he knows a proposition which is evidence for P. If we now suppose that N's evidence may be either deductive or non-deductive, then it is not clear that for N to know that P there has to be an infinite series of true propositions, of which P is the first member, and each subsequent member of which is evidence for its predecessor. If N's evidence for P can be non-deductive then it might just be a conjunction of observation statements about past instances of the generalization, and he might know this conjunction without propositional evidence. How then can the sceptic construct his argument for ignorance?

One possible move would be for him to make his sense of knowledge *doubly* regressive. A sense of 'knowledge' is regressive if and only if it is governed by this condition: For any person S and any proposition p, if S knows that p only because he believes another proposition which is evidence for p, then he knows that other proposition. Doubly regressive senses of 'knowledge' are governed not just by this condition but also by the condition that if S knows that p only because he believes another proposition which is non-deductive evidence for p then he knows *that* this other proposition is non-deductive evidence for p, and if he knows that p only because he believes another proposition which is deductive evidence for p then he knows *that* this other proposition is deductive evidence for p. This might be called an internalist sense of 'knowledge'. But given the large number of different ways in which 'internalism' and 'externalism' are currently used by epistemologists, 'doubly regressive' may be less distracting.³

If the sceptic takes 'knowledge' in a doubly regressive and semi-evidentialist sense, and if this sense also conforms to the principle of closure under known entailment, then he may well be able to construct another sound argument for ignorance. Suppose, as before, that P stands for a contingent empirical generalization partly about the future, a generalization whose form is 'All Fs are Gs'. And suppose that the sceptic takes a semi-evidentialist sense of 'knowledge' which makes it necessarily true that since P has this value, N knows that P only if he believes a proposition which is evidence for P. 'Evidence' is being used by the sceptic in the 'no-question-begging' sense which I earlier described. Here is a very rough outline of the kind of argument for ignorance that the sceptic can now construct. The argument has two parts. The first part aims to show that necessarily N knows that P only if there is a proposition X such that N knows *that* X is non-deductive evidence for P. Part two aims to show that there is no such proposition X.

The first part goes like this. Suppose N knows that P. Then N knows a proposition Q which is evidence for P. We have seen that when P is a contingent empirical generalization partly about the future there is no infinite

and non-repetitive series of propositions whose first member is P and each of whose subsequent members is deductive evidence for its predecessor. So N knows that P only if either Q is non-deductive evidence for P, or Q is deductive evidence for P and there is a series of propositions S1 whose first member is Q, each member of which after Q is evidence for its predecessor, and at least one member of which after Q is non-deductive evidence for its predecessor. If Q is non-deductive evidence for P, let X stand for Q, and if Q is deductive evidence for P let X stand for the first proposition in S1 which is non-deductive evidence for its predecessor. Since entailment is transitive, and since q is deductive evidence for p only if it entails p, X's predecessor will entail P. And for any propositions p, q and r, if r is non-deductive evidence for q and q entails p, then r is non-deductive evidence for p. So X will be non-deductive evidence for P, whether or not the same is true of Q. So N knows that P only if there is a proposition, namely X, which is non-deductive evidence for P. And since 'knowledge' is being used in a doubly regressive sense it will also be true that N knows that P only if he knows *that* X is non-deductive evidence for P.

Part two of the argument aims to show that N does not know that X is non-deductive evidence for P. Its general structure is analogous to that of demonic arguments. Here is a first, crude version:

- (1) N knows that 'X is non-deductive evidence for P' entails 'Induction is reliable';
- (2) For any person S, and any propositions p and q, if S knows that p, and knows that p entails q, then S knows that q;
- (3) N does not know that Induction is reliable;

So: (4) N does not know that X is non-deductive evidence for P.

(2) is just the principle of closure under known entailment, and since the sceptic is using a sense of 'knowledge' which conforms to that principle, (2) expresses a conceptual truth. Why (3)? Why does N not know that Induction is reliable? Because given the sceptic's sense of 'knowledge' this would require him to believe a proposition which is evidence for 'Induction is reliable'. And as we saw in Chapter I, that is ruled out by the no-question-begging condition which governs the sceptic's sense of 'evidence'. But why (1)? If 'X is non-deductive evidence for P' does indeed entail 'Induction is reliable', then nothing prevents (1) from expressing a contingent truth. Its status is comparable to that of the first premise of a demonic argument, which says that N knows that P entails the negation of the demonic hypothesis. So does the entailment hold? Suppose we just equate non-deductive with inductive evidence. P has the form 'All Fs are Gs'. So if X is non-deductive evidence for P it will be to the effect that all past observed Fs are Gs. For any p and any q, q is evidence for p only if, usually at least, when a proposition like q is true so is a proposition like p. So if X is inductive evidence for P then usually when a proposition like X is true so is one like P. So if X is inductive evidence for P then Induction is reliable.

What difference would it make if we adopted a broader conception of non-deductive evidence, and allowed for example that q can be non-deductive evidence for p when q has the form ‘ r and p is the best explanation of r ’? The sceptic may claim that a modified version of the argument will still go through. In the modified version, (1) is replaced by

- (1a) N knows that ‘ X is non-deductive evidence for P ’ entails either
 Induction is reliable’, or ‘Inference to the best explanation is reliable’,
 or ‘...is reliable’,

and (3) is replaced by

- (3a) N does not know that either ‘Induction is reliable’, or ‘Inference to the best explanation is reliable’, or ‘...is reliable’.

According to the sceptic (3a) will be true for essentially the same reason as (3), and this regardless of how the blank is filled. Whatever exactly the non-deductive method is, the no-question-begging condition will prevent N from having evidence for the reliability of that method.

To recapitulate. Regressive arguments for ignorance take ‘knowledge’ so that if S knows that p only because he believes another proposition which is evidence for p , then he already knows this other proposition. And sound arguments of this kind are not too difficult for the sceptic to construct. A sense of ‘knowledge’ is evidentialist if and only if it is governed by the condition that for any p S knows that p only if he believes a proposition which is evidence for p . It is semi-evidentialist if it restricts the condition to certain particular kinds of values of p . Sound regressive arguments for ignorance can be either evidentialist or semi-evidentialist. If the sceptic takes ‘evidence’ in a sense governed by the ‘no-question-begging’ condition, he has two ways to construct sound regressive arguments of the evidentialist kind. He can do this by denying that there is an infinite and non-repetitive series of propositions, whose first member is P , each member of which after P is evidence for its predecessor, and each member of which is believed by N . Or he can build into his sense of ‘knowledge’ a condition about the knower’s activity, require the knower actively to have assured himself of the truth of what he knows. There are also two ways for the sceptic to construct sound regressive arguments of the semi-evidentialist kind. He may place a deductivist constraint on evidence. Or he may take knowledge in a doubly regressive sense.

Now we must see whether any of these various sound and sceptical regressive arguments for ignorance will pass the negation test.

(2) The negation test

An argument for ignorance passes the negation test if and only if were N to want the knowledge which the argument concludes that he does not have then reflection would leave his desire unscathed. In the first part of this section I apply the negation test to regressive arguments of the evidentialist kind. In the second part I apply it to semi-evidentialist arguments.

(2.1) Knowledge without evidence

Regressive and evidentialist arguments for ignorance do seem to fail the negation test. There is a substitute for the regressive and evidentialist knowledge which they say that N lacks. The obvious substitute is knowledge which, whether or not regressive, is not evidentialist, not governed by the condition that S knows that p only if he believes a proposition which is evidence for p. This substitute knowledge could be either non-inferential knowledge, or it could be inferential knowledge dependent on knowledge of a non-inferential kind. It might or might not require for its definition the notion of acquaintance. Numerous relevant senses of ‘knowledge’ have been described in the enormous literature on epistemic foundationalism. A few roughly described examples will suffice.

Richard Fumerton has made an interesting suggestion about how non-inferential knowledge might involve acquaintance.

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, and certainly without possessing the ability to linguistically express that fact...Acquaintance is a relation that other animals probably bear to properties and even facts, but it also probably does not give these animals any kind of justification for believing anything, precisely because these other animals probably do not have beliefs to begin with. Without *thought* there is no truth, and without a bearer of truth-value there is nothing to be justified or unjustified.⁴

How then might acquaintance give us non-inferential knowledge or justified belief? Fumerton’s suggestion is that

one has a non-inferentially justified belief that P when one has the thought that P and one is acquainted with the fact that P, the thought that P, *and* the relation of correspondence holding between the thought that P and the fact that P. No single act of acquaintance yields knowledge or justified belief, but when one has the relevant thought, the three acts together constitute non-inferential justification.⁵

Senses of ‘non-inferential knowledge’ which are explicable without reference to the relation of acquaintance can be found in the literature on ‘reliabilism’. At its crudest, the idea is that some belief-forming mechanisms take as their input stimuli which do not include beliefs, and that non-inferential knowledge is true belief produced by a belief-forming process which is reliable in the

sense that most of the beliefs it actually produces are true. Less crudely, reliability might be defined not in terms of the actual frequency of true belief but in terms of probabilistic laws of nature relating input to true beliefs, with laws of nature being taken to support counterfactual conditionals.⁶

The inferential kind of substitute non-evidentialist knowledge would be recursively definable in terms of non-inferential knowledge: S knows that p if and only if either S has non-inferential knowledge that p, or he has non-inferential knowledge of a proposition p1 which is evidence for p, or he has non-inferential knowledge of a proposition p2 which is evidence for a proposition p1 which he believes and which is evidence for p, or...

Reflection would at least convince N that non-evidentialist knowledge that P is more easily obtainable than the evidentialist knowledge that the sceptic proves he does not have. Even if N finds on reflection that he is unable to believe in the availability of non-inferential knowledge by acquaintance, he is unlikely to doubt that he can have non-inferential knowledge in which there is a weaker connection between the belief that p and the fact that p. Reliabilists have indeed not found it easy to describe such weaker connections. But this is unlikely to prevent N from believing that they actually exist. Nor will N be prevented from believing in the availability of non-evidentialist non-inferential knowledge by the thought that he does not *know* that it exists.

More difficult is the question of whether non-evidentialist knowledge that P would be as conducive as evidentialist knowledge to what N wants. As in Chapter I, I assume that N will not on reflection want knowledge just for its own sake. For what, then, will he want it? Probably, for true belief. To know that p is, partly, to have a true belief that p. And nothing is conducive to a part of itself. But, as N will realize, his knowing that P can be conducive to his own subsequent true belief, and also to the general prevalence of true belief. If N already believes that P he will want go on believing it. For that, as he sees it, will be to go on believing the truth. Knowledge is non-accidentally true belief, and the exclusion of accidents makes for stability. He will also want others to share his own true beliefs. If these beliefs are knowledge then others can use him as an authority or source of information. When others use him as a source of information, they too acquire knowledge. And that makes *their* true beliefs stable. The question is whether evidentialist knowledge is in any of these various ways more conducive to true belief than non-evidentialist knowledge.

Having prepositional evidence for what you believe can make a special contribution to the future stability of your belief. When believing is threatened by wishful thinking the evidence can be remembered for what is believed. Having prepositional evidence can also make you a good source of information. If you believe that p, and if q is evidence for p, then by believing that q others may well be led to share your belief that p. That is perhaps all the more likely on the 'no-question-begging' sense of evidence that the sceptic employs in his arguments for ignorance. But if you know

that p in an evidentialist sense, and q is your evidence for p , then you will believe that q , and cite it as evidence. Doesn't all this give evidentialist knowledge an obvious advantage? It is good to have propositional evidence for what you know, and you have it if your knowledge is evidentialist. But if N knows that P in a regressive and non-evidentialist sense, then either he has no propositional evidence for P , or his knowledge that P requires him to know some other proposition for which he has no propositional evidence.

It seems to me that N can enjoy pretty much the same advantages even if he knows that P in a non-evidentialist sense. Suppose that at t , you have non-inferential, non-evidentialist knowledge that p . Since your knowledge is non-inferential, it does not follow just from your having this knowledge that you have propositional evidence for p . But neither does your having this non-inferential knowledge prevent you from having propositional evidence for p , or at least it does not prevent you from having evidence for the truth of a temporally modified version of the same proposition just afterwards. Suppose that at t , you know by acquaintance that you are having an experience as of a scarlet poppy before you. At this time, you have no propositional evidence for that proposition. But suppose that immediately afterwards you seem to remember the experience, and come to believe that you seem to remember it. By believing that you seem to remember the experience you are believing a proposition which is evidence that at t you had an experience as of a scarlet poppy before you. That secures your belief that you had the experience and allows you to become a source of information about your experience.

If, as we are supposing, N wants true belief, might he not want to believe the truth about *how* it is that he knows what he knows? If his knowledge that P were non-inferential, or dependent on non-inferential knowledge of some other proposition, might he not want to be aware of what it was that distinguished that knowledge from mere true belief? And is it possible for him to have that awareness, if the knowledge is non-inferential? N has non-inferential knowledge that he is having an experience as of a scarlet poppy before him. Suppose his knowledge is by acquaintance: introspective experience acquaints him with his own experience as of a scarlet poppy. If however he is aware that this is how he gets the knowledge, he will believe that he has this introspective experience. But if he does believe that he has this introspective experience then he will believe a proposition which is evidence for the proposition that he is having an experience as of a scarlet poppy, and his knowledge of the latter proposition will no longer be non-inferential. Maybe he will even know the proposition that he has the introspective experience. If so, it begins to look as if only evidentialist knowledge is compatible with complete awareness of how it is that one knows what one knows. Doesn't that give it an advantage? Once again, temporal considerations need to be brought in. At t , N has non-inferential knowledge that P . Why should N want an awareness of how he comes by that knowledge

which is simultaneous with the knowledge itself? Why can't he acquire the understanding afterwards, when it is too late for the process of acquiring understanding to disturb the object which he wants to understand?

These considerations suggest that regressive evidentialist knowledge will not on reflection seem more conducive than regressive and non-evidentialist knowledge to what N wants knowledge for. Reflection will convince N that regressive and non-evidentialist knowledge is more easily available. Can we conclude that regressive and non-evidentialist knowledge is an adequate substitute for regressive and evidentialist knowledge and hence that regressive and evidentialist arguments for ignorance fail the negation test?

There is an obstacle. N may not mind if his regressive knowledge that P is non-evidentialist. But perhaps he will want his knowledge that P to be not just regressive but doubly so, want it to be the case not just that if he knows that P by virtue of believing a proposition which is evidence for P then he knows this proposition, but also that if he knows that P by virtue of believing a proposition which is evidence for P then he knows *that* this proposition is evidence for P. There is after all an attractive lack of arbitrariness about doubly regressive knowledge: why want to know that P but not want to know what it is that enables you to know it? But will doubly regressive and non-evidentialist knowledge that P be more easily available than regressive and non-evidentialist knowledge that P? If N can know that P in a non-inferential way, then the question is irrelevant. But what if P stands for a proposition, such as a contingent empirical generalization partly about the future, which N does not think he can know in a non-inferential way? Can he in that case gain even non-evidentialist doubly regressive knowledge that P? As we saw in the second part of section (1), the sceptic can construct sound arguments for the conclusion that he has no such knowledge. For the moment, we must be content with a conditional conclusion. Regressive and evidentialist arguments for ignorance fail the negation test if and only if the same is true of doubly regressive and non-evidentialist arguments for ignorance.

(2.2) Consensual propositions

In section (1) I expounded some semi-evidentialist regressive arguments for ignorance: in these the sceptic took 'knowledge' so that for some but not all values of p, S knows that p only if he believes a proposition which is evidence for p. In these arguments P stood for a contingent empirical generalization partly about the future. As I said, the sceptic can secure the soundness of such arguments in either of two ways. He can place a deductivist constraint on 'evidence'. Or he can take 'knowledge' in a sense which is not just regressive but doubly so, a sense governed by the condition that if S knows that p only because he believes another proposition which is evidence for p, then he knows *that* this proposition is evidence for p. The deductivist constraint does I think make for obvious failure in the negation test: on reflection, N would be just as

happy to know that P in a sense which, even if it did require him to have evidence for P, allowed this evidence to be non-deductive. It is less obvious that the non-deductivist, doubly regressive arguments fail the test.

But fail they do. Let me, to begin with, assume

ENTAILMENT: For any person S and any propositions p and q, if S believes that p entails q, and if S does not want that q, then he does not want that p.

As we saw when discussing demonic arguments, this is a rather doubtful principle. But let me begin by showing that, were it true, our doubly regressive arguments would fail the negation test. After that I will confront the doubts about ENTAILMENT. I will secure the principle by adding a proviso, and will show that this addition does not obstruct the previous route to our conclusion about doubly regressive arguments.

These arguments have two parts. Part one aims to show that necessarily N knows that P only if there is a proposition X such that N knows *that* X is non-deductive evidence for P. Part two aims to show that N does not know that X is non-deductive evidence for P. Part two goes like this:

- (1a) N knows that 'X is non-deductive evidence for P' entails 'Either Induction is reliable or Inference to the Best Explanation is reliable or...is reliable';
- (2) For any person S, and any propositions p and q, if S knows that p, and knows that p entails q, then S knows that q;
- (3a) N does not know that either Induction is reliable or Inference to the Best Explanation is reliable or...is reliable;

So: (4) N does not know that X is non-deductive evidence for P.

To show that, given ENTAILMENT, this sceptical argument fails the negation test we need another two-part argument. First we show that if reflection will destroy N's desire to know that X is non-deductive evidence for P then given ENTAILMENT it will also destroy his desire to know that P. This is the easy part. For p, in ENTAILMENT, substitute

(5) N knows that P,

and for q substitute

(6) N knows that X is non-deductive evidence for P.

(5) does entail (6), if P is a contingent generalization partly about the future, and 'knowledge' is taken in the sense employed in a doubly regressive argument, and the first part of the doubly regressive argument for ignorance is sound. Reflection will convince N that the entailment holds. So if reflection destroys N's desire for the truth of (6) then given ENTAILMENT it will also destroy his desire to know that P.

The second part of the argument is more complex. For p, in ENTAILMENT, substitute the conjunction of

(6) N knows that X is non-deductive evidence for P,

and

(1) N knows that 'X is non-deductive evidence for P' entails R,

where R, in (1), stands for the relevant disjunction about reliable methods. And for q, in ENTAILMENT, substitute

(7) N knows that R.

If the second part of the doubly regressive argument is sound then reflection will convince N that (7) is entailed by the conjunction of (6) and (1). So, by ENTAILMENT, reflection will destroy N's desire for the truth of ((6) and (1)) if it will destroy his desire for the truth of (7). Will reflection destroy his desire for the truth of (7)? I claimed in Chapter I that however exactly we take 'knowledge', reflection will ensure that you do not want to know that Induction is reliable. 'Induction is reliable' is a consensual proposition: (i) Almost anyone who entertains it firmly believes it; (ii) almost anyone who entertains (i) firmly believes (i); (iii) almost anyone who entertains (ii) firmly believes (ii); and so on. The consensual nature of the proposition eliminated certain stock reasons for wanting to know that it is true. Nor did it seem that we would want to know that Induction is reliable for the sake of knowing some other proposition which we know to entail it. But what goes for 'Induction is reliable' goes also for R. It is a consensual proposition which we would not want to know for the sake of knowing some other proposition which we know to entail it. Grant then that reflection will destroy N's desire for the truth of ((6) and (1)). If we can move from here to the conclusion that reflection will destroy his desire to know that (6), then we can combine this with the conclusion of the first part of the argument and derive the overall conclusion that reflection will ensure that he does not want to know that P.

How, then, do we move from 'Reflection will destroy N's desire for the truth of ((6) and (1)) to 'Reflection will destroy N's desire for the truth of (6)'? How do we eliminate the conjunct? In the last chapter, I showed how N would be left not wanting to know the proposition which a demonic argument proved he did not know by the acquisition of a positive desire to know that this proposition entailed the negation of the demonic hypothesis. Something similar should work again. If it is true that N does not want that ((6) and (1)), and there is some separate reflection which will make him want that (1), then reflection will ensure that he does not want that (6). As I said in the last chapter, it isn't a general truth that if you don't want (x and y), but do want y, then you don't want x. But if there is nothing about x which makes it more desirable when not combined with y, then it does seem to be true that if you don't want (x and y), but do want y, then you don't want x. And there does not seem to be anything about the truth of (6) which makes it more desirable when it is not combined with the truth of (1).

What separate reflection will make N want that (1)? If N is convinced by the doubly regressive argument for ignorance that he does not know that P, then he will think that any desire he may have to know that P will be frustrated. This being so, he will want not to have such a desire, and, accordingly, he will also want not to want that (6). But now, if N wants *not* to want that (6), he will indeed on reflection want that (1). For, as we are assuming, he does not both

want that (6) and want that (1). And, as he will realize, he *can't* in these circumstances want that (6), if he also wants that (1). In other words, his wanting that (1) will appear to him as a means to the end of not wanting that (6), and hence to not wanting to know that P. Now even if his wanting that (1) is something that N wants only as a means to this end, still on reflection he will want to have this desire that (1), and then, on further reflection, he will actually want that (1). For, as he will realize, it is only if he does in fact want that (1) that his desire to have this want will not be frustrated. So if N is convinced by the doubly regressive argument for ignorance that he does not know that P, then reflection will eventually lead him to want that (1), and hence ensure that he does not want that (6).

That completes part two of the argument. Put it together with the conclusion of part one, and it follows that reflection will destroy N's desire to know that P. Given ENTAILMENT, our doubly regressive arguments for ignorance fail the negation test.

But, as I said, ENTAILMENT is a doubtful principle. There do seem to be cases in which you believe that p entails q, and want that p, but do not want that q. Are these seeming counterexamples genuine? Can we, without too much damage, modify the principle? Taking a cue from my previous treatment of demonic arguments, you might suppose that we can evade these questions. We do not need ENTAILMENT in order to show that N won't on reflection want that ((6) and (1)). We can reach that conclusion just by invoking INCOHERENCE. Unfortunately, this is not so. In demonic arguments the parallel conjunction had as conjuncts 'N knows that P' and 'N knows that P entails the negation of the demonic hypothesis'. And this conjunction was necessarily false because it entailed the necessarily false proposition that N knows the negation of the demonic hypothesis. But nothing necessarily false is entailed by ((6) and (1)): only the contingent proposition that N knows that R. Let me then look more closely at ENTAILMENT.

Some counterexamples to ENTAILMENT lose their plausibility when we distinguish between external and internal negation, between wanting that not-q and not wanting that p. You want to play the lottery. And, of course, you want not to have a high chance of losing. But you come to realize that the lottery is so constructed that necessarily if you play you do have a high chance of losing. This won't automatically stop you wanting to play the lottery. Doesn't that falsify ENTAILMENT? No, because ENTAILMENT does not say that if S believes that p entails q, and if S wants that not-q, then he does not want that p. It says that if S believes that p entails q, and if S does not want that q, then he does not want that p. Wanting not to have a high chance of losing is not the same as not wanting to have a high chance of losing. Wanting that not-q is compatible with wanting that q, and if you do indeed want that q, your belief that p entails q will have no tendency to prevent you from also wanting that p. Even though you are utterly unlikely to want to have a high chance of losing the lottery either for its own sake or for the sake of anything other than having

some chance of winning the lottery, and even though you want not to have a high chance of losing, you may in fact also want to have a high chance of losing the lottery if you see this as a necessary condition for having even a small chance of winning.

But we are still left with the seeming counterexamples which were mentioned in the last chapter. One of them relied on the principle that any proposition entails the disjunction of itself and any other proposition. You want that p . You realize that whatever proposition r stands for, p entails $(p \text{ or } r)$. So by ENTAILMENT, you want that $(p \text{ or } r)$. But that is absurd. If you do not want that r , you will not want that $(p \text{ or } r)$. And if there are no restrictions on what r stands for, you may well not want that r . To this we can add a case which relies on the principle that a necessary proposition is entailed by any proposition. You want Mary to be happy. You come to believe, perhaps rashly, that Goldbach's Conjecture is a necessary proposition. Realizing that a necessary proposition is entailed by any proposition, you come to believe that the Conjecture is entailed by 'Mary is happy'. But you don't want the Conjecture to be true. Professor X claims to have proved it, and you don't want him to have succeeded. By ENTAILMENT, your desire for Mary's happiness will evaporate, which is, of course, absurd. In the other seeming counterexample which I mentioned in the last chapter someone makes you realize that the bottle's containing brandy entails that it doesn't contain only methylated spirits. You waste no energy in wanting it not to contain only methylated spirits but, contrary to ENTAILMENT, this doesn't prevent you from wanting it to contain brandy. To this we can add a case in which you realize that your knowing that p entails that you do not know that not- p , but aren't at all interested in merely not knowing that not- p . What does interest you is knowing that p , and contrary to ENTAILMENT, your not wanting not to know that not- p doesn't stop you wanting to know that p .

We might try taking it that in ENTAILMENT 'p entails q' means not just that it is logically impossible for p to be true and yet q not true, but rather that there is an 'analytic entailment' between the two propositions, where p analytically entails q if and only if p strictly implies q and q contains no non-logical concepts which p does not contain.⁷ But though that would exclude the brandy and Goldbach counterexamples, it would still leave us with the counterexample in which you don't want not to know that not- p . 'You don't know that not- p ' contains no non-logical concepts that 'You know that p ' does not contain. It is better, I think, just to add a proviso, and say that for any person S and any propositions p and q , if S believes that p entails q , and if he does not want that q , then he does not want that p , the exception being that if S does not want that p for the reason that there is a logically sufficient but not necessary condition for q 's truth which he does not want to be fulfilled, then he may after all still want that p while believing that p entails q . This proviso seems to cope with the outstanding counterexamples. The brandy case. Even though you want the bottle to contain brandy, and you realize that its doing so entails that it does

not contain only methylated spirits, you aren't particularly interested merely in its not containing methylated spirits. This is because you realize that there are possible ways of its not containing only that substance which you don't in the least want to be actual. So for example it is a sufficient but not necessary condition for its not containing only methylated spirits that it contains nothing but milk. The knowledge case. You want to know that *p*, and believe that necessarily if you know that *p* then you don't know that not-*p*. But you are not particularly interested merely in not knowing that not-*p*, because you realize that there are possible ways of you not knowing that not-*p* which you don't want to be actual. It is for example a sufficient though not necessary condition for not knowing that not-*p* that you are not justified in believing that not-*p*. Not being justified in believing that not-*p* does not interest you because you believe that *p*, and want to know that *p* so as to be able to convince people that *p* is true, and you realize that your not being justified in believing that not-*p* is entirely compatible with *p*'s being false. In the same way, an arbitrary *r* is a logically sufficient but not necessary condition for '*p* or *r*', and, as likely as not, not a proposition which you want to be true. And that Professor X has proved Goldbach's Conjecture is an unattractive logically sufficient but not necessary condition for the truth of Goldbach's conjecture.

I showed that, given ENTAILMENT, doubly regressive arguments for ignorance fail the negation test. What would have happened if ENTAILMENT, in that reasoning, had been modified in the way now proposed? In the first part of the reasoning I maintained that since, for our chosen value of *P*,

(5) N knows that *P*

entails

(6) N knows that *X* is non-deductive evidence for *P*,

and since, on reflection N will believe that this entailment holds, it follows by ENTAILMENT that if reflection will destroy N's desire that (6) it will also destroy his desire to know that *P*. In the second part of the reasoning I was concerned with

(7) N knows that *R*,

(6) N knows that *X* is non-deductive evidence for *P*

and

(1) N knows that '*X* is non-deductive evidence for *P*' entails *R*,

where *R*, in (1) and (7), stands for the relevant disjunction about reliable methods. I maintained that since reflection will ensure that N does want that (1), it will, if it ensures that he does not want that ((6) and (1)), also ensure that he does not want that (6). And I further maintained that since reflection will ensure that N does not want that (7), and since (7) is entailed by ((6) and (1)), and since, on reflection, N will believe that this entailment holds, it follows by ENTAILMENT that reflection will ensure that N does not want that ((6) and (1)). So far as I can see, our qualification of ENTAILMENT does no damage to the original argument. The second part of the argument will still work provided that it is not the case that N's reason for not wanting that (7) is that there is

some sufficient but not necessary condition for (7) which he doesn't want to be true. But this condition is satisfied if as we are supposing it is the consensual nature of R which prevents him wanting to know that R. The first part of the original argument will still work provided that it is not the case that N's reason for not wanting that (6) is that there is some sufficient but not necessary condition for (6) which he doesn't want to be true. This condition is also satisfied if as we are supposing the reason why N doesn't want that (6) is that he doesn't want that ((6) and (1)) but does want that (1). I conclude then that even though ENTAILMENT must be qualified, still doubly regressive arguments for ignorance fail the negation test.

We still have no counterexample to our general conjecture about sceptical arguments.

IV

EVIDENCE-OR-ACQUAINTANCE ARGUMENTS

Let me return to evidence-or-acquaintance arguments. They were defined in chapter I. An evidence-or-acquaintance argument is a sceptical argument for ignorance which does not rely on any optional condition for knowledge other than either

(A) S knows a contingent proposition p only if either (i) he believes a proposition which is evidence for p, or (ii) p is entailed by a true proposition of the form ‘S is acquainted with x’,

or the rather stronger

(B) S knows a contingent proposition p only if either (i) he believes a proposition which is evidence for p, and no proposition is stronger evidence for not-p, or (ii) p is entailed by a true proposition of the form ‘S is acquainted with x’.

As optional conditions, (A) and (B) contrast with ‘S knows that p only if p is true’. It is utterly eccentric to say that though p is false still p is known. It is less eccentric to use ‘knowledge’ in a way governed by (A) or (B). Evidence-or-acquaintance arguments contrast with demonic and regressive arguments, which rely on other optional conditions for knowledge. The sceptic was unable to construct sound demonic or regressive arguments which pass the negation test. Can he do better with evidence-or-acquaintance arguments?

One might easily think he can. Arguments for ignorance have this form of conclusion: ‘N does not know that P’. My first chapter was about evidence-or-acquaintance arguments in which P stands for the proposition that Induction is reliable. Sound though these arguments were, they failed the negation test. As we can say, they weren’t negation-proof. Part of what made them non-negation-proof was the consensual nature of ‘Induction is reliable’: (i) almost anyone who entertains it will firmly believe it; (ii) almost anyone who entertains (i) will firmly believe (i); (iii) almost anyone who entertains (ii) will firmly believe (ii); and so on. The consensual nature of ‘Induction is reliable’ did at least ensure the inapplicability of two stock reasons for wanting knowledge. If the sceptic wants to construct a sound and negation-proof evidence-or-acquaintance argument he would then be wise to choose a non-consensual value of P. It also seems that your desire to know a proposition is at the mercy of your desire for

its actual truth. So the sceptic should additionally try to ensure that P is a proposition for the actual truth of which N has a reflectively indestructible desire. It would be surprising if no sound evidence-or-acquaintance argument could be constructed in which P met these two conditions. Isn't philosophy itself replete with non-consensual doctrines whose truth we are quite unable to establish through acquaintance or by adducing evidence? Isn't it unlikely that every such doctrine is one which reflection would stop us wanting to be true? The general conjecture starts to look precarious.

But things look different when we try to be more specific. The threat begins to fade. The first section of this chapter is a brief discussion of the prospects for a sound and negation-proof evidence-or-acquaintance argument in which P stands for a doctrine of the objectivity of values. Surprisingly enough, the prospects seem poor. Yet more surprisingly, the sceptic does not even seem able to construct a sound and negation-proof evidence-or-acquaintance argument in which P is the doctrine that God exists. His difficulties are described in section (2). Chapter V is a more extended discussion of evidence-or-acquaintance arguments in which P stands for a certain anti-Idealist doctrine.

(1) The objectivity of values

Can the sceptic build a sound and negation-proof evidence-or-acquaintance argument in which P is a doctrine of the objectivity of values? I distinguish between two doctrines which might answer to that description. First there is the thesis that to evaluate is in part at least to assert a proposition, and that some at least of the propositions thus asserted are true. This I call the doctrine of evaluative truth. There is also what I will call the doctrine of dynamic unity. This is a thesis about the conditions under which like desires, aversions or intentions will be expressed when like objects are evaluated by different people or by the same person at different times. I begin with the doctrine of evaluative truth.

(1.1) Evaluative truth

How can the sceptic build a sound evidence-or-acquaintance argument in which P stands for the doctrine of evaluative truth? Here is, one possibility. He finds or himself devises a sense of 'evaluation' on which the assertion of a proposition is no part of what it is to evaluate, and evaluation is just the expression of a desire, or an aversion, or an intention. On this sense of 'evaluation' the doctrine of evaluative truth is false. So N does not know that it is true. With an appropriate definition as his premise, the sceptic has no difficulty in constructing a sound argument for ignorance in which P is the doctrine of evaluative truth. And he will probably have no difficulty in presenting his reasoning as an evidence-or-acquaintance argument.

A sense of 'evaluation' on which the assertion of a proposition is no part of what it is to evaluate is not a sense on which, when he evaluates, the evaluator cannot *also* assert a proposition. You say, of the hawthorn log, 'It's so gnarled!', and your tone of voice makes it plain that you admire the log, and do not want it to be burned.¹ When you say that the log is gnarled you do of course assert that it has a knotted and twisted character. But the sceptic can, I assume, find or devise a sense of 'evaluation' on which that assertion is distinct from your actual evaluation of the log, an evaluation which consists just in the expression of your admiration, an admiration itself analysable in terms of desire. It will be objected that at least in serious ethical cases, this kind of disentangling of the evaluative from the factual is impossible. I shall simply assume that this is wrong.² But it will also be objected that even if the evaluator says something more abstract than that the log is gnarled he cannot but say something true or false so long as he uses a declarative sentence to make his evaluation. He cannot avoid asserting a proposition, even if he just says 'The log is good'. Just as there is nothing more to the truth of 'Snow is white' than that snow is white, so there is nothing more to the truth of 'A is good' than that A is good. On any sense of 'evaluation' some value judgements will have to be made by using sentences which can be substituted for p in the schema "'p" is true iff p'. The schema gives us what Blackburn calls a ladder of philosophical ascent: 'p', "it is true that p", "it is really and truly a fact that p"...,...none of these terms...marks an addition to the original judgement.³ 'A is good', 'it is true that A is good', "'A is good" represents the facts', "'A is good" conforms to the eternal normative structure of the world'... 'We can add flowers without end.'⁴ Since a sentence of the form 'A is good' can be substituted for p in the schema "'p" is true iff p', this sentence will have a truth condition, and if the sceptic's evaluator says that A is good then he must say something true or false, must assert a proposition. This objection cannot be quite so speedily dismissed.

But the sceptic can nevertheless dismiss it. He can reply that *if* 'A is good' has a truth-condition, then certainly one way to give the truth-condition of that sentence is to write down the sentence itself, but that at least so far as he is concerned this does not settle the question of whether 'A is good' has a truth condition. At least in *his* usage of 'truth' and 'falsity', a sentence can be used to say something true or false only if it can be used to manifest some *belief* which the user has. And the sense of 'evaluation' on which we never assert propositions when we evaluate particular objects is according to the sceptic a sense on which when we evaluate we do not manifest beliefs but rather express desires, aversions or intentions. You judge that A is good. In the sceptic's sense of 'evaluation', you do not then manifest any attitude other than a desire for A to exist. You judge that A is bad: the only attitude you manifest is an aversion to A's existence. You judge that she ought to ring him up: the only attitude you manifest is an intention to do something similar if you are in a situation of that kind. It may well be that there is nothing in the least idiosyncratic about the sceptic's insistence on this usage of 'truth' and 'falsity', and that the

conventions governing our actual language make for an analytical tie between asserting something which is true or false and having a belief whose content can be captured by means of the sentence used. But the sceptic does not have to prove his normality in order to set up his argument.⁵

The sceptic says that in his sense of ‘evaluation’, the evaluator does not manifest beliefs but rather expresses desires, aversions or intentions. But can’t desires or aversions actually be beliefs? That some desires *are* beliefs is a view which has in fact been advanced by certain philosophers who, while they want to maintain the ‘expressivist’ doctrine that we do or should evaluate only in the sense of expressing desires, aversions or intentions, or emotions, or imperatives, are nevertheless reluctant to accept that when in the ordinary way we talk of moral opinions or beliefs we are victims of inconsistency or superstition.⁶ Here the sceptic can say that there is something that it is like occurrently to desire, or to feel an aversion, and something it is like to form an intention, and that if there is anything it is like to have an occurrent belief, what *that* is like is different.

It is then possible for the sceptic to find or devise a sense of ‘evaluation’ which yields a sound argument whose conclusion has the form ‘N does not know that to evaluate is partly at least to assert a proposition, and that some of the propositions thus asserted are true’. Will this argument pass the negation test? Why, even initially, should N be alarmed by this conclusion? Does he perhaps think that unless evaluation is at least partly the assertion of a proposition then there can be no such thing as moral reasoning? Must we remove that worry, in order to show that the argument does after all fail the negation test? If so, how? Would you on the sceptic’s understanding still be making a *logical* mistake when you said for example that it isn’t wrong to get Mary to tell lies even though lying is wrong, and if lying is wrong then it is wrong to get someone else to tell lies? Or would you just be manifesting a self-paralysing set of attitudes?⁷ Fortunately, such questions can be set aside. The sceptic does nothing but deduce his conclusion from his definition of ‘evaluation’. Given that definition, his conclusion is necessarily true. So if N does want that knowledge of the doctrine of evaluative truth which the sceptic’s argument shows that he does not have, then his desire has an incoherent content. This opens the door for

INCOHERENCE: No want survives its owner’s belief that it has an incoherent content.

Reflection will convince N that the content of his desire for knowledge is incoherent. And so, by INCOHERENCE, reflection will destroy the desire itself. The sceptic’s argument fails the negation test.

The moral is that if the sceptic wants to build a sound and genuinely negation-proof evidence-or-acquaintance argument in which P is the doctrine of evaluative truth then he must look for or himself devise a sense of ‘evaluation’ on which it is contingently rather than necessarily false that N knows that this doctrine is true. This will be a sense on which to evaluate *is* in

part at least to assert some proposition, and it will be either a *purely assertive* or a *doubly attitudinal* sense. A purely assertive sense of 'evaluation' is one on which evaluation is nothing but the assertion of some proposition, and which does not make the expression of desire, aversion or intention any part of what it is to evaluate. A doubly attitudinal sense of 'evaluation' is one on which to evaluate is to do two things, one to assert a proposition and the other to express a desire, aversion or intention. Purely assertive and doubly attitudinal senses can be contrasted with *purely dynamic* senses, on which evaluation is nothing but the expression of desire, aversion or intention. (In the standard somewhat poverty-stricken terminology, the purely dynamic and doubly attitudinal senses are 'internalist', and the purely assertive sense is 'externalist'.⁸)

Suppose the sceptic takes the purely assertive option. He might then see what he can do with the stipulation that to evaluate something is to assert a conditional to the effect that in certain circumstances everyone would have a certain dynamic attitude towards it, where a dynamic attitude is a desire, aversion or intention. Conditionals of this kind have in fact often figured in would-be analyses of the ordinary concept of a value judgement, or of the ordinary concept of rightness. According to David Lewis it is an 'unobvious analytic truth' that something is (positively) valuable if and only if we would be disposed under conditions of the fullest possible imaginative acquaintance to desire to desire it.⁹ Frank Jackson makes a rather similar suggestion about our ordinary concept of rightness: 'to believe that something is right is to believe in part that it is what we would in ideal circumstances desire', where what we would desire in ideal circumstances is 'perhaps...what we would desire when our first-order desires square with our reflective second-order desires, or when our desires square with what we would converge on stably desiring after reflection, perhaps taking into account the desires of our community'.¹⁰ Suppose the sceptic stipulates that evaluation is the assertion of some such universal conditional. Can he then build a sound evidence-or-acquaintance argument in which P is a contingent doctrine of evaluative truth? It may seem that he can. The diversity of human nature may seem to ensure the falsity of any such universal conditional. There is nothing such that even under conditions of the fullest possible imaginative acquaintance absolutely everyone would have the same dynamic attitude towards it. Will his argument pass the negation test? It looks as if it might. We would like there to be agreement in desires, aversions and intentions.

Consider now the sceptic's doubly attitudinal option: he looks for or himself devises a doubly attitudinal sense of 'evaluation', a sense on which the evaluator does two things, one of which is to express a desire, aversion or intention, and the other of which is to assert a proposition. I am not here thinking of cases in which one and the same sentence is used on some occasion both to make a value judgement and to do something else, as in the case, already mentioned, when you say, of the hawthorn log, 'It's so

gnarled!’, and your tone of voice reveals that you are evaluating the log as well as saying how knotted and twisted it is. I am supposing rather that the sceptic might take it that when a sentence is used to evaluate something two distinct attitudes are involved in the evaluation itself. He might take it that to evaluate a particular object A is (i) to express a desire or aversion or intention which is directed towards A, and (ii) to assert a proposition to the effect that (a) your knowledge that A is F caused you to have this desire or aversion or intention towards A, and (b) for any person S and any object x, if S knew that x was F he would have an attitude towards x which is like the attitude which you have towards A. You judge it a good thing that Mary went to see John. You express your desire for her to have done this. You may also assert the proposition that (a) what made you want her to have done it was your knowledge that she was making someone happy, and (b) anyone who knew that an action would make someone happy would want that action to be done.¹¹ Here again it seems at first as if this might allow the sceptic to construct a sound evidence-or-acquaintance argument in which P is a doctrine of evaluative truth. Human diversity will ensure the falsity of the universal conditional. And again it seems that, thanks to our desire for agreement in desires, aversions and intentions, the sceptic may have here an argument which passes the negation test.

But despite all this, I do not think that our general conjecture about sceptical arguments can be falsified by a negation-proof evidence-or-acquaintance argument in which P is the doctrine of evaluative truth. Even if the sceptic does take a purely assertive or doubly attitudinal sense of ‘evaluation’, and does take the evaluator to assert a universal conditional about dynamic attitudes, still he won’t be able to construct a sound and negation-proof argument. Once we have had a look at arguments for ignorance in which P stands for a doctrine of dynamic unity, the difficulty will become plain.

(1.2) *Dynamic unity*

The doctrine of dynamic unity is a thesis about the conditions under which like desires, aversions or intentions will be expressed when like objects are evaluated by different people or by the same person at different times. To define it, I need the notion of a dynamically objective value judgement.

Suppose we take ‘evaluation’ in either a purely dynamic or a doubly attitudinal sense: whether or not to evaluate is in part at least to assert a proposition, it is, in part at least, to express a desire, aversion or intention. You judge that A is good; in part at least, you thereby express your desire for A to exist. You judge that A is bad; in part at least you express your aversion to A’s existence. You judge that she ought to ring him up; in part at least you express your own intention to do something similar should you find yourself in a similar situation. Let us also take it that when you thus evaluate an object,

and express a desire or aversion or intention, it is a belief about the object which has caused you to have that attitude. There is a causative belief which underlies your evaluation. Will this belief cause your actual evaluation? Perhaps not. It depends on whether we define 'evaluation' so that as well as expressing a dynamic attitude, the evaluator must assert the content of the belief which underlies his evaluation. If we do define 'evaluation' so that the evaluator must assert the content of this belief, then the belief itself could not cause the evaluation without causing something of which it would itself be an essential part. And that seems paradoxical. But the causative belief will in any case cause the desire or aversion or intention which the evaluator expresses. If A is the particular object which you in this sense evaluate, the content of the causative belief will be a proposition of the form 'A is F'. Call F the causative property which underlies your evaluation. We can now define a dynamically objective value judgement as any value judgement about a particular object such that (i) the evaluator knows that the object has the causative property F, and (ii) for any person S and any object x, if S knew that x was F he would have an attitude towards x which is like the evaluator's desire, aversion or intention towards the particular object that he evaluates. For a value judgement about A to be dynamically objective, it is not necessary for the evaluator actually to assert that he knows that A is F or that his knowledge that A is F caused him to have the desire or aversion or intention which he expresses. Nor is it necessary for him actually to assert the universal conditional specified by (ii). For a value judgement to be dynamically objective, it is not necessary for the evaluator to assert a true proposition. It is not even necessary for him to assert a proposition.

We can now define the doctrine of dynamic unity as the thesis that (1) some value judgements are dynamically objective, and (2) when a dynamically objective value judgement is made about some particular object only historical contingencies or purely personal deficiencies will prevent people from possessing the knowledge which according to the relevant universal conditional is enough to ensure that when like objects are evaluated like desires, aversions or intentions are expressed.

If this doctrine is true, then we can perhaps hope for a measure of agreement in the desires, aversions or intentions which different people express when they make value judgements. Disagreement will be to some extent contingent on ignorance; there will be a certain scope for non-coercive harmony. And once we recognize that the doctrine of dynamic unity has this consequence we may want it to be true. We may also want to know that it is true. Certainly it is not a consensual doctrine, knowledge of which will for that reason strike us as superfluous. There seems then to be a chance that if the sceptic can construct a sound evidence-or-acquaintance argument, in which P stands for the doctrine of dynamic unity, his argument will pass the negation test.

Only if there are dynamically objective value judgements does N know the truth of the doctrine of dynamic unity. Can it be shown that there are no dynamically objective value judgements? Suppose N judges that some particular object A is good, thereby expressing a desire for A to exist, and that this desire is caused by N's belief that A is F. F is the causative property for that judgement. N's value judgement will then be dynamically objective only if (i) N knows that A is F, and (ii) for any person S and any object x, if S knew that x was F he would want x to exist. One might well suppose that whatever F stands for the relevant universal conditional will be false. That seems to be ensured by the fact of human diversity. Whatever F stands for, there is somebody, somewhere, who would not want x to exist even if he knew that x was F. What if F stands for a *sui generis* property, for which no name suggests itself other than the name of goodness itself, and it is by a special kind of experience that N knows that A is F, a kind of experience for which no name suggests itself other than the name of moral or aesthetic experience? Still there will be somebody, somewhere, who is hostile or indifferent to what he knows to be F. How though do we distinguish hostility or indifference from lack of understanding? How do you know that people who say they hate Mozart hear what other people hear? What if Fness is a property which cannot even be grasped by those without special training or a special inborn gift, so that whenever it appears that some person knows that x is F and yet does not want x to exist we cannot know that the appearance of knowledge is anything but deceptive, cannot know that really this person understands what Fness is? Surely there are ways of finding out. What if F stands for a totally comprehensive conjunctive property, constructed out of all the non-conjunctive causative properties which ever have underlain or ever will underlie the evaluation of particular objects as good? Won't *that* ensure the truth of the universal conditional? No doubt it will, but only at the price of ensuring that N does not know that A is F, which he must do if he is to make a dynamically objective judgement that A is good. It does then seem that whatever F stands for, human diversity ensures that either (i) or (ii) is false. The argument can be generalized from judgements of goodness to all other value judgements about particular objects.

Suppose that the sceptic can thus construct a sound evidence-or-acquaintance argument for N's ignorance of the doctrine of dynamic unity. Will his argument pass the negation test? That one may doubt. The doctrine of dynamic unity is pretty strong. If it is true, then there are true universal conditionals linking knowledge and dynamic attitudes. How bad would it be if, instead of true universal conditionals of this kind, there were true conditionals about *almost* everyone, or about *most* people, or even just about most people in the evaluator's own community? Could we not content ourselves with the truth of something weaker, and weak enough to make the sceptic's appeal to human diversity irrelevant? Perhaps. But I do not think

that we should yet concede that the original diversity-based argument is both sound and properly sceptical.

There is a doubt about the soundness of the argument. The appeal to facts about human diversity seems to assume that if a value judgement is dynamically objective then the relevant universal conditional is contingent. N evaluates A, thereby expressing a desire, aversion or intention. F is the causative property which underlies his evaluation. N's value judgement is dynamically objective if and only if (i) he knows that A is F, and (ii) for any person S and any object x, if S knew that x was F he would have the same kind of desire, aversion or intention towards x that N has towards A. If (ii) is contingent then perhaps the facts of human diversity do show that either (ii) is false or (i) is false. But if (ii) is non-contingent then human diversity won't make it false, and nor is it obvious that (i) is false if (ii) is necessarily true. The sceptic might try to meet this objection by stipulating that in a dynamically objective value judgement the relevant universal conditional must be contingent, and then redefining the doctrine of dynamic unity. If however the soundness of the argument is preserved by this manoeuvre then the argument ceases to qualify as properly sceptical. As I stressed at the beginning of chapter I, not all arguments for ignorance are sceptical. An argument for ignorance is sceptical only if it is one in which philosophy does some work. And in the present argument, philosophy would do no real work. Given the premise about human diversity, which philosophy itself can do nothing to establish, it would be obvious enough that the doctrine of dynamic unity was false and so not known. No philosophical explication would be needed to show that N did not know the doctrine of dynamic unity. The general conjecture that presently concerns us is that all sound sceptical arguments fail both the negation and the affirmation tests. We are looking for counter-examples in the shape of sound and negation-proof sceptical arguments for ignorance, and in this chapter we are looking more particularly for sound and negation-proof evidence-or-acquaintance arguments, which we have defined as a kind of sceptical arguments for ignorance. A non-sceptical argument for ignorance can't serve as such a counterexample.

Despite this objection, it may still be possible to construct a sound and properly sceptical argument for ignorance in which P stands for a doctrine of dynamic unity. But not, I think, an argument which will pass the negation test.

How, to begin with, could we build an argument both sound and sceptical? Here is one possibility. Let P stand for the doctrine of dynamic unity, and consider the universal conditionals which are referred to in the definition of a dynamically objective value judgement. Perhaps it can be shown that if N knows that P then either (a) he knows that some such universal conditionals are contingently true, or (b) he knows that some such universal conditionals are necessarily true. It might then be argued that both (a) and (b) are false. To show that (a) is false we appeal as before to the facts of human diversity. In this part of the argument philosophy does no work. But philosophy is needed

in order to show that (b) is false, and that is enough for the argument to qualify as properly sceptical when taken as a whole. How, then, does the sceptic show that (b) is false? Perhaps he can appeal to a general psychologistic doctrine about modality.

What is the meaning of 'p is necessarily true'? How does it differ from the meaning of 'p is true'? As the sceptic may say, the only clear answer is psychologistic: 'p is necessarily true' means that p is true and that either p's negation is unintelligible or p is derivable from a proposition whose negation is unintelligible, where 'p is derivable from q' means that there is a series of conditionals each of which after the first has its predecessor's consequent as antecedent, the first of which has q as antecedent, the last of which has p as consequent, and which is such that each conditional has a negation which is unintelligible. But now there is an epistemological problem. Psychologically interpreted, 'p is necessarily true' does not seem to be knowable by acquaintance. But nor does there seem to be evidence for its truth, or not at least on a sense of 'evidence' which is governed by the condition, which we have already encountered, that q is evidence for p only if the argument 'q, so p' is not question-begging. What is the difficulty about evidence? If 'p is necessarily true' is given a psychologistic interpretation, then there will not be evidence for its truth unless there is evidence for some proposition of the form 'r is unintelligible'. But 'r is unintelligible' means 'Everyone is unable to understand r', which has form the 'All Fs are Gs'. If then the requisite evidence does exist it will presumably be inductive evidence, in the shape of a proposition of the form 'All of the many Fs so far examined have been Gs'. But in the present instance this kind of evidence seems unavailable.

Why so? For this reason. Suppose that if an F is a G then you can tell that it is a G, but if an F is not a G then you cannot tell that it is not. Then there is not much point in your arguing that since all of the many Fs so far examined have been Gs, all Fs are Gs. The argument will in fact be question-begging. Any normally intelligent person who considered the argument and was doubtful about the truth of the conclusion would thereby be made equally doubtful about the truth of the premise. So there is inductive evidence for 'Everyone is unable to understand r' only on this condition: if someone is unable to understand that r you can tell that he lacks this ability, but also, if someone is able to understand that r you can tell that he has this ability. Suppose then that you try to construct an inductive argument for 'Everyone is unable to understand that r'. Either you can yourself understand that r or you can't. If you can, then the conclusion of the argument is false. But one might well think that if you can't yourself understand that r then it is not true that if someone can, you can tell that he can. The reason is as follows. You can't tell that S can understand that r unless you can understand 'S can understand that r'. And you can't understand 'S can understand that r' unless you have understood that r yourself. (Compare 'You can't understand "S can imagine red" unless you have had experiences of or have imagined red yourself'.)

The sceptic claims then that there is no inductive evidence and hence no evidence for the truth of any psychologically interpreted proposition of the form 'p is necessarily true'. Since it is also true that we do not know any such proposition by acquaintance, we don't know the truth of any such proposition. But if we don't know the truth of any particular proposition of this kind then we don't know that there are any universal conditionals, of the kind referred to in the definition of a dynamically objective value judgement, which are necessarily true. And so we don't know the truth of the doctrine of dynamic unity.

That is one possible argument. There is another. Consider again the universal conditionals which are referred to in the definition of a dynamically objective value judgement. If there are dynamically objective value judgements then some such universal conditionals are true. Either all these universal conditionals are contingent, or some of them are non-contingent. But (c) if they are all contingent then there are no dynamically objective value judgements, and (d) if some are non-contingent there are no dynamically objective value judgements. To show that (c) we appeal as before to facts of human diversity. In this part of the argument philosophy does no work. But philosophy is needed to show that (d), so the argument as a whole qualifies as properly sceptical. How does the sceptic show that (d)? N evaluates A as good, thereby expressing his desire for A to exist. F is the causative property which underlies his evaluation, and his value judgement is dynamically objective only if (i) he knows that A is F, and (ii) it is necessarily true that for any person S and any object x, if S knew that x was F then S would desire x to exist. Suppose that the universal conditional is non-contingent. The sceptic argues that in this case there will be no F which allows for the truth of both (i) and (ii). If the universal conditional is non-contingent F will have to stand for the kind of strange, anomalous and 'non-natural' property which Mackie thought we ordinarily attribute to things when we judge that they are good, a property which ensures that whatever has it 'would be sought by anyone who was acquainted with it, not because of any contingent fact that this person, or every person, is so constituted that he desires this end, but just because the end has to-be-pursuedness somehow built into it'.¹² Isn't such a property too strange to be instantiated? Even if it is instantiated how could we ever know that it is unless we possess some equally anomalous intuitive faculty? And then again, won't this non-natural property F have to supervene on natural properties, and won't that make for a further mystery if, as it is plausible to assume, there is no natural property G such that 'A is G' entails that A is F?¹³

We have then two possible arguments for ignorance in which P is the doctrine of dynamic unity. One invokes the epistemological implications of a general de-mystificatory doctrine of modality, the other relies on familiar considerations about the supervenience of non-natural on natural properties or the non-existence of what is strange or anomalous. When fully spelled out,

both may reveal themselves as evidence-or-acquaintance arguments. Both are at any rate properly sceptical. It may even be that both are sound.

The difficulty is, however, that neither seems to pass the negation test. Reflection would destroy N's desire to know the doctrine of dynamic unity. This can be shown by pressing the question, Why should one want there to be true universal conditionals about dynamic attitudes, as distinct from true conditionals about almost everyone, or about most people, or about the members of one's own community? But I would rather delay my treatment of this question. I would rather delay it until chapter VI, in which I discuss the non-negation-proof nature of sceptical arguments against justified belief. Is there then some other way to fail our arguments for ignorance in the negation test? I think there is.

The doctrine of dynamic unity entails that some value judgements are dynamically objective, and a dynamically objective value judgement is any value judgement about a particular object such that (i) the evaluator knows that the object has the causative property F, and (ii) for any person S and any object x, if S knew that x was F he would have an attitude towards x which is like the evaluator's desire, aversion or intention towards the particular object that he evaluates. Now it is necessarily true that if there are dynamically objective value judgements then either (a) all the relevant universal conditionals are contingently true or (b) at least some of the relevant conditionals are necessarily true. And it also seems to be true that

DISJUNCTION: For any person S, and any propositions p and q, if S wants that either p or q then he has some desire that p and some desire that q.

Suppose that reflection will destroy any desire that N may have for any of the relevant conditionals to be necessarily true. Then by DISJUNCTION N will not want that ((a) and (b)). Further suppose that there is a relation R between the doctrine of dynamic unity and the disjunction ((a) and (b)) such that if N believes that the two are linked by R, and if he does not want the disjunction to be true then he does not want the doctrine of dynamic unity to be true. Then we can conclude that on reflection N will not want the doctrine of dynamic unity to be true. And from there we can move to the conclusion that on reflection he will not want to know that it is true.

It does indeed seem that reflection will destroy any desire that N may have for any of the relevant conditionals to be necessarily true. He will not on reflection be able to find a convincing answer to the question of why he should want any of the conditionals to be necessarily rather than contingently true. The conditionals correlate knowledge that an entity is F with desires, aversions or intentions of the kind that you express when you evaluate some particular object which you know to be F. N might I suppose say that were certain such correlations somehow fixed, this would protect his own preferences, desires or intentions from the destructive effects of discoveries about their ultimate origins. He evaluates A as good, thereby expressing a

desire for it to exist. F is the causative property which underlies his evaluation. He knows that A is F. He might say that if he were to discover that his own desire for the existence of F things was originally caused by something to whose existence he is now hostile or indifferent then this discovery would undermine his desire, if it weren't somehow independently fixed by his knowledge. If his enthusiasm for certain colour combinations or for certain kinds of outspoken behaviour were not thus independently fixed then it would evaporate on his discovery that it had evolutionary causes or was the product of some traumatic episode of his infancy. Wanting to have stable desires and fearing the effects of discoveries about their origins, he might want the stability to be somehow guaranteed.¹⁴ It would indeed be guaranteed if it were necessarily true that whenever he knew that an object was F he would want it to exist. A fortiori, it would be guaranteed if there were a necessarily true universal conditional correlating knowledge that an object is F and desire for it to exist.

But this isn't really a very convincing answer. Why should the stability of your attitudes be threatened by discoveries about their origins? Surely you may still be glad that F things exist even if you think that you wouldn't have had this attitude if something had not happened towards which you have no independently positive feelings. You can regard the evolutionary or psychological cause of your desire for the existence of F things as something which has luckily helped you to be fully sensitive to Fness. There is a parallel with the causation of belief. Suppose you believe that p, and think that there is no evidence either for or against p's truth. It isn't clear that your belief that p will evaporate just because you come to realize that you wouldn't have had this belief if it hadn't been drummed into you by your Uncle. If that was indeed the origin of your belief then you are more likely now to think that since p is in fact true, it was lucky that your Uncle was there to drum it into you.¹⁵

What must R stand for if N's belief that the doctrine of dynamic unity is R-related to the disjunction ((a) or (b)) will ensure that his desire for the truth of that doctrine would be undermined by his lack of desire for the truth of the disjunction? Not the entailment relation. In the last chapter I discussed

ENTAILMENT: For any person S and any propositions p and q, if S believes that p entails q, and if S does not want that q, then he does not want that p.

This, as I said, must be qualified. It does not hold when S's reason for not wanting that q is that there is a logically sufficient but not necessary condition for q's truth which he does not want to be fulfilled. Suppose we replace p by 'N knows the doctrine of dynamic unity' and q by the disjunction ((a) and (b)). And suppose that N believes that the former entails the latter. Can we then use the qualified version of entailment to move from 'N does not want that ((a) and (b))' to 'N does not want to know the doctrine of dynamic unity'? Not unless we can find some new way in which reflection will destroy

N's desire for the truth of the disjunction. Our original thought was that his desire for the truth of the disjunction would be undermined by his lack of desire that (b). And (b) is of course a sufficient but not necessary condition for the truth of the disjunction. Could R stand for the relation in which one proposition stands to a second when the second is part of the analysis or definition of the first? I think it is true that if S believes that p stands to q in that relation, and does not want that q, then he does not want that p. But unfortunately the disjunction ((a) or (b)) is not likely to be believed by N to be part of the actual definition of the doctrine of dynamic unity. R is elusive. And yet it does seem that the disjunction stands to the doctrine of dynamic unity in *some* relation which is intimate enough for N's belief that it stands in that relation to ensure that his lack of desire for the truth of the disjunction has the right destructive effect.

If there is indeed a sufficiently intimate relation here we are left with the move from N's not wanting the doctrine of dynamic unity to be true to his not wanting to know that it is true. Here we can use the qualified version of ENTAILMENT. N will believe that his knowing the doctrine entails that the doctrine is true. And since (b) is not even a sufficient condition for the truth of the doctrine of dynamic unity, N won't not want the doctrine of dynamic unity to be true because he doesn't want a sufficient but not necessary condition for the truth of that doctrine to be fulfilled.

I return finally to the question left dangling at the end of the first part of this section: Can the sceptic construct a sound negation-proof evidence-or-acquaintance argument in which P is the doctrine of evaluative truth? It seemed that perhaps he could, provided that he employed a purely assertive or doubly attitudinal sense of 'evaluation', and took the evaluation to be in part at least the assertion of a universal conditional about desires, aversions or intentions. But in the light of what I have said about arguments in which P is the doctrine of dynamic unity, the prospects turn out to be poor. Sound though it may be, the sceptic's argument for ignorance of the doctrine of evaluative truth will fail the negation test. Either reflection will reconcile N to a kind of evaluation which requires the assertion of less than universal conditionals. Or, by the method just sketched, we can show that N's desire to know the doctrine of evaluative truth will be no more stable than his desire for some of these conditionals to be necessarily true.

(2) Theism

Evidence-or-acquaintance arguments are a kind of sceptical arguments for ignorance. Some evidence-or-acquaintance arguments rely on just this non-optional condition for knowledge:

- (A) S knows a contingent proposition p only if either (i) he believes a proposition which is evidence for p, or (ii) p is entailed by a true proposition of the form 'S is acquainted with x'.

We can call them (A) arguments. Others rely on the rather stronger

- (B) S knows a contingent proposition p only if either (i) he believes a proposition which is evidence for p, and no proposition is stronger evidence for not-p, or (ii) p is entailed by a true proposition of the form 'S is acquainted with x'.

We can call them (B) arguments. Like other arguments for ignorance, (A) and (B) arguments have conclusions of this form: 'N does not know that P'. If there is a sound and negation-proof (A) argument whose conclusion is that N does not know that P then there is a sound and negation-proof (B) argument with the same conclusion. If N can't meet condition (A), then he can't meet condition (B). But even if N does believe a proposition which is evidence for P, there may still be another proposition which is stronger evidence for not-P. So even if there is no sound and negation-proof (A) argument whose conclusion is that N does not know that P, there may still be a sound and negation-proof (B) argument with that conclusion. Can the sceptic construct a sound and negation-proof (B) argument, in which P stands for the doctrine that God exists?

(2.1) *Evil*

Some philosophers have argued to God's non-existence from what they have taken to be the incompatibility of his essential attributes. They have claimed for example that omniscience is incompatible with immutability,¹⁶ or omnipotence with omniscience.¹⁷ The sceptic might try to construct a sound (B) argument on the basis of such a claim. But if 'God exists' is indeed necessarily false, then neither this nor any other argument for ignorance of God's existence will pass the negation test. If P is necessarily false, so is 'N knows that P'. And if 'N knows that P' is necessarily false and N is informed by the sceptic's argument that this is so, then by INCOHERENCE reflection will destroy his desire for the knowledge which the sceptic says he lacks.

Suppose that 'God exists' is not necessarily false. Could the sceptic then construct a sound and negation-proof (B) argument, in which P is the doctrine that God exists? He must somehow exclude the possibility of knowledge by acquaintance. Suppose he can. He must also show that contrary to (B) there is stronger evidence for God's non-existence than any evidence for his existence. That evil exists may seem to be such evidence. There are in fact two different thoughts about evil on which the sceptic might try to rely. The first is deontological. The world contains evils which an omnipotent being would have been able to prevent. But no wholly good and omnipotent being would allow, even for the sake of a greater good, an evil that it was able to prevent. So there is no being which is both omnipotent and wholly good. There is also this non-deontological thought. A wholly good and omnipotent being would allow an evil that it was able to

prevent if and only if this was necessary for the existence of a greater good. But the world contains more evil which an omnipotent being could have prevented than is necessary for the existence of any greater good. So there is no being which is both omnipotent and wholly good.

That the sceptic encounters difficulties when he tries to use these thoughts tells one little or nothing about anything but evidence-or-acquaintance arguments. But the difficulties are real enough. As I am assuming, evidential relations hold only between propositions. The deontological and the non-deontological thoughts both include value judgements. And there are some senses of 'evaluation' on which to evaluate is not even in part to assert a proposition, but just to express a desire, aversion or intention. In the jargon of the last section, some senses of 'evaluation' are purely dynamic. If the sceptic is to cast either thought about evil in the form of a claim about evidence for God's non-existence, he must not employ a purely dynamic sense of 'evaluation', but rather one on which to evaluate is at least partly to assert some proposition. He must employ a purely assertive or doubly attitudinal sense. And now the question arises of what kind of proposition the sceptic will take the evaluator to assert. Will the sceptic say that to evaluate a particular object A is to assert something of the form 'A is F', together with a universal conditional correlating knowledge that something is F with desire, aversion or intention? Given what I said in the last section about the implications of human diversity, he would then have to concede that it is always a false proposition that the evaluator asserts. And false propositions can't be evidence. Suppose then that the sceptic takes it that to evaluate a particular object A is to assert something of the form 'A is F', together with a conditional which correlates the attitudes not of everyone but just of most people. If for example you judge that A is evil you assert something of the form 'A is F', together with the proposition that for most S, and any x, if S knew that x was F then he would feel fear, disgust or hatred towards x.

On this supposition, the deontological thought loses its cogency, and will not help the sceptic to construct a sound (B) argument in which P stands for the doctrine that God exists. The deontological thought will help the sceptic only if a true proposition is expressed by 'No wholly good being would allow, even for the sake of a greater good, an evil that it was able to prevent'. If however the evaluator uses this sentence to assert a restricted conditional of the kind just specified then what he asserts will be something like this: 'If an object A has a property F such that for any x most people who knew that x was F would feel fear, disgust or hatred for x, then most people would feel less than undiluted love or admiration for a being who was able to prevent A from existing but allowed A to exist for the sake of something whose existence most people would desire more than they desired A's non-existence'. And it is fairly unlikely that anything like that is really true, because it is probable that only a minority of people accept

deontological principles, and that most people have purely consequentialist inclinations.

Can the sceptic make more of the non-deontological thought? It was this: (i) A wholly good and omnipotent being would allow an evil that it was able to prevent if and only if this was necessary for the existence of a greater good. But (ii) the world contains more evil which an omnipotent being could have prevented than is necessary for the existence of any greater good. So (iii) there is no being which is both omnipotent and wholly good. Suppose, as before, that the sceptic takes the evaluation of a particular object A to involve the assertion of something of the form 'A is F', together with a less than universal conditional correlating knowledge that something is F with desire, aversion or intention. On this assumption, (i) might well come out as true. And the sceptic might claim that the conjunction of (i) and (ii) is stronger evidence for (iii) than any evidence for the existence of an omnipotent and wholly good being. It is doubtful, however, that the conjunction of (i) and (ii) is *any* kind of evidence for (iii). For the argument '(i) and (ii), so (iii)' seems question-begging. Any normally intelligent person who was doubtful about whether (iii) is true would thereby be made equally doubtful about whether (ii) is true. He would wonder whether he knew enough about which goods might exist or about how evils could be necessary for goods to be confident that (ii) is true. Doubts about (ii)'s truth are I grant most unlikely to visit anyone who is *not* doubtful about whether (iii) is true. A person who already believes that God does not exist may recognize that even if the world contains evils which would have been preventable by an omnipotent being and which seem utterly gratuitous relative to any greater good that he can imagine, still it is logically possible that they are necessary for some greater good that he can't imagine. His recognition of this logical possibility will hardly be enough to make him doubt that (ii). But a person who was doubtful about whether or not God exists would take what he cannot imagine more seriously. He would be moved by the thought that if an omniscient and omnipotent being *did* exist then much else might exist which was entirely beyond human powers of imagination and understanding, and that we would be in the position of people confronted with the activity or productions of a master in a field in which they had no expertise;¹⁸ 'from the theistic perspective there is little or no reason to think that God would have a reason for a particular evil state of affairs only if we had a pretty good idea of what that reason might be'.¹⁹ And to determine whether '(i) and (ii), so (iii)' is question-begging, and, hence, whether the conjunction of (i) and (ii) is *any* kind of evidence for (iii), we must indeed consider the attitude to this conjunction not of the atheist but of the person doubtful about whether or not (iii) is true.²⁰ If this is right, it isn't necessary to appeal to some theodicy, just in order to cast doubt on the prospects of basing a sound (B) argument on the non-deontological thought. We do not need anything as deep as that.

(2.2) Explanatory exclusion

Arguments for ignorance have conclusions of the form ‘N does not know that P’. (A) arguments are those sceptical arguments for ignorance which rely on no non-optional condition for knowledge other than

- (A) S knows a contingent proposition p only if either (i) he believes a proposition which is evidence for p, or (ii) p is entailed by a true proposition of the form ‘S is acquainted with x’.

Can the sceptic construct a sound and negation-proof (A) argument in which P stands for the proposition that God exists? Only if he can show that no proposition is evidence for God’s existence. His difficulty here is partly that there are certain not totally implausible cosmological arguments whose premises he cannot show not to be evidence for their imposing but subtheistic conclusions, and partly that were such conclusions granted it would be difficult for him to prevent their theistic supplementation.

Natural theologians advance cosmological arguments which rely on explanationist doctrines. All such doctrines may be false. But, I suggest, the sceptic cannot show that they are all false. And this prevents him from showing that in no such argument do the premises constitute evidence for the conclusion. One such explanationist doctrine is the Principle of Sufficient Reason: If it is consistent to suppose that a fact has some explanation or other then this fact has a true explanation. Grant the natural theologian this principle, together with some version of Ockham’s Razor, and he can deduce the existence of a being with several of the attributes traditionally supposed to be peculiar to God. One does not have to believe the Principle of Sufficient Reason in order to recognize that its falsity cannot be demonstrated.²¹ And that its falsity cannot be demonstrated is quite compatible with the view that it is not accepted by scientists and everyday reasoners. One could accept that its falsity cannot be demonstrated and yet agree with van Fraassen that scientists and everyday reasoners believe no more than that certain of their explanatory hypotheses are empirically adequate, and tell the truth about what is observable. It would also be quite consistent to say that though its falsity is indemonstrable the Principle of Sufficient Reason is a bad thing to believe, at least in so far as it applies to scientific theories.

If I believe the theory to be true and not just empirically adequate, my risk of being shown wrong is exactly the risk that the weaker entailed belief will conflict with actual experience. Meanwhile, by avowing the stronger belief, I place myself in the position of being able to answer more questions, of having a richer fuller, picture of the world, a wealth of opinion so to say, that I can dole out to those who wonder. But since this extra opinion is not additionally vulnerable, the risk is—in human terms—illusory, and *therefore so is the wealth*. It is but empty strutting and posturing, this display of courage not under fire and avowal of

additional resources that cannot feel the pinch of misfortune. What can I do but express my disdain for this appearance of greater courage in embracing additional beliefs which will *ex hypothesi* never brave a more severe test?²²

Someone might say that even if the natural theologian *is* allowed to assume the Principle of Sufficient Reason or some other general explanationist principle, he will be unable to establish more than some such sub-theistic conclusion as that there is an omnipotent and omniscient creator. If that were true, then the sceptic might still be able to construct a sound (A) argument in which P is the theistic hypothesis. The sceptic could at least still claim that there is no evidence for the existence of an omniscient and omnipotent creator who is also wholly good. Swinburne's treatment of this problem of supplementation illustrates its difficulty. Swinburne maintains that the creative hypothesis will be simpler if the explanatory being is omnipotent rather than just immensely powerful, and that omniscience is more 'consonant' with omnipotence than is partial ignorance. He then maintains that an omnipotent being would be perfectly free and that if 'moral judgements are... propositions which are true or false', then 'God's perfect goodness follows deductively from his omniscience and his perfect freedom'.²³ But to make the deduction Swinburne has to assume that it is incoherent to suppose that an agent 'might see refraining from A as over all better than doing A, be subject to no non-rational influences inclining him in the direction of doing A, and nevertheless do A'.²⁴ And, surely, there is no incoherence here.

I doubt however that the sceptic can afford to grant the natural theologian his explanationist premise. Instead of trying to deduce perfection from omniscience and freedom, the natural theologian could suggest that there are mystical experiences which, if veridical, would have objects which could not but be wholly good. The perfect being which would be the object of a veridical mystical experience could be identified with the all-powerful explainer. One does not need to believe that there are veridical mystical experiences in order to think that the sceptic is unable to demonstrate that all ostensible mystical experiences are delusive.

Is it so sure that even if he is allowed to assume an explanationist doctrine the natural theologian can establish his sub-theistic bridgehead? Consider Swinburne's Bayesian system, which relies at a crucial point on the doctrine that it is improbable that the existence and order of the universe has no explanation.²⁵ Even with the benefit of this assumption, Swinburne is obliged to rely on certain doubtful-looking non-explanationist principles of evidence in order to reach the conclusion that on the evidence he considers theism is more probable than not. Swinburne argues that on evidence other than that of religious experience the probability of theism is not very low. And he further argues that 'unless the probability of theism on other evidence is very low indeed, the testimony of many witness to experiences apparently of God

suffices to make many of these experiences probably veridical'.²⁶ But this further argument relies on the Principle of Credulity, and the Principle of Testimony. The Principle of Credulity is that 'if it seems (epistemically) to a subject that x is present, then probably x is present'.²⁷ The Principle of Testimony is that '(in the absence of special considerations) the experiences of others are (probably) as they report them'.²⁸ The sceptic could challenge these principles. There is, however, a fairly simple, non-Bayesian way to bring out the potential of explanationist principles. We can consider what can be called *explanatory exclusion arguments*.

Explanatory exclusion arguments begin by specifying a class whose non-emptiness is incontestable and which is such that if there is a non-member of this class which is logically capable of explaining why the class has any members at all then this non-member will have at least one of the attributes traditionally supposed to be peculiar to God. The actual existence of a non-member which explains why the class has members is then inferred with the help of the Exclusion Principle, together with some explanationist principle. The Exclusion Principle is that any explanation of why a class has any members at all must postulate a non-member of that class. Suppose for example that the initial class is the class of things which are caused to exist. Given the Exclusion Principle we can infer with the help of the Principle of Sufficient Reason that there is something which is not caused to exist and which explains why there is anything which is caused to exist.

If an explanatory exclusion argument is to serve the interests of a monotheistic natural theology then the initial class must be so chosen that we are not committed to an infinite explanatory regress. Suppose that, as in my example, the initial class is the class of things which are caused to exist, and that the explanationist principle is the Principle of Sufficient Reason: for any fact, if it is consistent to suppose that this fact has an explanation, then there is some true explanation of this fact. From the non-emptiness of the initial class, the Exclusion Principle and the Principle of Sufficient Reason we infer that there is something X which is not caused to exist and which explains why the initial class has any members at all. Is it consistent to suppose that there is an explanation for the fact that there is an X which is not caused to exist and which explains why there are things which are caused to exist? Not if all explanations are causal.

But some people believe that as well as causal explanations there are non-causal, personal explanations, in which the explanans is a volition or intentional action. On this doctrine it is consistent to suppose that there is a personal explanation for X 's existence. And if we assume that nothing can be the personal explanation of its own existence, then by the Principle of Sufficient Reason there is a further entity $X1$ whose intentional action gives a personal explanation for the existence of X . Is $X1$ caused to exist? No, because if it were then the fact that the class S of things caused to exist has any members at all would be explained by the existence of something $X1$ whose own existence is explained by something $X2$ which is itself a member

of S, which is absurd. Is there a personal explanation for the existence of X1? There has to be, given that this is a consistent supposition, and given the Principle of Sufficient Reason. So there is a further entity X2 whose intentional action gives a personal explanation for the existence of X1. And so on ad infinitum.

To avoid the regress we could take a different initial class. We could take as the initial class those things for whose existence there is an explanation. That this class has any members at all is explained by the existence of something X whose existence does not have an explanation. But it is not consistent to suppose that there is an explanation for the fact that something exists for whose existence there is no explanation, so the Principle of Sufficient Reason does not commit us to saying that there is some further entity X1 which explains why X exists. In your anxiety to show that there is just one entity which does not belong to the initial class you might fall into the trap of making that entity or its explanatory action totally unintelligible. Taking as the initial class those beings which are not self-explanatory, you might postulate a self-explanatory explainer. Taking as the initial class beings located in either space or time, you might postulate an explainer not located in either space or time. Leibniz seems to have proposed an explanatory exclusion argument with contingent things as the initial class and, as the explainer, 'a necessary being, bearing the reason of its existence within itself'.²⁹ The notion of an explainer whose existence has no explanation seems by contrast perfectly intelligible.

How does the natural theologian show that his unexplained explainer is omnipotent rather than just immensely powerful? And how does he show that his being is omniscient? By using further explanatory exclusion arguments. He takes as his initial class finitely powerful things, and argues that since the non-emptiness of this class has an explanation, there is an omnipotent being which explains the non-emptiness of the class. Similarly, he establishes the existence of an omniscient being by taking as his initial class beings which are wholly or partially ignorant. He then argues from considerations of simplicity that really it is one and the same being which is omnipotent, omniscient and an unexplained explainer. The appeal to simplicity may indeed rest on false assumptions. But I do not think that it is any easier for the sceptic to show that this is so than it is for him to demonstrate the falsity of the Principle of Sufficient Reason.

I suggest then that because the sceptic is unable either to discredit explanatory exclusion arguments or to demonstrate the delusive nature of ostensible mystical experiences he cannot construct a sound (A) argument for ignorance of God's existence. If this is right, and if sound (B) arguments with that conclusion are unobtainable, then the sceptic is unable to build any sound and negation-proof evidence-or-acquaintance argument with that conclusion.³⁰

V

ANTI-IDEALISM

It would be surprising, I thought, if philosophy could not from its own doctrines supply something for P to stand for in a sound and negation-proof evidence-or-acquaintance argument. Two obvious possibilities have been considered and rejected. Another will be considered now.

Let 'Idealism' stand for this doctrine: there are contingent entities which can exist when neither thought of nor experienced by any being of finite power, and while all these entities have intrinsic properties, none have any but mental intrinsic properties. It is fairly plain that Idealism is a doctrine to which most of us feel a certain repugnance. The doctrine repels us because we want there to be, beyond or outside us, rocks and rivers, eyes and arms, all with the shapes and colours that they seem to have. There are no such things, on the Idealist doctrine. It is all a false imaginary glare. If Idealism repels us, another doctrine is correspondingly attractive: there are contingent entities which can exist when neither thought of nor experienced by any being of finite power, and while all these entities have intrinsic properties, some at least have intrinsic properties of a non-mental kind. We can call this Anti-idealism. As well as being something that, initially at least, we may well want to be true, Anti-idealism is non-consensual. It isn't believed by everyone who entertains it. Idealists are not unheard of. There are also people who believe that some contingent entities have only dispositional properties. There is a chance then that if the sceptic can construct a sound evidence-or-acquaintance argument in which P stands for Anti-idealism, this argument will pass the negation test. Nor in his construction of this argument for ignorance will the sceptic be hampered by his inability to discredit explanationist principles, as he was when P was the theistic hypothesis. Unlike theism, Anti-idealism meets no explanatory demands. The intrinsic as opposed to the dispositional properties of what can exist when neither thought of nor experienced seem not to have any explanatory relevance. The sceptic may even insist that just because Anti-idealism lacks explanatory power, there is no evidence for its truth.

As it turns out, I have to conclude that there is no sound and negation-proof evidence-or-acquaintance argument in which P stands for Anti-idealism. And with that conclusion I bring to an end my search for a sound and sceptical

argument for ignorance which passes the negation test. In subsequent chapters I go on looking for sound and negation-proof sceptical arguments, but change the focus from arguments for ignorance to arguments against justified and against rational belief.

Section (1) contains more about Anti-idealism. Section (2) defends the soundness of an evidence-or-acquaintance argument in which P stands for Anti-idealism. Section (3) shows that this argument does not pass the negation test.

(1) Idealism and Anti-idealism

Both doctrines postulate a plurality of entities which can exist even if neither thought of nor experienced by any being of finite power. They postulate what we can call for short a plurality of alpha beings. Why the jargon? Why won't 'a plurality of mind-independent beings' do? Because the latter is ambiguous, and on no natural interpretation captures what I have in mind. Suppose we say that a being is mind-independent if and only if it could exist even if there were no minds. Then 'There is a plurality of mind-independent beings' will be incompatible with 'No beings can exist except minds'. But 'There is a plurality of alpha beings' is compatible with 'No beings can exist except minds', at least if a finite mind can exist without thinking of or experiencing itself. A mind-independent being might alternatively be taken as a finite being which can exist if neither thought of nor experienced. In that case the necessary existence of an essentially omniscient God would exclude the existence of mind-independent beings. But the existence of alpha beings is not excluded by the necessary existence of an essentially omniscient God: a being which can't exist unthought of by an omniscient God may still be capable of existing when not thought of or experienced by any being of finite power. The belief that there are alpha beings should also be distinguished from the so-called absolute conception of reality, according to which there are beings with properties whose intelligibility is not dependent on a particular 'perspective' or on the enjoyment of a particular kind of experience. Even a being not in this sense absolute can be an alpha being. Even if all its properties are only perspectively intelligible, it may still exist when neither thought of nor experienced by any being of finite power.

There is a plurality of alpha beings. On that, both doctrines agree. And both agree that some alpha beings are contingent, not necessarily existent. It is on the intrinsic properties of contingent alpha beings that the two doctrines differ. The Idealist believes that all contingent alpha beings have mental intrinsic properties, and that none has any intrinsic property of a non-mental kind. The Anti-Idealist believes that all contingent alpha beings have intrinsic properties, but that some at least have non-mental intrinsic properties.

By an intrinsic property, I mean one which is neither relational nor dispositional. 'Relational property' I use in a slightly more restrictive sense

than usual. It is often said that 'F' stands for a relational property if and only if 'x is F' entails the existence of something other than x or a part of x. I will however replace 'the existence of something other than x or a part of x', in this definition, by 'the existence of something other than x or a part of x which could exist if neither thought of nor experienced by x'. This revised definition makes the property of being taller than Mary relational. But unlike the more usual definition it makes non-relational such properties as that of having an experience as of a scarlet poppy. That S has a visual experience as of a scarlet poppy does not as I take it entail that there is a scarlet poppy which can exist if neither thought of nor experienced by S. For there are non-veridical as well as veridical experiences of this kind. What is entailed is that *either* there is such a poppy, *or* S's experience has an internal object which represents such a poppy, where an internal object of S's experience can represent something whether or not anything of the type it represents can exist if neither thought of nor experienced by S. So unless we say that such an internal object would be a part of S, the property of having a visual experience as of a scarlet poppy will on the usual definition have to be relational: that S has such an experience will entail that something exists which is neither S nor a part of S, the something being either an internal object or a scarlet poppy. But on the revised definition the property of having a visual experience as of a scarlet poppy will not have to be relational, because 'S has an experience as of a scarlet poppy' is consistent with there being nothing but S and an internal object which, even if it is a part of S, cannot exist if neither thought of nor experienced by S.

'Dispositional property' I use in a standard way, and I say that intrinsic properties are non-dispositional as well as non-relational in order to leave open the question of whether all dispositional properties are relational. There are, I assume, intrinsic properties of having experiences and thinking things, properties which are non-dispositional as well as non-relational. One can have an experience as of a scarlet poppy, or think of unicorns, without being disposed to any kind of bodily behaviour. I also assume that there are intrinsic properties of colour and shape. I grant that colour and shape words like 'scarlet' or 'spherical' are used to name dispositional properties, as well as to name intrinsic properties, 'x is dispositionally scarlet' means something like 'x has a disposition to produce experiences as of its being intrinsically scarlet'. But the dispositional sense depends on the intrinsic sense. If 'scarlet', as used in 'experience as of its being scarlet' itself had a dispositional sense, then an experience as of something's being scarlet would be an experience as of its being disposed to give us experience as of its being scarlet, which would in turn be an experience as of its being disposed to give us an experience as of..., and so on ad infinitum. But experiences as of scarlet things have no such complexity. 'Spherical', 'triangular', 'cubical' and other shape terms, have as I take it analogous pairs of senses, and so do all the other terms which we use in purely phenomenological descriptions of experience.

So for example there are both intrinsic and dispositional senses of ‘painful’, ‘sweet’, ‘cold’ and ‘shrill’.¹

I take it that an intrinsic (non-relational and non-dispositional) property is a mental property if and only if it consists in having an experience, or in thinking something, or it is a property of that part of an episode of thinking or experiencing which is not its content. The property which consists in having an experience as of a scarlet poppy is an intrinsic property of the mental kind. If such an experience is non-veridical, then its content is an internal object. It is a contested question whether its content is an internal object even if it is veridical: some people would say that S’s veridical experience as of a scarlet poppy has no content other than a scarlet poppy which can exist if neither thought of nor experienced by S. But however that may be, no intrinsic properties of shape and colour which are instantiated by the content of such an experience are properties of the mental kind. An intrinsic property is also mental if it is a property of that part of an episode of thinking or experiencing which is not its content. Suppose for example that there is something that it is like to will something, so that my willing that p has an intrinsic property which does not belong to its content, and is not shared by my believing that p, or by my wanting that p. Then this volitional property is a mental intrinsic property.

According to the Idealist, then, all contingent alpha beings have intrinsic mental properties, but none has any intrinsic non-mental property.² This is incompatible with the Cartesian doctrine that contingent alpha beings include some which neither think nor have experiences as well as some which think and experience. And it is incompatible with ‘eliminative materialism’, according to which there are no contingent alpha beings which think or have experiences. Since alpha beings can exist if neither thought of nor experienced by any being of finite power, Idealism, as I define it, could not be accepted by McTaggart. For according to McTaggart there are no non-conscious alpha beings, and all conscious alpha beings are both finitely powerful and forever thinking of themselves as well as of each other. Idealism is however general enough to be compatible both with the Berkeleian doctrine that all contingent alpha beings are spirits with human or superhuman powers, and with the Leibnizian doctrine that while all contingent alpha beings are conscious, some are but very dimly and confusedly so.

Can the Idealist say that some or all contingent alpha beings constitute a physical world? If a dispositional physical property consists merely in a disposition to cause experience, as Russell on occasion said, then the Idealist can believe in a physical world, in the sense of a world of contingent alpha beings with dispositional physical properties. As we will shortly see, he may even find a way to say that alpha beings are spatio-temporally located.³

The Anti-Idealist believes that at least some contingent alpha beings have non-mental intrinsic properties. His doctrine is therefore consistent with what we can call *chromatic qualativism*. This says that some contingent alpha beings have intrinsic colour properties. Chromatic qualativism is indeed a species of

Anti-idealism. The Anti-Idealist need not deny that some alpha beings have only mental intrinsic properties. So his doctrine is consistent with some types of mind-body dualism. But his doctrine is also consistent with the type of materialism which says that all contingent alpha beings have non-mental intrinsic properties and even with the type of materialism which says both this and that none has any mental intrinsic property.

I should perhaps stress that both Idealism and Anti-idealism are on the present definition doctrines about contingent existence as more than the internal object of a finitely powerful mind, rather than doctrines about contingent and irreducible existence. They are not, for example, doctrines about what John Foster calls 'ultimate contingent reality'.⁴

(2) Anti-idealism: an evidence-or-acquaintance argument

If Anti-idealism were demonstrably incoherent, we would know in advance that our general conjecture about sceptical arguments could not be falsified by an evidence-or-acquaintance argument in which P stood for that doctrine. No argument for ignorance will pass the negation test if P, in that argument, is demonstrably incoherent. By INCOHERENCE, N's desire for the actual truth of P would not survive his recognition of its incoherence, and neither would his desire to know its truth. Is it so certain, then, that Anti-idealism *is* coherent?

Take chromatic qualitivism, that species of Anti-idealism which says that some contingent alpha beings have intrinsic (non-dispositional, non-relational) colour qualities. I do not think it can be shown that even this doctrine is incoherent. According to McDowell, believing that it is coherent to suppose that there are colour properties 'which characterize things independently of their perceivers' and yet resemble colours as they figure in our experience, is no better than believing in the coherence of 'a conception of amusingness which was fully intelligible otherwise than in terms of the characteristic human responses to what is amusing, but which nevertheless contrived somehow to retain the "phenomenal" aspect of amusingness as we experience it in those responses'. He would 'sympathise with anyone who found the idea incoherent'.⁵ But what actual argument could back this up? Berkeley's argument about the unimagability of extramentally existing sensible objects would tell against the coherence of chromatic qualitivism, but it is hardly necessary to criticize it yet again. How could it be intelligible to say that a contingent alpha being has an intrinsic colour property when it is unintelligible to say that it is intrinsically sweet or bitter or has the intrinsic properties which we use to describe the content of experiences of pain? A familiar challenge. But the more one thinks about it, the less obvious it seems that there *is* anything unintelligible about these latter statements. Perhaps we cannot *imagine* a temporally located and non-conscious alpha being with any of the intrinsic properties with which we are acquainted by our taste or pain experiences. But is it clear that we are unable in thought to detach such qualities from contents of

experience and conceive of their possession by alpha beings? What else can be said for the incoherence of chromatic qualativism? An argument seems to arise from the ashes of McTaggart's proof of immaterialism. But it is probably chimerical.⁶

Even if Anti-idealism is coherent, the sceptic may still be able to build a sound evidence-or-acquaintance argument in which P stands for that doctrine. He may try to construct an argument in which no non-optional condition for knowledge is relied on other than

- (B) S knows a contingent proposition p only if either (ia) he believes a proposition which is evidence for p, and no proposition is stronger evidence for not-p, or (ii) p is entailed by a true proposition of the form 'S is acquainted with x'.

Alternatively, he may try to build an argument which relies only on the weaker

- (A) S knows a contingent proposition p only if either (i) he believes a proposition which is evidence for p, or (ii) p is entailed by a true proposition of the form 'S is acquainted with x'.

Suppose that the sceptic can't construct a sound (A) argument, because there is some evidence for Anti-idealism. Can he then construct a sound (B) argument? I doubt it. Even if he could show that Anti-idealism can't be known by acquaintance, he would not be able to show that the evidence for the falsity of the doctrine is stronger than the evidence for its truth. And he wouldn't be able to show this because there isn't any evidence for the falsity of the doctrine. There isn't even any evidence for the falsity of chromatic qualativism.⁷

To see that this is so, we must keep in mind the distinction between evidence for p's falsity, and lack of evidence for p's truth. Suppose that

- (P1) If q is a set of experiential data then q is good evidence for p if and only if p is entailed by the best explanation of q.

And suppose that if (P1) is true then there is no good evidence for the contingent truth of chromatic qualativism. It does not in the least follow that if (P1) is true then there is evidence that chromatic qualativism is false. There might of course be such evidence if it were true that

- (P2) The explanatory redundancy of a proposition not entailed by experiential data is good evidence for the falsity of that proposition.

But why should we accept (P2)? There are philosophers who accept some such principle on grounds of parsimony. Jackson maintains in Chapter 8 of his *Perception* that 'we have no reason to believe that material things are coloured', and from here he moves to the conclusion that it is 'reasonable to assert that...colour...is not a property of material things'. He says that 'although the precise status of Ockham's razor is a matter of dispute, it seems clear that properties we have no reason to believe are possessed by material things are properties we ought not to ascribe to them'.⁸ Mackie makes a similar move in his *Problems from Locke*:

the literal ascription of colours, as we see colours, and the like, to material things, to light, and so on, forms no part of the explanation of what goes on in the physical world in the processes which lead on to our having the sensations and perceptions that we have... admittedly physics does not itself tell us that no such properties are there. This denial is a further, philosophical step; but it is one which is at least *prima facie* reasonable in the light of the successes of physical theory...The philosophical principle of economy of postulation...supplies a reason for not introducing supposedly objective qualities of kinds for which physics has no need.⁹

I should have thought, however, that it is in a way less parsimonious to believe a proposition false than merely to believe that there is no evidence for its truth.

What plausible arguments are there for the falsity of chromatic qualitivism which do not depend on some such principle as (P2)? To my knowledge, there are just two.

The first can be extracted from this famous passage of Russell's:

There is no colour which pre-eminently appears to be *the* colour of the table, or even of any one particular part of the table—it appears to be of different colours from different points of view...And... even from a given point of view the colour will seem different by artificial light, or to a colour-blind man, or to a man wearing blue spectacles, while in the dark there will be no colour at all, though to touch and hearing the table will be unchanged. Thus colour is not something inherent in the table, but something depending upon the table and the spectator and the way the light falls on the table. When, in ordinary life, we speak of the colour of the table, we only mean the sort of colour which it will seem to have to a normal spectator from an ordinary point of view under usual conditions of light. But the other colours which appear under other conditions have just as good a right to be considered real; and therefore, to avoid favouritism, we are compelled to deny that, in itself, the table has any one particular colour.¹⁰

Russell thinks that if, for each particular colour quality which the table appears to have, we lack good evidence that it has that quality, it would be favouritism to believe that it has any colour quality at all. Suppose that it is indeed favouritism to believe that *p* when there is no better evidence for *p* than there is for not-*p*. Then for all the argument shows, it is favouritism to believe that the table has no colour quality. For all the argument shows, there is no better evidence for that hypothesis than there is for the hypothesis that it has some colour quality or other, we know not which.

The other argument depends on a realist interpretation of physics and on the assumption that chromatic qualitivism is committed to the proposition that there are medium-sized alpha beings, such as tables or apples and pears, which have continuously coloured surfaces or are even continuously coloured through and through. The thought is that on a realistically interpreted physics such medium-sized alpha beings are composed of spatially discrete particles, and thus can't be continuously and intrinsically coloured through and through or have continuously and intrinsically coloured surfaces. This argument can be resisted without challenging a realist interpretation of physics, and without exploring such possibilities as that 'all the spatial locations within [a material object] that are not occupied by discrete particles are occupied by "fields" set up by the relationship among particles'.¹¹ We need only point out that chromatic qualitivism entails nothing about the size or observability of its intrinsically coloured alpha beings. It does not entail the existence of Eddington's common-sense table,¹² or the continuously pink ice-cube of Sellars's manifest image of man-in-the-world.¹³ Its intrinsically coloured alpha beings may be the fundamental entities of a realistically interpreted physics.

I turn then to (A) arguments, to those evidence-or-acquaintance arguments which invoke just

- (A) S knows a contingent proposition p only if either (i) he believes a proposition which is evidence for p, or (ii) p is entailed by a true proposition of the form 'S is acquainted with x'.

There is I think a sound (A) argument in which P stands for Anti-idealism. With 'contingent alpha beings with non-mental intrinsic properties' abbreviated to 'contingent and non-mental alpha beings', it goes like this:

- (1) Nobody knows that there is a contingent and non-mental alpha being unless either (i) there is experience whose best explanation would postulate the existence such a being, or (ii) there is experience whose actual content is such a being;
- (2) There is no experience whose best explanation would postulate the existence of a contingent and non-mental alpha being;
- (3) There is no experience whose content is a contingent and non-mental alpha being;

So: (4) N does not know that there is a contingent and non-mental alpha being. 'S has an experience whose content is an alpha being' means not just that S has an experience as of an alpha-being, but that N is acquainted with an alpha being, so that there actually is an alpha being which is the content of N's experience, and the content of his experience is not an internal object which represents an alpha being. Given that much, given (A), and given that 'evidence' is governed by a no-question-begging condition, (1) seems true. Given (A) you know that there is a contingent non-mental alpha being only if either there is evidence for the existence of such a being, or someone has an experience whose actual content is such a being. Suppose that, as in our Chapter I arguments for inductive ignorance, q is evidence for p only if there

is a non-question-begging argument of the form, 'q, so p', where the argument 'q, so p' is question-begging if any normally intelligent person who was doubtful about the truth of p would thereby be made equally doubtful either about the truth of q, or about the truth of the conditional 'if q then p'. Suppose that p, in 'q, so p', stands for the proposition that there is a contingent and non-mental alpha being. What can q then stand for, if 'q, so p' is to be non-question-begging? There does not seem to be anything that it can stand for, other than that there is experience whose best explanation would postulate the existence of such a being.

Let q stand for a proposition of the form 'N has an experience as of—', where the experience does not have a contingent and non-mental alpha being as its content, but, whether or not it is veridical, has an internal object which represents such a being. An example would be 'N has an experience as of a scarlet poppy', where the experience does not have an actual intrinsically scarlet poppy as its content, but has as its content an internal object which represents such a poppy. It seems to me that if q is interpreted in this way, and p stands for the proposition that there is a contingent and non-mental alpha being, then 'q, so p' will be question-begging. Any normally intelligent person who was doubtful about the truth of p would thereby be made equally doubtful about the truth of the conditional 'if q then p'.

Now suppose by contrast that while p stands as before for the proposition that there is a contingent and non-mental alpha being, q stands for a proposition of the form 'There is experience whose best explanation would postulate the existence of a contingent and non-mental alpha being'. Then I think the argument 'q, so p' will not be question-begging. A normally intelligent person doubtful about the truth of p would not thereby be made equally doubtful about the truth of q or the truth of the conditional 'if q then p'. A person doubtful about the truth of p might well have independent doubts about the truth of q, but he wouldn't be doubtful about q's truth just because he was doubtful about p's truth. Is there anything else that q can stand for, on which 'q, so p' will be non-question-begging? Not that I can see.

I should perhaps stress that since (1) gives only a necessary condition for N to know that there is a contingent and non-mental alpha being, it is perfectly consistent to accept (1) on the grounds I have given and yet still deny that there is any sense of 'knowledge' on which inference to the best explanation of experience is enough to yield knowledge of the unobservable. One could accept (1) on the grounds I have given even if one rejected all principles of inference to the best explanation, and thought, with van Fraassen, that we should accept scientific theories only in the sense of believing that they are empirically adequate, or tell the truth about observables. Nor, of course, does (1) commit one to the dubious principle, which figures in arguments for the contingent falsity of chromatic qualativism, that nothing contingent exists except experience and what explains it.

I turn now to the second premise of the sceptic's argument for ignorance. It is

- (2) There is no experience whose best explanation would postulate the existence of a contingent and non-mental alpha being.

In support of (2), the sceptic could say this. Although there may be experience whose best explanation would postulate the existence of contingent alpha beings other than the actual subject of that experience, and although any contingent alpha being must have *some* intrinsic (non-relational and non-dispositional) property, still the explanation of experience in terms of alpha beings need not say anything about just what intrinsic properties these alpha beings have, and certainly it need not say either that they are non-mental, or that they are mental. If physics tells us anything about contingent alpha beings which goes beyond the description of regularities in our experience, it tells us only about their structural or causal features, their non-intrinsic properties.

The sceptic must consider this argument against (2):

- (A) There is experience whose best explanation would postulate the existence of a contingent alpha being other than the subject of that experience;
 (B) Any such alpha being must have *some* intrinsic property;
 (C) Any such alpha being will be either a space or spatially located;
 (D) Neither a space nor a spatially located being can have intrinsic properties all of which are mental;

So: (E) There is experience whose best explanation would postulate the existence of a contingent and non-mental alpha being.

It seems to me that even if (C) is true, (D) is false. Can a contingent alpha being be spatially located even if all its intrinsic properties are mental? John Foster has suggested that physical space could be identified with an internal object of the experience of a single non-human mind, an internal object which consisted of a single sensory field. The occupants of this space might be mobile parts of this field.¹⁴ In the event, this first suggestion does not help us to answer our question. According to our initial definition, an alpha being may be incapable of existing if not thought of or experienced by a being of infinite power. So if the non-human owner of the sensory field were a being of infinite power, the mobile parts of the field could be alpha beings, and this even if they were incapable of existing if not experienced by the owner of that field. They would not, however, be alpha beings all of whose intrinsic properties were mental, in the sense of 'mental' that I am using. In the usage explained in the first section of this chapter, intrinsic properties like redness or painfulness, which characterize sensory fields, are not mental properties. If Foster calls such properties mental, it is because he thinks it incoherent to suppose that they are possessed by anything other than internal objects of experience. But, as my discussion of chromatic qualativism showed, it is doubtful that there is any such incoherence. And this makes it appropriate to

take it that an intrinsic property is mental only if either it consists in experiencing something, or thinking something, or is an intrinsic property of thinking or experiencing which is not a property of the content of an episode of experiencing or thinking.

But Foster has another suggestion which does I think enable us to reject premise (D) of the spatial argument against

- (2) There is no experience whose best explanation would postulate the existence of a contingent and non-mental alpha being.

His suggestion is that space could be constituted out of individual subjects of experience, each of which can be assigned a unique location in a three-dimensional coordinate system by virtue of the character of the content of its experience.¹⁵

If the Idealist accepts premises (A), (B) and (C) of the spatial argument against (2), and, on the basis of this second suggestion of Foster's, rejects premise (D), then he will indeed have to accept that no two spatially located subjects of experience have qualitatively indistinguishable contents of experience at the same time. He may, however, be willing to say that subjects of experience can be spatially located by virtue of intrinsic properties of their thoughts or experiences which are not properties of the contents of these thoughts or experiences.¹⁶

The last premise of the sceptic's argument for ignorance of Anti-idealism falsity is

- (3) There is no experience whose content is a contingent and non-mental alpha being.

First I consider an argument for (3) which is based on recent work by Howard Robinson and J.J.Valberg.¹⁷ I then propose what seems to me a better argument.

To start with, some thoughts of Valberg's. Suppose that five seconds ago I looked at a book, focused on whatever object it was that was then present in my experience, and continued to focus on that object for the next five seconds. If half way through this last five seconds God had eliminated the book 'but maintained the activity in my brain just as it was when the book was there',¹⁸ then the activity in my brain in the latter part of the five second period would have continued just as it was in the former part.¹⁹ And in this case, 'within my experience' things would have been just as they were in the former part of the period.²⁰ It follows that the object which has actually been present in my experience for the last few seconds, the object on which I have actually focused for this period of time 'is such that it might have survived the elimination of the book'.²¹ So this object cannot be the book. It cannot be an external object, whose existence is independent of its presence in experience. It must rather be an internal object, whose existence is not thus independent.

Generalizing, secularizing and substituting 'alpha being' for 'external object', we get this. Suppose you have an experience as of an alpha being. Then

this experience will have been caused by activity in your brain, and either (i) this activity will have caused you to have an experience whose content is an alpha being, or (ii) it will have caused you to have an experience whose content is an internal object. But (i) can't be true because (iii) given any brain event which causes an experience whose content is an alpha being, it is empirically possible, not contrary to the laws of nature, for a qualitatively identical brain event to occur even in the absence of any appropriate alpha being, and to precede or be simultaneous with an experience whose content is an internal object. So (ii) is true, and there is no such thing as an experience whose content is an alpha being. Why is (i) supposed to be incompatible with (iii)? Because of the nomological relations which are supposed to be involved in causal relations between particular events. That a particular brain event is causally sufficient for a particular experience is supposed to imply that it is a law of nature that whenever there is a relevantly similar brain event there is a relevantly similar experience. To minimize controversy about the relation between causality and laws of nature we could, however, bypass causal notions and express the whole argument in purely nomological terms. The two premises of the argument would then be these:

- (A) Whenever you have an experience as of an alpha being x there is a preceding or simultaneous event b in your brain such that either (i) it is a law of nature that whenever a b -like brain event occurs its owner has an experience whose content is an x -like alpha being, or (ii) it is a law of nature that whenever a b -like brain event occurs its owner has an experience whose content is not an x -like alpha being, but rather an internal object which represents an x -like alpha being;

and

- (B) For any event in your brain, which precedes or is simultaneous with your experience as of an alpha being x , it is empirically possible that another qualitatively identical event should occur in your brain and precede or be simultaneous with an experience whose content is not an actual x -like alpha being, but rather an internal object which represents an x -like alpha being.

The rest would then go as follows. Given (B), it is empirically possible for a b -like event in your brain to precede or be simultaneous with an experience whose content is an internal object which represents an x -like alpha being. But then contrary to (A) (i) it is not a law of nature that whenever a b -like brain event occurs its owner has a simultaneous or later experience whose content is an x -like alpha being. And so given (A) we must say that whenever you have an experience as of an alpha being x there is an earlier or simultaneous brain event b such that whenever a b -like brain event occurs its owner then or subsequently has an experience whose content is not an x -like alpha being, but rather an internal object which represents an x -like alpha being. You don't, then, ever have an experience whose content is an alpha being. *A fortiori*,

- (3) There is no experience whose content is a contingent and non-mental alpha being.

The same argument can be extracted from Robinson's defence of the sensedatum theory of perception.²² We can call it the Missing Object Argument.

A critic of this argument might concede (B) but refuse to accept (A). And to support his resistance to (A) he might suggest that this proposition seems true only because it hasn't been distinguished with sufficient clarity from the true but for our present purposes perfectly unhelpful

- (A1) Whenever you have an experience as of an alpha being *x* there is a preceding or simultaneous event *b* in your brain such that it is a law of nature that whenever a *b*-like brain event occurs its owner has a simultaneous or later experience whose content is either an *x*-like alpha being or an internal object which represents an *x*-like alpha being.

Suppose that (B) is true. And suppose you have an experience as of an alpha being *x*, and this experience is preceded by or simultaneous with a brain event *b*. Then it is empirically possible that a qualitatively identical event should occur in your brain, without your having any simultaneous or subsequent experience whose content is an *x*-like alpha being. Given (A) we could now infer that *b* is such that it is a law of nature that whenever a *b*-like brain event occurs its owner has an experience whose content is nothing more than an internal object which represents an *x*-like alpha being, and hence that your experience as of an *x* has no alpha being as its content. But if (A) is replaced by (A1), this inference is blocked. All that follows is that your experience as of an *x* has as its content *either* an internal object which represents an *x*-like alpha being *or* an actual *x*-like alpha being. A defender of the Missing Object Argument needs then to show why we should accept (A), rather than the true but innocuous (A1), with which, we may suspect, he has confused it.

He may reply that, so far from being true but innocuous, (A1) is in fact quite clearly false: the type of experience which (A1) correlates with *b*-like brain events is too unspecific for the correlation really to be a law of nature.²³ But what *is* the relevant condition on laws of nature? (A1) correlates with *b*-type brain events a disjunctive type of experience, the type made up of those experiences which either have *x*-like alpha beings as their contents, or have as their contents internal objects which represent *x*-like alpha beings. Does this fail as a nomological event-type merely because it is disjunctive? Presumably not, for in a probabilistic law to the effect that when an *A*-type event occurs there is a certain probability that an *X*-type event will occur but also a certain probability that a *Y*-type event will occur, it is a disjunctive event-type which is correlated with type *A*. But even if the relevant requirement on laws of nature can indeed be satisfactorily formulated, and even if it is violated by (A1), the critic of the Missing Object Argument can easily restate his point.

Instead of saying that (A) seems true only because it hasn't been properly distinguished from true, innocuous and equally nomological (A1) the critic can

say that (A) seems true only because it hasn't been properly distinguished from the true, innocuous and non-nomological

(A1)* Whenever you have an experience as of an alpha being x there is a preceding or simultaneous event b in your brain such that whenever a b -like brain event occurs its owner has a simultaneous or later experience whose content is either an x -like alpha being or an internal object which represents an x -like alpha being.

The conjunction of (A1)* and (B) is again quite consistent with the existence of experiences whose contents are alpha beings, and so even is the conjunction of (A1)* and the proposition that some b -like brain event actually occurs when neither then nor later is there an appropriate x -like alpha being. And unlike (A1), (A1)* has no vulnerable implications about the specificity of the event-types able to figure in laws of nature.

But defenders of the Missing Object Argument also have a less formal qualm about (A1), which is equally applicable to the non-nomological (A1)*. Suppose you have an experience as of an alpha being x , but there is no x -like alpha being to be the content of your experience. Then given either (A1) or (A1)* your experience will have an internal object as its content, and will follow or be simultaneous with an event b in your brain. But isn't it extremely odd that b should in these circumstances be accompanied or followed by *any* kind of experience as of an alpha being x ? Are we to suppose that the brain somehow knows that no x -like alpha being is present and maliciously deceives its owner by producing an experience as of an alpha being x ? Or that although the brain has a natural tendency to produce experiences whose contents are merely internal objects, an alpha being will when it is present inhibit by some kind of action at a distance the brain's tendency to produce an experience whose content is merely an internal object, and instead produce an experience of which it is itself the content? Surely these are absurd hypotheses. But how else are we to explain the case?²⁴

Faced with these questions, our critic may stop trying to find some true but innocuous proposition that his opponents have failed to distinguish from (A), and instead ask why we should accept either (A) or any of the theses with which we have seen that (A) is liable to be conflated. He may admit that for any human experience as of an alpha being x there is a preceding brain event b such that whenever someone has an experience as of an x -like alpha being this experience is preceded by a b -like brain event. But he may deny that b -like brain events are always followed by experiences as of x -like alpha beings. Some b -like brain events, he may say, are not followed either by experiences with x -like alpha beings as their contents or by experiences with internal objects which represent x -like alpha beings. He may thus refuse to accept even (A1)*. Does science exclude this lack of regularity? It may seem that we already know enough about the chemical or surgical production of hallucinatory experience to conclude that, for any alpha being which we perceive, there is a sufficient cerebral condition for the production of a

hallucination of that object. But even if this is so, it is another thing to know that when a hallucination of an x-like alpha being is generated by an artificially induced brainevent, that brain event is of precisely the same kind as the brain event involved in the veridical perception of an alpha being.

I think, however, that it is wiser to drop any challenge to (A) which also threatens (A1)*, and to go on insisting that if (A) seems plausible this is because it isn't being clearly enough distinguished from the true but unhelpful (A1)*. You may still find the latter less absurd than the rhetorical questions about action at a distance and maliciously deceiving brains are intended to make it seem.

Put yourself in the position of the believer in experiences with alpha beings as their contents. You are to suppose that it is not just empirically possible but actually the case that some b-like brain events occur in the absence of x-like alpha beings. Then, given (A1)*, some b-like brain events are followed by experiences whose contents are actual x-like alpha beings, while others are followed by non-veridical experiences as of x-like alpha beings. You are forbidden to accept this consequence unless you can provide a good explanation of why it is that in the first case actual x-like alpha beings are present and that in the second case they are absent. And the only explanations that you are supposed to be able to suggest are absurdities about action at a distance and maliciously deceiving brains. You can reply either by denying that you are obliged to provide any explanation of the difference, or by providing some non-absurd explanation of the difference.

Why must you provide any explanation? The answer will presumably be that if you don't then you violate some such general principle as

- (E) Whenever some X-type events are followed by Y-type events but others are not there is a good explanation for this difference between the consequences of the X-type events.

But can such a principle be reconciled with the view that there are irreducibly probabilistic fundamental laws of nature? Someone may say: if we arrange it so that no x-like alpha being is present, but that a b-like brain event occurs, it is no random matter that there is then a non-veridical experience as of an x-like alpha being, whereas there is a randomness in the behaviour of sub-atomic particles under any conditions that we are able to arrange. Perhaps. But this shows only that there are different ways in which (E) can be violated. The point remains that it is only on some such seemingly arbitrary principle as (E) that you are obliged to find an explanation for the difference between the consequences of b-like brain events.²⁵ This being so, we can perhaps continue to uphold (A1)*, and, while rejecting the essential premise (A), explain its plausibility in terms of its conflation with (A1) and (A1)*. A doubt remains about whether the Missing Object Argument really shows that

- (3) There is no experience whose content is a contingent and non-mental alpha being.

But perhaps this final premise of the sceptic's argument for ignorance of Anti-idealism can be supported in a different way. Suppose you have an experience as of a scarlet poppy, and, contrary to (3), the content of your experience is not an internal object which represents a scarlet poppy, but an actual qualitatively scarlet poppy which could exist unthought of or unexperienced by you or any other finite being. Then you are directly affected by this scarlet poppy. When you have an experience as of x then something is impressed upon you. You are directly affected by it. That is what it is like. To describe the something that is impressed upon you and thus directly affects you is to describe the content of your experience. This content may be an internal object, or, as in our present case, it may be an object which could still exist even if it were not the content of your experience. In our present case a brain event ultimately caused by the poppy will indeed somehow cause you to be directly affected by what is impressed upon you. But it is not a brain event which directly affects you by virtue of being impressed upon you. It can in fact only be the poppy which thus directly affects you, for in the present case the content of your experience is not an internal object which represents a scarlet poppy. What we must say, then, is that the subject of consciousness can be directly affected by a non-mental alpha being other than its own brain.

Consider now the proposition that an agent can have a direct effect on a non-mental alpha being other than its own brain. It seems utterly implausible. We assume without question that we can't through volition have any effect on any non-mental alpha being other than our own brains except through the mediacy of changes in our own brains. You don't when you will your arm to go up, or decide to raise your arm, will or decide that those specific brain events shall occur which are causally necessary for your arm to rise. But it is still these brain events which are the direct effect of your volition, and not the arm movement which they cause. Only through the mediacy of changes in your own brain can your decisions or volitions have any effect on other parts of your own body or on non-mental alpha beings other than your own body. This seems to have the status of what C.D. Broad called a *basic limiting principle*. It is one of those Very general principles, mostly of a negative or restrictive kind, ... which form the framework within which the practical life, the scientific theories, and even most of the fiction of contemporary industrial civilisations are confined'.²⁶ Another principle which has this status is that one cannot foresee (as distinct from infer on the basis of one's past experience) what will happen in the future. And a principle which seems not to have this status is (E). It seems to me that once we are forced by philosophers of perception to formulate the claim that a subject of consciousness can be directly affected by a non-mental alpha being other than its own brain, we can see that its negation is a basic limiting principle. If you have a veridical experience as of a scarlet poppy, you aren't aware of the brain events which are causally necessary for you to have that experience. But it is still only these brain events which directly affect you, and not the scarlet poppy which causes them. That you are not directly affected by

any non-mental alpha being other than your own brain is I suggest as much a basic limiting principle as that you do not by your volition directly affect any non-mental alpha being other than your own brain. And that, I suggest, is why it is true that

- (3) There is no experience whose content is a contingent and non-mental alpha being.

If this is right, or if the Missing Object Argument does after all work, then the sceptic can indeed construct a sound evidence-or-acquaintance argument in which P stands for Anti-idealism. I now turn to consider whether his argument will pass the negation test.

(3) The Negation Test

No one will deny that on first acquaintance Idealism is an unattractive doctrine. Knowledge of Anti-idealism is then likely to be at least an initial object of desire. Suppose the sceptic confronts you with the evidence-or-acquaintance argument which I defended in the last section. How might reflection destroy your initial desire for it to have a false conclusion?

The missing knowledge of Anti-idealism is of the evidence-or-acquaintance kind. Someone may suggest that a reliabilist kind of knowledge is an adequate and more easily available substitute. N has no evidence that there is a contingent and non-mental alpha being. Nor does he have experience whose actual content is a being of that type. Maybe this would not matter if by some generally reliable process he could acquire the true belief that there is such a being.

What, I wonder, would this process be? Suppose you have a visual experience as of a scarlet poppy before you, and this makes you believe that there is a scarlet poppy before you, and hence that you have before you a contingent alpha being which is intrinsically scarlet, and hence that there is at least one contingent and non-mental alpha being. As we saw, chromatic qualativism is extremely hard to disprove: is hard to prove that there are no contingent alpha beings with intrinsic colour properties. It is not out of the question that in the case considered there is indeed a contingent alpha being before you which has some intrinsic colour property or other. It is even conceivable that there is an intrinsically scarlet alpha being before you. But even if your visual experience does thus lead you to the true belief that you have before you an intrinsically scarlet alpha being, and hence to the true belief that there is at least one contingent and non-mental alpha being, still this whole process of belief-formation seems not to be generally reliable. Even the chromatic qualitivist, who believes that there are contingent alpha beings with intrinsic colour properties, will find it hard to maintain there is a regular and exact resemblance between the particular intrinsic colour properties which contingent alpha beings have, and the particular intrinsic colour properties which our visual experiences make them appear to have. In the last section I argued for

- (3) There is no experience whose content is a contingent and non-mental alpha being.

If the chromatic qualativist accepts (3) then he must say that the contents of experiences as of intrinsically coloured external things are internal objects of those experiences. He will not claim that the intrinsic colours of contingent alpha beings play any role in explaining the particular intrinsic colours which characterize the internal objects of visual experiences. So it will be an accident if such experiences are veridical in point of colour, and a highly unlikely accident at that. Why after all should the intrinsic colours of the alpha beings be just those which characterize the internal objects, rather than any of the numerous other colours which we are able to distinguish (200 or so, even if we only count hues)? Why indeed should the alpha beings not have intrinsic colours of a kind with which experience leaves us entirely unacquainted? There are moreover cases in which one observer has an experience as of say a wholly scarlet poppy, and another observer a simultaneous experience as of a wholly muddy brown poppy, and in which the chromatic qualativist will like everyone else want to say that they are looking at one and the same flower. Since the chromatic qualativist will deny that the poppy is both wholly intrinsically scarlet and wholly intrinsically muddy brown, he will have to agree that at least one of the experiences is not produced by a generally reliable process. And in these circumstances it is hard to see what grounds there could be for saying that either experience has been produced by such a process. And then again, if our experiences as of intrinsically coloured external objects were regularly veridical in point of colour, and were reliable guides to the intrinsic colours of the contingent alpha beings whose existence made the experiences veridical, then the intrinsic colours of these contingent alpha beings would be both highly changeable and incredibly variegated. The intrinsic green of the forest would really darken as dusk gathers. The parts of a contingent alpha being would be as different in their real intrinsic colours as the seeming intrinsic colours of their parts: 'If you look at port wine in a slim conical glass held against the light, its colours range from pale yellow at the bottom, through orange to ruby-red, and, as the reader can verify for himself, a bathful would look almost black.'²⁷ It is doubtful, then, that reliabilist knowledge of Anti-idealism is even available.

But, all the same, the sceptic's evidence-or-acquaintance argument fails the negation test. It fails because, repellent though Idealism initially is, reflection will destroy our desire for its falsity, and, hence, our desire for the truth of Anti-idealism, and hence our desire to know that Idealism is false. To show that the argument fails the negation test we can employ a method which I mentioned in chapter I when I was discussing arguments for inductive ignorance. Grant, to begin with,

ENTAILMENT: For any person S and any propositions p and q, if S believes that p entails q, and does not want that q, then he does not want that p.

For *p*, in ENTAILMENT, substitute 'N knows that Anti-idealism is true', and for *q* substitute 'Anti-idealism is true'. Not much reflection is needed for N to believe that the first of these two propositions entails the second. So if reflection does indeed destroy his desire for the truth of Anti-idealism then by ENTAILMENT it will also overcome his desire to know that Anti-idealism is true. The argument for ignorance will fail the negation test. It is true, as we saw, that ENTAILMENT needs to be qualified. Even if S does believe that *p* entails *q*, and does not want that *q*, he may still want that *p* if his reason for not wanting that *q* is that there is a logically sufficient but not necessary condition for *q*'s truth which he does not want to be fulfilled.²⁸ But as we will see, this proviso makes no difference. Reflection will destroy N's desire for the fulfilment of a condition which is necessary as well as sufficient for the truth of Anti-idealism, namely that some alpha beings have some non-mental intrinsic qualities or other.

How does Idealism lose its repellent aspect? Consider the source of our dislike. Is this just that you cannot help but disbelieve the doctrine once you entertain it, and that, wanting in general that what you believe corresponds to how things are, you want its falsity for the sake of avoiding false belief? I doubt it. The manifest image of man in the world, that partly depersonalized remnant of the original image in terms of which man first came to be aware of himself, still stubbornly contends with the scientific image of a world whose only ultimate constituents are imperceptible objects postulated for the explanation of correlations among observables.²⁹ And Idealism is doubtless incompatible with the manifest image. But the manifest image is not in its entirety an even intermittently unavoidable object of belief. Intermittently, the solipsist is compelled to believe that he is not alone. Likewise the man who disbelieves in his own freedom of choice: practical situations force him to believe for the moment that he is free, even if later he believes that this was an illusion. But I doubt if the Idealist is condemned by his relations to the manifest image even to such diachronic inconsistency.

Do we dislike Idealism because it would deprive us of our familiar stock of sub-human objects which we can break, bend, burn, melt, eat and generally turn to all our purposes, uninhibited by the fellow-feeling which sometimes restrains us in our treatment of the human world? That seems confused. The Idealist postulates a plurality of contingent alpha beings with no intrinsic properties but those of the mental kind. But he does not say that when I have an experience as of an oak tree before me, there is a single contingent alpha being which corresponds to the internal object of my experience, or that were I to have an experience as of felling the tree, and of lopping and burning its branches, real damage would be done to what would exist unthought of or unexperienced by any being of finite power.

Idealism repels us not, or not just, by what it seems to imply about the non-human world but rather or also by what it seems to imply about human beings.

Her eyes are blue and you admire them (they are a beautiful violet blue). You naively believe that they would still exist, with just that blueness that you so admire, if neither you nor anyone else could see them. This, you realize, is incompatible with Idealism. If you are reluctant at this point to embrace Idealism and abandon your belief about her eyes, it isn't, or isn't only, because you fear a diachronic inconsistency in your belief system or want to do anything *to* or *with* her eyes. Perhaps this is your thought. If Idealism were true then the blueness would be nothing but a feature of an internal object of your own experience, and hence a property not of *her*, or of *her* eyes, but of *you*. Your admiration reduces to admiration of yourself. It is true that even from an Idealist standpoint her eyes could still be alpha beings with a dispositional property of blueness, a disposition to produce experiences as of intrinsically blue eyes. But admirable though the mysterious mechanism might be which grounded that disposition it isn't, or isn't only, a dispositional blueness that you admire. What you admire is what you take to be the intrinsic blueness of her eyes, and if that becomes a feature merely of an internal object of your own experience it is yourself that you are admiring, and not, as of course you want it to be, something or someone distinct from yourself.

Here again reflection may lead you to change your mind. It may lead you to think that you have misconstrued the relation between yourself and the internal objects of your experiences. In your experience a content is presented. Are you, the owner, just a certain series of such episodes of presentation and of analogous episodes in which a content is thought? Or are you a substance, a body or spirit, which is not identical to any such series? Suppose the latter. Then clearly the internal objects of your experience are not after all a part of you. Suppose the former: you are identical to a series of experiencings and thinkings. Can't we even in this case draw a distinction between the admiration of (i) the internal objects of voluntary thoughts and experiences, contents contingent on a freedom not to think or experience, and (ii) the internal objects of thoughts and experiences not produced at will? Do we not feel that it is somehow less egocentric to admire internal objects of the second type than to admire those of the first? When you deliberately imagine her blue eyes, the internal object is of the first type. But when you have an actual visual experience as of her blue eyes, the internal object is of the second type. We could indeed say that there is no self-admiration in the admiration of an internal object of the second type if the self is identified with a series of voluntary thoughts and experiences, and certain involuntary thoughts and experiences are regarded as its property rather than a part of itself. I won't now try to develop such a concept of the self. Nor will I say any more about how the appearance of self-admiration could be dispelled by developing a concept of the self as spirit or body. The point is simply that the appearance can probably be dispelled by further metaphysical reflection on these lines.

Even if Idealism doesn't require you to admire *yourself* when you admire what you naively take to be the intrinsic blueness of her eyes, doesn't it still require you to be admiring something other than *her*? Someone will now assure you that even when you naively admired the blueness of her eyes, you never imagined that it was actually *her* or even a part of her that you were admiring. You never imagined that she was identical to her body. Or, if you did, reflection will cause a rapid change of mind. But this still leaves the fact that you did at least imagine that you were admiring for its intrinsic blueness something which is not only distinct from you, as even an internal object of your experience may be, but an *alpha being* distinct from you. If Idealism is true there are no alpha beings with non-mental intrinsic properties, and in particular no alpha beings with intrinsic colour properties. Don't you want there to be a world in which alpha beings have intrinsic colour properties and indeed just those particular intrinsic colour properties for which you naively admire them, and this regardless of whether the existence of such a world would allow you to admire intrinsic colours without admiring yourself, and regardless of whether the intrinsically coloured alpha beings are people or parts of people? How, if at all, can reflection destroy *this* desire?

What we must say is that even if Anti-idealism is true, still no alpha beings will have the particular non-mental intrinsic qualities for which you admire them, as distinct from just some non-mental intrinsic qualities or other. Even if Anti-idealism is true, you have to make do with your continuing ability to admire the intrinsic qualities of the internal objects of your experience.³⁰ To recognize this is to undermine your desire for the doctrine's truth, and hence your desire to know its truth. As already pointed out, bare chromatic qualitivism is extremely hard to refute. It may even be possible, by a complete rejection of the realist interpretation of physics, to maintain that the world includes macroscopic contingent alpha beings which have continuously and intrinsically coloured surfaces or which are continuously and intrinsically coloured through and through.³¹ But unless we suppose, what I have argued to be false, that we have experiences of which contingent and intrinsically coloured alpha beings are the actual contents, it seems highly probable that when we naively and admiringly attribute particular intrinsic colours to ostensible alpha beings like tables and eyes we are making a mistake. Even if chromatic qualitivism is true, and such objects have some intrinsic colours or other, it is highly improbable that they have just the particular colours that you admire in them. Why should her eyes be violet blue, when they could as well be orange or lime green? Even if Sellars's ice cube has *some* continuous and intrinsic colour through and through, it is highly unlikely that this colour is its seeming pink. Here, Anti-idealism is no help. What may perhaps help is to see consciousness as supplying 'a home in which objects can enter into actuality, so that we as consciousness are to be thought of as existing for the sake of objects which need us in order to exist rather than its being the objects which exist for our sake.'³²

THE PRICE OF DOUBT

To summarize. Though the sceptic can probably construct a sound evidence-or-acquaintance argument in which P stands for Anti-idealism, still this argument will fail the negation test. Reflection can undermine our initial hostility to Idealism by reminding us that even if Anti-idealism is true still alpha beings will not have the particular qualities that we want them to have. And if reflection can thus destroy our desire for the truth of Anti-idealism then it can also destroy our desire to know its truth.

VI

ARGUMENTS AGAINST JUSTIFIED BELIEF

As I conjectured, all sound sceptical arguments fail both the negation or the affirmation test. One kind of counterexample would be a sound, sceptical and negation-proof argument for ignorance. No such counterexample has emerged from the last five chapters. But not all sceptical arguments are arguments for ignorance. Some are arguments against rational or against justified belief.

Arguments against rational belief have conclusions of this form: 'N is not rational in believing that P'. Like arguments for ignorance, they are sceptical only if philosophical: they must not appeal to N's personal deficiencies or special historical circumstances; they must be arguments in which philosophy does some actual work.¹ Can we build sound, sceptical and negation-proof arguments against rational belief? If so, my conjecture falls. One way to define rational belief is in terms of rational action: S's belief that p is rational if and only if some action of S's is rational which is suitably related to his belief that p. What is it to call an action rational? Is it to make a value judgement? Not necessarily. It can be to say no more than that the action is, or is believed by the agent to be, at least as conducive to what are in fact the agent's ends as any action which he is able to perform instead. So one way to define 'rational belief' is in terms of the non-evaluative rationality of belief-related action. We can I think construct various sound and sceptical arguments in which 'rational belief' is taken in such a way, and there are some among them whose failure in the negation test is by no means obvious. This we shall see in chapters VII and VIII. But first, a look at sceptical arguments more naturally describable as arguments against justified belief. Even here, my general conjecture comes under a degree of pressure.

In contemporary epistemology 'justified belief' is a term of art with no generally agreed function or definition. The same goes even for 'epistemically justified belief'. As Alston remarks, with some restraint, current disputes about the nature of epistemic justification have features which 'are best explained by supposing that there is no unique item called "epistemic justification" concerning which the parties are disagreeing.'² To call an action justified is to evaluate it, and usually at least to make a moral evaluation. Belief is also open

to moral evaluation or to evaluation of some analogous kind. This helps to explain some current uses of 'S is justified in believing that p', and 'S is epistemically justified in believing that p'. 'Do what you can to believe what's true'; 'Do what you can not to believe what's false.' The latter, if not the former, is thought to express a moral duty, or an intellectual duty somehow analogous to a moral duty. An action is justified if it is permissible, not one that the agent had a duty not to do. So 'S is justified in believing that p', and 'S is epistemically justified in believing that p' are naturally enough used to express the thought that so far as believing that p is concerned, S has not failed to do his moral or intellectual duty. That however is not the only way in which such sentences are used. Epistemologists have been heavily engaged in 'the analysis of knowledge', in trying to determine what, in 'the' ordinary sense of 'knowledge', distinguishes knowledge from mere true belief. Often they have used 'justification', or 'epistemic justification', to stand for whatever they have thought that this distinguishing feature is, or would have been if there had been no Gettier counterexamples to various classical suggestions about what it is. And some writers who use the terms in this way, or who think that there are at least some senses of 'knowledge' on which knowledge requires justified belief, also maintain that there is nothing normative, or nothing interestingly normative, about epistemic justification. Thus Moser insists that one's belief that P is justified if there is something that makes P 'evidentially more probable for one, on one's total evidence, than not only \sim P but also P's probabilistic competitors'.³ The concept of justified belief is, he says, 'semantically independent of any normative concept'.⁴ And Fumerton maintains that the concept of epistemic justification is normative only in the sense that epistemic judgements are sometimes expressed in the kind of language also used to express moral or prudential judgements. 'To describe someone as being epistemically justified...in believing P is not in and of itself to make *any* moral or prudential claim about what he ought to believe. It is not to praise or blame the person for having the belief. It is not...even to praise or criticize the belief.'⁵ There are on the other hand some writers who take the concept of knowledge to be itself essentially normative or evaluative.⁶

Arguments against justified belief have conclusions of this form: 'N is not justified in believing that P'. Let us take it that in *some* such arguments 'justified belief' is so used that (i) the conclusion of the argument does not negate an evaluation, (ii) for some sense or other of 'knowledge', 'S is justified in believing that p' is entailed by 'S knows that p', and (iii) 'S is justified in believing that p' does not entail that p. Are there sound sceptical and negation-proof arguments of this kind? Not, it seems, unless there are also sound, sceptical and negation-proof arguments for ignorance. And these we have so far not discovered. Suppose for example that 'justified belief' is so taken that S is justified in believing that p if and only if either (i) he believes a proposition which is evidence for p, and no proposition is stronger evidence for not-p, or (ii) he has an experience as of an x, where 'there is an

x' either entails that p or is evidence for p, and no proposition is stronger evidence for not-p. (If p stands for 'There is a scarlet poppy', (ii) would be satisfied if he has an experience as of a scarlet poppy.) There is an evidence-or-acquaintance sense of 'knowledge' on which 'S knows that p' entails that S is in this evidence-or-experience sense justified in believing that p. If there is a sound and sceptical evidence-or-experience argument against justified belief, there will it seems be a sound and sceptical evidence-or-acquaintance argument for ignorance. And if this argument for ignorance fails the negation test, so also will the argument against justified belief. It is true that 'S knows that p' entails that p, whereas 'S is justified in believing that p' has no such implication. This means that the argument for ignorance, unlike the argument against justified belief, may fail the negation test simply because N has no reflectively indestructible desire for P's truth, and hence no reflectively indestructible desire to know that P. It does, however, still seem that if N has no reflectively indestructible desire for P's truth, then he will not have a reflectively indestructible desire even to have the evidence-or-experience kind of justification for believing it.

But in this chapter I want to focus on a different kind of argument against justified belief. I want to focus on arguments in which 'N is not justified in believing that P' *does* negate an evaluation, and in which 'justified belief' does not have to be used so that for some sense of 'knowledge' 'S is justified in believing that p' is entailed by 'S knows that p'. More particularly, I want to consider arguments in which 'justified belief' is taken not just in an evaluative, but also in a *volitional* way. S is volitionally justified in believing that p if and only if an action of S's is justified which is suitably related to his believing that p. On the face of it, arguments against evaluatively and volitionally justified belief pose more of a threat to our conjecture than those in which 'justified belief' is used in a non-evaluative and knowledge-dependent way. In section (1) I say more about volitionally justified belief, and clarify the notion of a suitably belief-related action. In section (2) I show how the sceptic can construct sound arguments against evaluatively and volitionally justified belief. In section (3) I show that these arguments fail the negation test.

(1) Volitionally justified belief

We are concerned with arguments against evaluatively and volitionally justified belief, with those arguments against justified belief in which 'N is not justified in believing that P' negates an evaluation, and in which 'justified belief' is taken in a volitional way. I will in fact take it that in all arguments against volitionally justified belief the conclusion negates an evaluation. S is volitionally justified in believing that p if and only if an action of S's is justified which is suitably related to his believing that p. As applied to belief-related action, 'justified' means nothing other than what it means as applied to any other kind of action. It is used to express an evaluation, and usually a moral

evaluation. We saw that sometimes epistemologists take ‘the’ concept of justified belief to be both normative and somehow *analogous* to a moral concept. But if when we say that an action is justified we make a moral evaluation, then the concept of volitionally justified belief isn’t just analogous to a moral concept. It *is* a moral concept. Sometimes ‘the’ concept of justified belief is taken as the concept of a belief whose justification abstracts from all moral considerations other than intellectual duties. I will take it for the sake of simplicity that *no* morally relevant considerations are ignored in the evaluation of the belief-related action whose justification makes a belief volitionally justified.

S is volitionally justified in believing that p if and only if an action of S’s is justified which is suitably related to his believing that p. What is the belief-related action? Not, I assume, a willing on S’s part to start believing that p, which then immediately causes him to start believing that p. It isn’t incoherent to suppose that willing to believe is sometimes immediately effective: there is no logical obstacle to anything’s directly causing anything. It is however contingently false that we bring ourselves to believe things just by willing that we believe them. It is logically possible for a door to collapse as an immediate result of your willing it to collapse or of your deciding to break it down. But in fact this never happens. You have to do something else as well, such as kicking it or hitting it with a sledgehammer. Breaking a door down differs in this respect from raising your arm, which you can do without doing anything other than willing your arm to go up or deciding to raise it. Doubtless something else has to happen, before you raise it. Nerve impulses have to be transmitted. But you don’t have to perform the action of transmitting them, in order to raise your arm.

Some writers seem to disagree. Newman, for example, who talks of assent as ‘the mental assertion of an intelligible proposition...an act of the intellect direct, absolute, complete in itself, unconditional, arbitrary, yet not incompatible with an appeal to argument, and at least in many cases exercised unconsciously’.⁷

Assent is an act of the mind, congenial to its nature; and it as other acts, may be made both when it ought to be made, and when it ought not. It is a free act, a personal act for which the doer is responsible, and the actual mistakes in making it, be they ever so numerous or serious, have no force whatever to prohibit the act itself.⁸

Is ‘assent’ Newman’s word for an immediately effective willing to start believing? Or does he perhaps take assent to be a believing which is itself just one among such other species of willing as attending or deciding? In either case, there would I think be no such thing. It seems better to keep ‘assenting’ as the name for giving an affirmative answer to the question, ‘Is p true?’. The question may have been asked by someone else or it may have been self-addressed. The answer may

be public or just spoken in the mind. But if the answer is sincere, as it may not be, then it requires that you *already* believe that p.

The justified and suitably belief-related action in terms of which I define volitionally justified belief is then neither an immediately effective willing to start believing that p nor a believing which is itself somehow a species of willing. Nor is it an act of assent, in the sense of a spoken answer to the question, 'Is p true?' Nor is it an act of judging that p, an act of assertion which terminates a process of wondering whether p is true. Nor is it a decision to act *as if* p is true, or, as some writers call it, an act of accepting p.⁹ What then is it? I have in fact several different kinds of actions in mind. The action may be a decision on S's part not to prevent himself from starting to believe that p, or a decision not to prevent himself from continuing to believe that p. In these cases it will be a decision which enables some other event to cause or help cause S to start or go on believing that p. The action may, alternatively, be a willing on S's part to start believing that p, or a willing to continue to believe that p. In these cases it will be an action which indirectly causes S to start or continue to believe that p. All of these actions are more than just logically possible.

by dwelling upon a proposition continually and repeatedly, by considering again and again what it would be like if it were true and imagining in detail what it would be like (if you can), by acting as if the proposition were true on all occasions to which its truth or falsity is relevant, and by increasing the number of those occasions whenever possible—by such means you will gradually get into a state of believing the proposition. You will wake up one fine day and find that you do believe it.¹⁰

Deciding not to prevent yourself from starting to believe that p may be equally effective. You may for example allow yourself to start believing that p by deciding to do nothing to check the operation of your desire to believe that p, or even your desire for p to be true.¹¹ And it is just as easy to cause yourself to continue to believe, either by indirectly efficacious positive action or by deciding not to prevent yourself from continuing to believe.

It is sometimes suggested that if you do get or allow yourself to start believing that p, and do this without having any evidence for p, then you restrict your freedom with respect to what, after the event, you can believe about the origins of your belief. This may seem to show that whether or not the judgement that one is volitionally justified in believing that p involves an undesirable mode of evaluation, reflection may still weaken the desire to be thus justified, at least in the case where it would involve getting yourself to start believing without having evidence.

How exactly is your freedom supposed to be restricted? In its strongest version, the idea is this. If you get or allow yourself to start believing that p, and

do so without having any evidence for *p*, then the continued existence of your belief that *p* is incompatible with your simultaneously believing that it had this evidence-free origin. This strong doctrine is encouraged by some of the things that Bernard Williams says about ‘believing at will’. Beliefs, according to Williams, ‘aim at truth’.

If I could acquire a belief at will, I could acquire it whether it was true or not...If in full consciousness I could will to acquire a ‘belief irrespective of its truth, it is unclear that before the event I could seriously think of it as a belief, i.e. as something purporting to represent reality. At the very least, there must be a restriction on what is the case after the event; since I could not then, in full consciousness, regard this as a belief of mine, i.e. as something I take to be true, and also know that I had acquired it at will. With regard to no belief could I know—or, if all this is to be done in full consciousness, even suspect—that I had acquired it at will.’¹²

This Williams offers as an argument for the thesis that ‘it is not a contingent fact’ that there is no such thing as acquiring a belief at will. Whatever the merits of that argument, Williams does say something quite plausible about restrictions on what is the case after the event. It could well seem that there are similar restrictions on what is the case after the event in all cases of getting or allowing oneself to start believing, whether or not these are cases of believing at will or ‘believing just like that’. It could seem that for any *S* and any *p*, *S* can’t at once believe that *p*, believe that he believes that *p*, and believe that when he got or allowed himself to start believing that *p* he did not at that previous time believe that any proposition he believed was evidence for *p*.

There is an obvious objection. Suppose that although you did indeed originally get yourself to believe that *p* without then believing any proposition which you believed was evidence for *p*, you do nevertheless now believe a proposition which you now believe is evidence for *p*. Won’t your present belief about evidence for *p* allow you to believe that *p* and believe that you believe that *p* and yet still be aware of the evidence-free origin of your belief that *p*? Let us allow this objection, and retreat to the following principle: For any *S* and any *p*, *S* can’t believe that *p*, and believe that he believes that *p*, while believing that he acquired the belief that *p* without at that previous time believing that anything he believed was evidence for *p*, unless he now believes that something he believes is evidence for *p*. Call this the *Eventual Evidence Principle*. The suggestion now is that once you grasp that the Eventual Evidence Principle is true, you will see that there are severe disadvantages in getting or allowing yourself to start believing that *p* when you do not believe any proposition which is sufficient evidence for *p*. If you are contemplating such action, these disadvantages will put you off. What, you will ask yourself, might the proposition be which you must subsequently believe and believe to be sufficient

evidence for *p*? What ensures that it will be a true proposition, which really is evidence for *p*?

I think we should reject the Eventual Evidence Principle. It fails because the very acquisition of a belief may give you the means to supplement or reinterpret the story of that acquisition. What you believe may itself throw light on how you came to believe it. Suppose you believe there are poppies in the garden, and believe that you have this belief. You also believe that you acquired the belief that there are poppies in the garden without at that previous time believing that any proposition you believed was sufficient evidence for its truth. Since you do now believe that there are poppies in the garden, you may well think that it was precisely these poppies which caused you first to acquire the belief that they were there in the garden, and that just for this reason, and despite the fact that you did not then believe that any proposition you believed was sufficient evidence for this, you did not acquire the belief 'irrespective of its truth'. But if you do indeed now believe that you were originally caused to acquire your belief about the poppies by what made it true, you surely do not also need now to believe that something you now believe is evidence for the existence of poppies in the garden, in order to believe that you have the belief that there are poppies in the garden. And it is just as well that there is no such evidential requirement. Even if you do now believe that there was or still is a causal chain running from the poppies through your experience as of the poppies to your belief about the poppies, there does not seem to be anything here which allows you to think that anything you now believe is evidence for the existence of the poppies. 'I have an experience as of poppies in the garden' is hardly evidence for 'There are poppies in the garden'. More to the present point, the Eventual Evidence Principle also breaks down in certain cases in which, as you subsequently think, you not merely acquired the belief that *p* but actually got or allowed yourself to start believing it. Wagering in the Pascalian way, you get yourself to start believing that God exists, and without believing any proposition which you believe to be evidence for God's existence. More particularly, you cause yourself to start believing that God exists by deciding to do nothing to prevent your strong desire to believe that God exists from pushing you into actually believing this. How do things stand after you have thus caused yourself to start believing? Suppose they stand like this. You believe that God exists, believe that you believe it, and believe that you allowed yourself to start believing that God exists without then believing any proposition which you believed to be evidence for 'God exists'. Is this possible only if you have in the meantime come to believe that some proposition you believe is evidence for 'God exists'? That condition must indeed be satisfied, according to the Eventual Evidence Principle. But, given that you do now believe that God exists, you could, in the light of what you now believe, reinterpret the process by which you first came to believe it. You could see the belief-producing operation of your desire to believe that God exists as itself

dependent on divine assistance. You could even see your actual desire to believe as something whose divinely ordained function is to produce the desired belief.¹³ And if you do thus reinterpret the process of belief-acquisition, in the light of the proposition which you now believe, this won't necessarily ensure that you believe something which you believe to be evidence for the proposition in whose light you reinterpret the process. It isn't evidence for God's existence that you want to believe that he exists.¹⁴

I shall assume that one's desire for volitionally justified belief will not be undermined by thoughts about how, after getting or allowing oneself to start believing, one lacks the freedom to believe that one's belief had an evidence-free origin. One may indeed have moral objections to getting or allowing oneself to start believing without having evidence.¹⁵ But, as we will shortly see, what really undermines the desire for volitionally justified belief is more a meta-ethical than an ethical objection, and one which applies even in the case when the believer does in fact have evidence for what he believes.

(2) Arguments against volitionally justified belief

Arguments against justified belief have conclusions of this form: 'N is not justified in believing that P'. Suppose that 'justified belief' is here taken in a volitional way. We then get: 'No action of N's is both morally permissible and suitably related to his believing that P'. And in the last section I gave a general idea of what a suitably related action might be. To show how the sceptic can construct sound arguments against volitionally justified belief I go back to Chapter IV. In that chapter I discussed evidence-or-acquaintance arguments in which P stands for a doctrine of the objectivity of values. And various senses of 'evaluation' were specified on which your evaluation of a particular object requires you to assert a universal conditional. Some of these senses were purely assertive: the evaluator does nothing but assert a proposition. Others I called doubly attitudinal: the evaluator both asserts a proposition and expresses a desire, aversion or intention. Suppose for example that you evaluate A as good. There is an actual or possible doubly attitudinal sense of 'evaluation' on which you would then be doing two things: expressing a desire for A to exist, and asserting a proposition. The proposition might have this form: A has a property F such that (i) your knowledge that A is F caused you to desire A to exist, and (ii) for any person S and any action x, if S knew that x was F then he would want it to exist. If we also stipulate that the universal conditional is contingent, then it is fairly clear that you would then be asserting a false proposition. If the universal conditional is contingent then human diversity would ensure that either (i) or (ii) is false. If this is right, then we should be able to find or devise a sense of 'evaluation' on which you can't evaluate a belief-related action as morally permissible without asserting a contingently false proposition. And that in turn will enable us to construct sound arguments against volitionally justified belief. Will these arguments be properly

sceptical? Will they be ones in which philosophy does some work? That question we can leave until we have set them up.

What universal conditional might you assert if you said that N's belief-related action was morally permissible? What conditional might you assert if you said, of *any* action, that it is morally permissible?

To say that an action is morally permissible could be just to deny that it is morally impermissible, to deny that it is morally obligatory to refrain from performing it. In this case no universal conditional would be asserted. What you would assert is just the negation of the proposition that you would have asserted if you had evaluated refraining from the action as morally obligatory. On a doubly attitudinal sense of 'evaluation', you might when you evaluate refraining from A as morally obligatory express an intention as well as assert a proposition. The intention might be either an intention to refrain from A, or an intention that were it in your power to do something relevantly similar to refraining from A then you would do it. (Maybe it is my refraining that you are evaluating. Then you can't intend more than to do something relevantly similar were that in your power.) If that were the intention, what would the simultaneously asserted proposition be? Perhaps it would be that refraining from A has some property F such that (i) your knowledge that refraining from A was F caused you to have this intention, and (ii) for any action x and any person S, if S knew that x was F then either he would intend to x or he would intend that were an action which was F in his power then he would perform it. If saying that A is morally permissible is just denying that it is morally obligatory to refrain from it, then the only proposition asserted would be that refraining from A has no property such that (i) and (ii). And no amount of human diversity can make *that* proposition false.¹⁶

To construct our argument we must then take 'morally permissible' in a different way. We must take it so that to judge an action morally permissible is indeed partly to assert a universal conditional. Two possibilities suggest themselves.

To say that an action A is morally permissible might be both to express the *absence* of an intention and to assert a conditional linking absence of intention with knowledge. More particularly, it might be in part to express the absence of an intention to refrain from A, or the absence of an intention that were it in your power to do something like refraining from A then you would do it. The simultaneously asserted proposition might then be that refraining from A has some property F such that (i) your knowledge that refraining from A was F prevented you from having the intention to refrain from A, and (ii) for any person S and any action x, if S knew that x was F this would prevent him from having the intention to x, or prevent him from having the intention that were it within his power to perform an action which was F he would perform it.

The other possibility is to bring in the notion of a second-order intention. We could say that to evaluate A as morally permissible is in part to express

either an intention not to intend to refrain from A or an intention not to intend that were it within your power to refrain from an action relevantly similar to A you would refrain from it. And then we could say that the simultaneously asserted proposition might be that refraining from A has a property F such that (i) your knowledge that refraining from A was F caused you to intend not to intend to refrain from A, and (ii) for any person S and any action x, if S knew that x was F then he would intend not to intend to x, or he would intend not to intend that were it within his power to do something like x he would do it.

Corresponding to these two possibilities, we have two suggestions as to what you assert if you say that N is volitionally justified in believing that P, and two kinds of arguments with conclusions of the form ‘N is not volitionally ‘justified in believing that P’. It will be quite enough to spell this out just for the second-order intention possibility. We have:

- (1) It is not the case that there is an action A and a property F such that (i) N’s knowledge that refraining from A was F made him intend not to intend to refrain from A, and (ii) for person S and any action x, if S knew that x was F then either he would intend not to intend not to x, or he would intend not to intend that were it within his power to do something like x he would do it;

So: (2) N is not justified in believing that p.

Such arguments are valid, if ‘justified’, in (2), means ‘volitionally justified’, and given the appropriate second-order intention interpretation of what it is to evaluate a belief-related action as morally permissible. And if we stipulate that the universal conditional in (1) is contingent, then it is fairly clear that human diversity ensures that (1) is true. We may guarantee the truth of (ii) by taking F as a gerrymandered, comprehensively conjunctive property, constructed out of all the properties which ever have underlain or ever will underlie the evaluation of an action as morally permissible. But that just ensures that N does not know that refraining from A is F.

Are such arguments properly sceptical? Like arguments for ignorance, arguments against justified belief are sceptical only if they are arguments in which philosophy is not idle. In Chapter IV, we were obliged to doubt the sceptical character of a parallel argument which moved from the contingent falsity of a universal conditional to the conclusion that N did not know the truth of a doctrine of dynamic unity. The inference was so obvious that philosophy had no work to do. But here philosophy is not idle. ‘Justified’ must be interpreted with some care before it is clear that it has a sense on which (2) follows from (1). The arguments are sceptical.

Does this give us sceptical arguments which pass the negation test? I denied in Chapter IV that we have any good reason for wanting the necessary truth of the universal conditionals which evaluators may be taken to assert. And on that basis I maintained that we have no reflectively indestructible desire for the truth of universal conditionals which may or may not be contingent. No parallel

considerations are available in the present case, because in (1) the universal conditional is contingent. This brings us to the question, which I circumvented in Chapter IV, of why we should not be content with the truth of conditionals which correlate knowledge and dynamic attitudes but are less than universal in their scope.

(3) The negation test

Suppose that N does initially desire to be volitionally justified in believing that P. As we can continue to suppose, the relevant universal conditionals correlate knowledge and second-order intention. Would reflection leave his desire unscathed? Only in that case will an argument against N's being volitionally justified in believing that P pass the negation test. But reflection will not leave his desire unscathed unless he can answer the question of why he should want there to be a true universal conditional correlating knowledge and second-order intention, rather than a true conditional which correlates the attitudes of almost everyone, or of most people, or even just of everyone in his own community. In order to show that this condition is not satisfied, I shall focus on a simpler case, and consider some possible answers to a parallel question about the universal conditional which you might assert if you were to evaluate A's existence as better than its non-existence. It is fairly plain that if in this case reflection would make you content with the truth of a less than universal conditional, then the same would be true in the case of judging that an action is morally permissible, and hence in the case of judging that belief is volitionally justified. If on reflection you are content with a less than universalizing evaluation of one thing as better than another, you would likewise be content with a less than universalizing evaluation of doing something as no worse than not doing it, and with a less than universalizing evaluation of doing something as permissible.¹⁷

Let us assume that if you judge that A's existence is better than its non-existence then you assert that A has a property F such that for absolutely any person S and any x, if S knew that x was F then he would prefer x's existence to its non-existence. Why should you want that to be true? Four answers suggest themselves, none of them satisfactory.

(A) COOPERATION

You want A to survive, but believe that it will not survive just by your own efforts: perhaps A is a pattern of life or an institution or even just a group of buildings which would disintegrate without a concerted effort in which many people play their part. You think that no one is likely thus to cooperate, or even to be non-obstructive, unless he actually prefers A's existence to its non-existence. Wouldn't A's survival be more likely if A had a property F such that

for any person S and any object x, if S knew that x is F then he would prefer x's existence to its non-existence?

Obviously there is no good reason here for wanting the truth of a *universal* conditional, which correlates the beliefs and preferences of the entire human race, past, present and future. What is the object whose continued existence depends upon the favour of absolutely everyone? Why not just hope to have enough friends and allies for the task in hand?

(B) SOLIDARITY

There is a certain solidarity in any group of people each member of which believes that every other member has the same fundamental preferences as himself. There is a certain solidarity even in groups each member of which believes no more than that all the other members would prefer what he prefers if they believed the truth. Consider then a group each member of which prefers F things and each member of which believes that every other member does or would with true beliefs prefer the F things which he himself prefers. It has a certain solidarity. Now suppose it is true that for *any* object x and any person S, if S knew that x was F then he would prefer x's existence to its non-existence. Then the group can be as large as you like, coextensive even with the whole human race.

Does this secure your desire for the universality of the conditional? Not unless you insist that there must be no limit to the size of the group. But why insist on that? Why want solidarity with the entire human race? The question sounds sinister. But there is a difference between human fellow-feeling and the solidarity of shared preferences. To feel towards me as a fellow human being need you want more than that I actually exist? Even if I wouldn't exist without some preferences or other, you might perfectly well want me to exist even if there are no particular preferences which you want me to have, and certainly without wanting me to share your own preferences. Solidarity is still possible even if there are no true universal conditionals of the relevant kind. It is enough that F is such that there are some or even a lot of people who, if they know that an object has F, prefer that object to exist rather than not. None of your ends will be shared by everyone, and some hardly anyone will share. But a few at least will be quite widely shared. Isn't that solidarity enough?

(C) IMPARTIALITY

Suppose your knowledge that A is F makes you prefer A to go on existing, and that A is something that will not survive just by your own efforts. A majority prefers A not to exist. You realize that A can indeed be preserved by the combined efforts of a relatively small group of your friends and allies, but only by overriding the opposition of the majority. This prompts democratic scruples.

Why, you ask, should you prefer the satisfaction of your own preference for A's existence to the satisfaction of the contrary preferences of someone else? Why shouldn't each person's preferences be treated equally? Suppose that there is an established voting rule which gives each person's preferences an equal weight. Why not let the matter be settled by a vote? But now you entertain the proposition that for any person S and any object x if S knew that x was F then he would prefer x's existence to its non-existence. Wouldn't the truth of this conditional make everything alright? Those who prefer A's non-existence show merely that they have failed to gain knowledge that A is F. A's existence would still be in the real interests even of those who prefer it not to exist. This would entitle you to ignore their actual preferences.

Once again, I think that universality is an exaggerated demand. It would be enough surely for A to serve the real interests only of those whose votes you would have to ignore. That doesn't require absolutely everyone who knows that something is F to prefer its existence to its non-existence.

There is, anyway, another difficulty with the impartiality thought. You are supposed to be worried about preferring the satisfaction of your own preference for A's existence to the satisfaction of the contrary preferences of others. We can distinguish however between your preferring A's existence to its non-existence and your preferring your preference for A's existence to be satisfied rather than not satisfied. From the propositions that you prefer A's existence to its non-existence, and that if A exists then your preference for A's existence will be satisfied, it does not follow that you prefer the satisfaction of your preference for A's existence to the non-satisfaction of that preference. It is true that if you prefer A's existence, and if A exists, then your preference is satisfied. But part at least of what you prefer about A must be something different from that, and may have nothing at all to do with either the satisfaction of your preferences or those of anyone else. The satisfaction of your own preference for A, as distinct from the existence of A itself, may indeed be something that you want not at all. If counting your own preference for A's existence as having more weight than the contrary preferences of anyone else means preferring the satisfaction of your own preference for A's existence to the satisfaction of someone else's preference, then this may well be something that you want not to do. But you don't have to do it, just because you prefer A's existence to its non-existence and act on this preference when you know that someone else has the contrary preference.

(D) MULTIPLICATION

A final pro-universalizing thought. Suppose that your knowledge that A is F makes you prefer A's existence to its non-existence, and that you also want there to be as many F-things as possible. Suppose further that for person S and any object x, if S knew that x was F then he would prefer x's existence to its

non-existence. Realizing as much, you may want the universal conditional to be true. You may think that if it is true then there will be more F-things than if it is not. Some people at least will act on their preferences, and there will be more preferences to act on if the universal conditional is true than if some people who believe that objects are F do not prefer the existence of these objects.

To see why the multiplication thought is not in the end likely to secure a desire for there to be true and universal conditionals, it is necessary to look quite carefully at the general conditions under which reflection destroys desire. When you want one thing for the sake of another you may either believe that there is a causal connection between the two or believe that there is a logical connection between the two. You may want a ladder in order to climb up to the window. You may want to checkmate her in order to win the game. I think that if you believe that something will never be actual you may easily still want it for its own sake, or as a logical condition of something else that you want for its own sake. But I also think that it is difficult to want one thing as a causal condition for something else if you think that the first thing will never be actual. If I want, for its own sake, the wall to be painted blue, my desire is not threatened by my certainty that it never will be. Nor will my certainty that I will never checkmate her affect my desire to do so when I want this as a logical condition of winning the game. But if I want a ladder as a causal condition of reaching the window, then my desire for a ladder is threatened by my conviction that no ladder will come my way. It seems then to be a true general principle that for any person S and any propositions p and q, if S wants that p, and not for its own sake but as a causal condition for its being the case that q, then his desire that p will not survive his conviction that p is false.

This principle is applicable to the multiplication thought. Suppose you want it to be the case that for any person S and any object x, if S knew that x was F then he would prefer x's existence to its non-existence. This isn't something that you will want to be the case for its own sake. Nor is it something that you will want to be the case as a logical condition of something else that you want to be the case for its own sake. So if you are impressed by the multiplication thought you will want this universal conditional to be true as a causal condition for maximizing the number of objects which are F. If, however, the relevant universal conditional is contingent you will think that human diversity ensures its falsity. And so if you want the universal conditional to be true as a causal condition for maximizing the number of F-things, then by our principle your desire is doomed.

Has this rather abstract discussion missed something essential out?

From a number of premises, including Burkean conservatism, concern for 'stability', worries about 'cultural imperialism' and post-modernist 'irony', [contemporary anglophone political philosophers] ...have tended to converge on the conclusion that the so-called 'Enlightenment project' of addressing the reason of every human being of sound mind

was a gigantic error and that all we can aspire to is articulating the shared beliefs of members of our own society ...If this were true, it would have devastating implications for movements dedicated to securing human rights in countries where they are not respected and never have been. The government would be able to point to a long tradition of repression and cite my philosophical colleagues to show the absurdity of counterposing mere abstract principles to the dense network of lived morality.¹⁸

Can only the universal conditional arm us against paralysis and indifference in the face of repression? Surely not. Of course there are oppressors with no qualms. But that doesn't mean that even in countries with long traditions of oppression qualms are too rare to make criticism futile. Nor by rejecting the universalizing conception of evaluation do we condemn ourselves just to articulating the shared beliefs of our own society. Nor need we abandon our own commitments just because we stop believing that anyone would share them who believed what we believe about the facts.¹⁹ To fight your enemies, you do not have to believe that they are already your potential allies.

Nothing seems to underlie our initial desire for the truth of universal knowledge-preference conditionals but these various and dubious thoughts about cooperation, solidarity, impartiality, and multiplication. It seems then that we have no reflectively indestructible desire to make universalizing judgements of betterness. But that in turn suggests that we have no such desire to make any universalizing evaluation. And in that case we have no reflectively indestructible desire to make universalizing evaluations of beliefs as volitionally justified. The sceptic's arguments against volitionally justified belief will therefore fail the negation test.

VII

FROM RATIONAL ACTION TO RATIONAL BELIEF

There are beliefs which, were they true, would be knowledge. There are beliefs which it is permissible for you to get or allow yourself to start or go on having. Both kinds of beliefs are called justified. And both are called rational. ‘Justified belief’ and ‘rational belief’ are often used interchangeably. ‘Rational belief’ can however be taken in another way. This we can see if we consider what it is to call an action rational. To call an action rational *can* be to evaluate it. But it can also be to say no more than that it stands, or is believed by the agent to stand, in a certain relation to what are in fact the agent’s own ends. It can for example be to say no more than that the action is thought by the agent to be at least as conducive to what are in fact his own ends as any other action which he is able to perform instead. To call S’s belief that p rational can be to evaluate as rational an action of S’s which is suitably related to his believing that p. That makes ‘rational belief’ mean something not too different from what in the last chapter I labelled volitionally justified belief. But to call S’s belief that p rational can also be, not to make a value judgement, but just to say that an action suitably related to his believing that p is thought by S to stand in a certain relation to S’s own ends. Here we have a volitional but non-evaluative interpretation of ‘rational belief’. Volitional uses of ‘justified belief’ are always evaluative because to call an action justified is always to evaluate that action. Volitional uses of ‘rational belief’ are not always evaluative.

Can we build sound, sceptical and negation-proof sceptical arguments against non-evaluatively and volitionally rational belief? If so, my general conjecture about sceptical arguments falls. How to construct such arguments? One way might be this. First find a sense of ‘rational’, on which acting rationally looks as if it might be something that we want to do but can’t. Then make sure that, when action is indeed relevantly belief-related, we really do have a vain desire for it to be in this sense rational. This is the method that I now try out.

In the first section of the present chapter, I survey some senses of ‘rational action’. Section (2) is about various unattainable but in the end undesirable kinds of infinitely regressive practical rationality. Section (3) presents what

from our falsification-seeking standpoint is a more promising candidate: an attractive kind of practical rationality not easily attainable except by beings who believe much more than we do about the future. In the following chapter I consider whether, by defining volitionally rational belief in terms of this last kind of practical rationality, we can finally build a sound and sceptical argument which gets through the negation test.

(1) Rational action

Consider sentences of the form ‘It was rational for N to A’. Such sentences can be used to evaluate N’s action A. When they are thus used, then the suggestion may also be made that N used his reason to discover the truth of what is asserted about his action, or even that he made that discovery by using Reason, conceived of as a special cognitive power distinct from our ordinary powers of sensation and logical thought. But for all that, and despite what some writers say, such sentences can also be used in a non-evaluative way.¹

Let me start by surveying some non-evaluative senses of ‘rational action’. I begin with an instrumental sense. In this, your Aing is rational if and only if it is at least as conducive to your ends as any other action which you are able to perform instead. Your ends are those things which you want for their own sakes, and the wants in question are those which you have at the time of your Aing. No restrictions are placed on the nature of the agent’s ends. They may or may not be selfish or self-abnegating, may or may not refer essentially to the agent himself. They may or may not refer to any particular individuals or groups or times or places. To say that your Aing is in this sense rational is not to express a favourable or hostile attitude; nor is it to assert a conditional about the attitudes of some group to which you belong.

The ‘for its own sake’ terminology is perhaps a bit clumsy when it comes to elucidating non-evaluative practical rationality. So instead of saying that N wants that p for its own sake I will say that N has an internally unconditional want that p. And instead of saying that he wants that p for the sake of something else I will say that his want that p is internally conditional. This still doesn’t quite settle things, because ‘N wants that p for its own sake’ is ambiguous: the possessive pronoun can refer either to p’s being the case or to N’s desire that p. ‘N has an internally unconditional want that p’ I take to mean the same as ‘N wants that p for the sake of p’s being the case’, rather than ‘N wants that p for the sake of his wanting that p’. In this terminology, N’s Aing is in the instrumental sense rational if and only if it is at least as conducive to those things for which he has internally unconditional desires as any other action which he is able to perform instead.²

It is not sufficient for N’s Aing to be instrumentally rational that he has an internally unconditional desire to A. Nor is it even sufficient that he has an internally unconditional desire to A which is strong enough somehow to outweigh the combined strength of any internally unconditional desires he

has to perform other actions which he is able to perform. This is because N's Aing is instrumentally rational only if he has an internally unconditional desire for something to which his Aing is conducive, and a desire to A is not a desire for something to which Aing itself is conducive. We might distinguish at this point between instrumental and non-instrumental rationality, and say that N's Aing is non-instrumentally rational if and only if he has an internally unconditional desire to A, or if and only if he has an internally unconditional desire to A which is at least as strong as any internally unconditional desire he has to perform any other action he is able to perform. But it is also possible to construct a single sense of 'rational action' which allows us to take into account not only the action's conduciveness to other things for which the agent has an internally unconditional desire, but also any internally unconditional desire which the agent may have to perform the action itself.

To construct this single sense, we need to distinguish between an action's satisfying a want, and its being conducive to the satisfaction of a want. Your want that p is satisfied by its being the case that p. So if you want to A then your Aing will itself satisfy that want. But suppose that as well as wanting to A, you want to have x. If your Aing is conducive to your having x, then as well as itself satisfying your desire to A, your Aing will be conducive to the satisfaction of your desire to have x. Let us now say that your action *makes for* the satisfaction of some want you have if either it satisfies some want you have or it is conducive to the satisfaction of some want you have. On our initial instrumental sense of 'rational action', N's Aing is rational if and only if it is at least as conducive to the satisfaction of N's internally unconditional wants as any other action which he is able to perform instead. On our new and broader sense, N's Aing is rational if and only if no other action which he is able to perform instead makes more for the satisfaction of his internally unconditional wants. I will call this new and broader sense the *basic* sense of 'rational action'. If N has an internally unconditional desire to A, then his Aing can make for the satisfaction of some internally unconditional want he has, and may therefore be basically rational, even though it is not conducive to the satisfaction of any internally unconditional want he has, and hence is not rational in the initial instrumental sense.

Philosophers and social theorists often try to distinguish between instrumental and expressive rationality. An action is said to be expressively rational when it expresses the agent's commitment to an end, or when it is itself the creative adoption of an end.³ I do not think we need a separate expressive sense of 'rational action'. So far as I can see, expressively rational actions are, or are likely to be, basically rational, if not indeed instrumentally rational in my initial sense. In 1956, on the streets of Budapest, a woman spits at a photograph of Rakosi. Her action expresses her commitment to the freedom of Hungary. Is it identical to the expression of her commitment, or conducive to this? Suppose she has an internally unconditional desire to express her commitment. Then if

her action is conducive to the expression of her commitment it may well be instrumentally rational, even if it does nothing to help free her country. If on the other hand her action is identical to the expression of her commitment, then even if it is not instrumentally rational it will satisfy her desire to express her commitment and will thus be basically rational.

Nozick distinguishes between the instrumental rationality of an action and a kind of practical rationality which is partly determined by what he calls symbolic utility. I am also doubtful about the need for this distinction, once basic rationality has been defined. According to Nozick, an action has symbolic utility if it ‘symbolizes a certain situation, and the utility of this symbolized situation is imputed back, through the symbolic connection, to the action itself’.⁴ Neurotic hand washing symbolizes a removal of guilt, whose hypothetical utility determines the actual symbolic utility of the handwashing itself.

Kant felt that in acting morally a person acts as a member of the kingdom of ends, a free and rational legislator. The moral action does not *cause* us to become a (permanent) member of that kingdom. It is what we would do as a member, it is an instance of what would be done under such circumstances, and hence it symbolises doing it under those circumstances. The moral acts get grouped with other possible events and actions and come to stand for and mean them. Thereby, being ethical acquires a symbolic utility commensurate with the utility these other things it stands for actually have.⁵

Suppose that N’s Aing has more symbolic utility than any other action he is able to perform instead. Then it looks as if, other things being equal, his Aing will be basically rational, if not indeed instrumentally rational in my initial sense. Suppose, what may well be true, that the removal of guilt is something for which the handwasher has an internally unconditional desire. Is his handwashing conducive to the symbolization, or identical with this? In the first case the action may well be instrumentally rational, even if it is not conducive to the actual removal of guilt. If on the other hand the action is identical to the symbolization, then even if it is not instrumentally rational it may well be basically rational.

My initial definition of basic rationality is rough, and I will try to improve it in a moment. But it is already quite easy to see that by modifying the basic sense we can construct many other senses of ‘rational action’. Some of these further senses require belief on the agent’s part: there is for example a sense on which N’s action is rational if and only if he believes that it is basically rational. Some senses which require belief on the agent’s part are also infinitely regressive: they require the agent to have performed an infinite number of actions. There is for example a sense in which N’s Aing is rational if and only if (1) he believed that (1.1) it would be basically rational, and (2) he decided

whether to go on believing that (1.1), and (3) he believed that (3.1) the decision referred to in (2) was basically rational, and (4) he decided whether to go on believing that (3.1), and (4) he believed that (4.1)...and so on ad infinitum.

We could also generate further senses of ‘rational action’ by placing rationality conditions on the wants in terms of which basic rationality is defined. Hume remarked that reason is and ought only to be the slave of the passions. And sometimes this is glossed as the doctrine that although your actions can be rational in the sense of being conducive to what you want, your want that p can be rational only in the sense that it depends on your having a true or justified or rational belief about how its being the case that p is related to other things that you want. If you want that p for the sake of its being the case that p, have an internally unconditional want that p, then there is no sense in which your want can be either rational or irrational. This seems unnaturally restrictive. Sentences of the form ‘It would be rational for N to A’ can be used evaluatively. But if actions can be evaluated as rational so can wants, including wants which are internally unconditional. If ‘It would be rational for N to A’ can be used in an abstractly evaluative way, to mean something like ‘It would be good for N to A’, then N’s internally unconditional want can also be evaluated as rational, with nothing more meant than that it is good that N has that want. Apart from this, an internally unconditional want can perfectly well be rational in various purely descriptive senses. It may for example be rational in the sense that your having it is at least as conducive to the satisfaction of your wants as your not having it. Or it can be rational in the sense of being generated or sustained by a descriptively rational action. It may even be that part of what makes it rational to sustain an internally unconditional want that p is that your continuing to have the want makes it more likely that you will do something to ensure that p.

Are there senses of ‘rational action’ which can’t be constructed out of or subsumed under the basic sense? Perhaps we are willing to call an action rational just to signal that it is not ‘collectively self-defeating’, to signal that it is enjoined by a rule the general observance of which would be at least as conducive to what each person wants as the general observance of any rule which enjoins some other action which the agent is able to perform. But with senses of this Kantian kind I am not concerned.

To improve on the initial definition of basically rational action, it is necessary to refine the notion of one action’s making more than some other action for the satisfaction of the agent’s internally unconditional wants. In decision theory, one is often presented with some such scheme as this.

	S1	S2	...SN
A	A & S1	A & S2	A & SN
B	B & S1	B & S2	B & SN

A and B are actions, S1...SN are mutually exclusive and jointly exhaustive states of the world, each of which is believed by the agent to be possible. The

subjective expected utility of A is then defined as the sum of the series $D(A \& S_1) P(S_1/A) \dots D(A \& S_N) P(S_N/A)$, where $D(A \& S_i)$ measures the utility that the agent believes before acting that $(A \& S_i)$ would have for him, and $P(S_i/A)$ measures the strength of the agent's belief that S_i given A. Suppose we take utility for the agent to mean satisfaction of his internally unconditional wants, and reinterpret the symbolism so that in the series $D(A \& S_1) P(S_1/A) \dots D(A \& S_N) P(S_N/A)$, $D(A \& S_i)$ measures the actual utility for the agent of $(A \& S_i)$, relative to the internally unconditional wants that he has at the time of A_{ing} , and so that $P(S_i/A)$ measures the objective rather than the subjective probability of S_i given A. Could we then say that the degree to which A makes for the satisfaction of the agent's wants is defined as the sum of the series? The difficulty is that the scheme represents only a finite number of possible states of the world. The agent may indeed not be able to take more than a finite number of states of the world into account when forming his subjective probabilities. But there may not be any limit to the different ways in which it is objectively possible for wants to be satisfied in consequence of an action.

Perhaps we should respond on these lines. Let a possible outcome of an action be a series of actual or possible events each of which is or would be later than its predecessor and each of which is an event made more probable by the action. Each action has perhaps an infinite number of possible outcomes with no common member. And each such possible outcome of an action can itself consist of an infinite number of possible events. Let the U subset of a possible outcome of N's A_{ing} be that subset which consists of all the events in the outcome which have utility for N. We can now say that the basically weighted utility which N's A_{ing} has for him is measured by the strength of any internally unconditional desire which he has for A itself and the sum of a series which satisfies these conditions: (i) each member of the series is a product whose factors are a number measuring the amount of utility for N of the U subset of some possible outcome of his A_{ing} , and a number measuring the objective probability, conditional on N's A_{ing} , of that U-subset; (ii) the possible outcomes, the utilities and probabilities of whose U-subsets determine these products, have no common members; and (iii) there is no U subset of a possible outcome of N's A_{ing} whose utility for N and probability is not measured by some member of the series. And finally we can say that N's A_{ing} is basically rational if and only if it has at least as much basically weighted utility for him as any other action which he could have performed instead.⁶

(2) Infinite regression

Is there among the various senses of 'rational action' which I have now surveyed a sense on which acting rationally looks to be something that we want to do but can't? If so, the sceptic could seize on it, and hope to use it in the

construction of a sound and negation-proof argument against volitionally rational belief. In his writings on freedom, Galen Strawson makes strong claims for the unattainability and profound desirability of what is in effect a certain infinitely regressive practical rationality. I will begin with what he says.

According to Strawson, ‘true responsibility’ is the kind of freedom which ‘most people think matters most to them’.⁷ And it is, he thinks, a kind of freedom which it is impossible for anyone to have. An action is truly responsible if and only if (i) it is rational in the sense of being a ‘function’ of how the agent is, ‘mentally speaking’; (ii) the agent has consciously and explicitly chosen, in the light of some principles of choice P1 (‘preferences, values, proattitudes’), to be the way he is, mentally speaking, ‘in certain respects’; (iii) he has consciously and explicitly chosen in the light of some principles of choice P2, to have principles of choice P1; (iv) he has consciously and explicitly chosen, in the light of principles of choice P3, to have principles of choice P2; and so on, ad infinitum. True responsibility can’t be had because it requires the actual completion of an infinite series of choices.⁸ If Strawson is right, there is an infinitely regressive kind of descriptive rationality which all actions lack, but which we want at least some of our actions to have. Is this what the sceptic about volitionally rational belief is looking for?

There is I think a certain unclarity in Strawson’s definition of true responsibility. Is ‘choosing to have a principle of choice’ choosing to go on having a principle you already have? Or is it choosing to start having a principle you do not already have? If the former, P2 is not necessarily distinct from P1, and true responsibility is not necessarily infinitely regressive. In choosing whether to continue to have principle of choice P1 you may apply, not some further principle P2, but principle P1 itself. Suppose that, as Strawson seems to allow, having a principle of choice is wanting things of some particular type. And suppose accordingly that having P1 is wanting things of type X. Then to apply P1 to the question of whether to continue having P1 is simply to determine whether continuing to want things of type X is in fact conducive to things of type X. There is nothing problematic about that. Things are different if choosing to have a principle of choice is choosing to start having a principle which you do not already have. In this case true responsibility is indeed infinitely regressive. But then it isn’t obvious that we want to be truly responsible: why not manifest your autonomy just by choosing whether to continue to have the principles you already have? In the rest of this section I will take a more systematic look at infinitely regressive practical rationality.

Basic rationality was defined in the last section. N’s Aing is basically rational if and only if no action which he is able to perform instead makes more for the satisfaction of his internally unconditional wants. We do, I assume, feel some attraction to basic rationality. Suppose that just two actions are open to you: A and B. Then you may well want that either (1) you A and your Aing would be basically rational, or (2) you B and your Bing would be basically rational. It may however occur to you that you are free either to cause yourself

to continue to have, or to cause yourself to stop having, the wants relative to which your Aing or your Bing would be basically rational. More briefly, it may occur to you that you are free either to keep your ARB wants or to stop having them. More briefly still, it may occur to you that you are free either to K(ARB/W) or to S(ARB/W). And once you have had this thought you may well want that either (1.1) you A and your Aing would be basically rational and you K(ARB/W) and your K(ARB/W) ing would be basically rational, or (2.1) you B and your Bing would be basically rational and you K(ARB/W) and your K(ARB/W) ing would be basically rational, or (3) you S(ARB/W) and your S(ARB/W) ing would be basically rational. But now it may further occur to you that you are free either to keep or to stop having the wants relative to which your K(ARB/W) ing or your S(ARB/W) ing would be basically rational. Or more briefly, that you are free to keep or stop having your K(ARB/W)RS(ARB/W) wants. Or more briefly still that you are free either to K(K(ARB/W)RS(ARB/W)/W) or to S(K(ARB/W)RS(ARB/W)/W). And once you have had this further thought you may further want that either (1.2) you A and your Aing would be basically rational and you K(ARB/W) and your K(ARB/W) ing would be basically rational and you K(K(ARB/W)RS(ARB/W)/W) and your K(K(ARB/W)RS(ARB/W)/W) ing would be basically rational, or (2.2) you B and your Bing would be basically rational and you K(ARB/W) and your K(ARB/W) ing would be basically rational and you K(K(ARB/W)RS(ARB/W)/W) and your K(K(ARB/W)RS(ARB/W)/W) ing would be basically rational, or (3.1) you S(ARB/W) and your S(ARB/W) ing would be basically rational and you K(K(ARB/W)RS(ARB/W)/W) and your K(K(ARB/W)RS(ARB/W)/W) ing would be basically rational. And now it may occur to you that you are free either to keep or stop having the wants given which either your K(K(ARB/W)RS(ARB/W)/W) ing or your S(K(ARB/W)RS(ARB/W)/W) ing would be basically rational. And now, it seems, you are embarked on a train of thought which, if time permitted, could go on forever.

Time doesn't permit, and indeed the train of thought will come quite quickly to a halt. It seems though that when it has halted, and when you reflect about how it could have continued, you may find that you would like to have taken not just all the actions necessary to satisfy that finite number of desires revealed or generated before it actually stopped, or all the actions necessary to satisfy the finite number which would have been revealed or generated if it had stopped at some later time, but all the actions necessary to satisfy the infinite number of disjunctive desires which you would have had if the train of thought had continued forever. For, given any desire revealed or generated by the time the train of thought does stop, a further desire, which would have been revealed or generated if it had not stopped at that point, would have been just as natural. If this is right, then reflection on how your train of thought could have continued, and on the arbitrariness of any particular stopping place, may leave you with a desire to have performed an infinite number of actions, to have performed not just the action necessary for it to be true that either (1.1) or (2.1) or (3), but the

two actions necessary for it to be true that either (1.2) or (2.2) or (3.1), and not just these but the three actions necessary for it to be true that either (1.3) or (2.3) or (3.2), and so on. Reflection may leave you wanting either to be rational in Aing or rational in Bing, in an infinitely regressive sense of 'rational'.

On the other hand, it is clear enough that such a desire will never be satisfied. Since there is no time at which you will have lived forever, you will never have had the time to take the infinite number of actions which you would need to have taken in order for the desire to be satisfied. There is it seems a minimum and finite period of time such that no action can be performed in less time than that. So even if actual infinities can exist outside mathematics you will never have performed the requisite infinite number of actions.

Is this desire reflectively indestructible? Consider again the formative train of thought. It occurs to you that you are free either to cause yourself to continue to have or to cause yourself to stop having the wants relative to which your Aing or your Bing would be basically rational. So the rationality you want for your Aing or your Bing is a rationality relative to the internally unconditional wants you already have. You are interested in whether you would have internally unconditional desires for the events in the possible outcomes of Aing or Bing if you had true beliefs about them and if there were no change in the kinds of things for which you have internally unconditional wants at the time of wanting your Aing or your Bing to be basically rational. But on this supposition your ARB wants, the wants relative to which your Aing or your Bing would be basically rational, are just *the same as* your $K(\text{ARB}/W)\text{RS}(\text{ARB}/W)$ wants, the wants relative to which it would be basically rational for you to keep your ARB wants or stop having your ARB wants. Both your ARB wants and your $K(\text{ARB}/W)\text{RS}(\text{ARB}/W)$ wants will be just the whole set of internally unconditional wants which you have at the time of worrying about the basic rationality of your Aing or Bing. We can take it that your ARB wants and your $K(\text{ARB}/W)\text{RS}(\text{ARB}/W)$ wants are identical. But now it follows that

- (1.1) you A and your Aing would be basically rational and you $K(\text{ARB}/W)$ and your $K(\text{ARB}/W)$ ing would be basically rational
will be identical to
- (1.2) you A and your Aing would be basically rational and you $K(\text{ARB}/RW)$ and your $K(\text{ARB}/W)$ ing would be basically rational and you $K(K(\text{ARB}/W)\text{RS}(\text{ARB}/W)/W)$ and your $K(K(\text{ARB}/W)\text{RS}(\text{ARB}/W)/W)$ ing would be basically rational.

Similarly, (2.1) will be identical to (2.2), and (3.2) will be identical to (3.1). But from this it follows that the disjunction ((1.3) or (2.3) or (3.2)) will be identical to the disjunction ((1.2) or (2.2) or (3.1)). And by parallel reasoning we can show that in the supposedly infinite series of disjunctions each member after ((1.3) or (2.3) or (3.2)) would be identical to its immediate predecessor. This means that really there is no infinite series of distinct

disjunctions whose truth requires you to perform an infinite number of actions. The non-existence of the infinite series follows from the condition that basic rationality is to be assessed relative to the internally unconditional wants you already have. So in wanting a basic rationality which both satisfies that condition and is infinitely regressive in the way that you initially suppose, you have an incoherent object of desire. And by the principle of INCOHERENCE, already defended in Chapter II, your actual desire will not survive your recognition that this is so.

I now turn to a slightly different train of thought, equally capable of prompting a desire to perform actions rational in an infinitely regressive sense, but this time immune to the corrosive power of incoherence. The train of thought just criticized makes you want a seemingly infinite series of disjunctions to be true. The first two disjunctions in the series are: either (1) you A and your Aing would be basically rational, or (2) you B and your Bing would be basically rational; and either (1.1) you A and your Aing would be basically rational and you K(ARB/W) and your K(ARB/W) ing would be basically rational, or (2.1) you B and your Bing would be basically rational and you K(ARB/W) and your K(ARB/W) ing would be basically rational as your S(ARB/W) ing, or (3) you S(ARB/W) and your S(ARB/W) ing would be basically rational. We saw that your ARB wants (those relative to which your Aing or your Bing would be basically rational) will be identical to your K(ARB/W)RS(ARB/W) wants (those relative to which it would be basically rational for you to keep your ARB wants or basically rational for you stop having them). This is because both will be identical to the internally unconditional wants which you have when you want that (1) or (2). It was this identity which telescoped the seemingly infinite series of disjunctions. Suppose, however, that as before you want that either (1) you A and your Aing would be basically rational, or (2) you B and your Bing would be basically rational, but that instead of wanting to make a basically rational decision about keeping or dropping your ARB wants you want to have freely chosen to acquire your ARB wants and you want this choice to have been basically rational. More briefly, you want to have I(ARB/W) ed, and want your I(ARB/W) ing to have been basically rational. It can't be assumed that your ARB wants are just the same as your I(ARB/W)RN(ARB/W) wants, the wants relative to which your I(ARB/W) ing would have been basically rational. Your ARB wants are just the internally unconditional wants which you now have. But if you want your I(ARB/W) ing to *have been* basically rational then you want it to have been basically rational not relative to the set of internally unconditional wants which you now have, but relative rather to the set of internally unconditional wants which you would have had *before* you acquired your ARB wants and while you were deliberating about whether to acquire them. This being so, a genuine regress is liable to develop.

Consider again. You want that either (1) you A and your Aing would be basically rational, or (2) you B and your Bing would be basically rational. And suppose you want that at some time or other in the past you have freely caused or allowed yourself to start having the ARB wants, and that this action of yours was basically rational. More briefly, you want that (4) You I(ARB/W)ed, and your I(ARB/W)ing was basically rational. Now suppose that the I(ARB/W)RN(ARB/W) wants, the wants relative to which your I(ARB/W)ing would have been basically rational, are ones which you want to have had just before I(ARB/W)ing. Just as you want that you I(ARB/W)ed, so you want that you caused or allowed yourself to start having the I(ARB/W)RN(ARB/W) wants. For short, you want that you I(I(ARB/W)RN(ARB/W)/W)ed. And just as you want your I(ARB/W)ing to have been basically rational so you want your I(I(ARB/W)RN(ARB/W)/W)ing to have been basically rational. So as well as wanting that (4), you want that (5) you I(I(ARB/W)RN(ARB/W)/W)ed and your I(I(ARB/W)RN(ARB/W)/W)ing was basically rational. And so on, were time to permit, ad infinitum. This time there is a genuine infinite series of distinct desiderata. And if you reflect about the arbitrariness of any particular stopping place then you may well be left with a desire which, as you think, could only be satisfied if you had performed an infinite number of actions. You are in fact liable to be left with the unsatisfiable desire that either you A or you B and that whichever of those actions you take, that action will be both basically rational and rational in an infinitely regressive sense.

Is this a reflectively indestructible desire? Again, I think not. Why do you want the kind of infinitely regressive rationality which you believe you will never have? I think that you want it so as *either* to be *or* to have been autonomous with respect to your wants. More particularly, you want it for the sake of its being true *either* that (i) you have freely and with basic rationality acquired the ARB wants you now have, and likewise the wants relative to which that acquisition was basically rational, and so on, ad infinitum, *or* that (ii) you will now freely and with basic rationality keep or drop the ARB wants you now have, and likewise the wants relative to which that keeping would be basically rational, and so on, ad infinitum. But the infinite series in (ii) is as bogus as the infinite series in the train of thought which prompted the first conflict considered in this section. There is no genuinely infinite series here because the ARB wants are just all the internally unconditional wants you now have, and hence identical to the wants relative to which it would be basically rational to keep them. And once (ii) has been telescoped, the disjunctive goal shrinks to this: either (i) or (iia) you will now freely and with basic rationality keep or drop the ARB wants you now have. And now the goal is realisable just by (iia). But (iia) is easier to achieve than the non-finitized (i), and I think that once you realize this you will stop wanting the infinitely regressive rationality which you initially wanted. The general principle here operative is on some such lines as these. If you want one thing

only for the sake of a second thing, and then come to believe that a third thing is both equally conducive to the second thing and easier to bring about than the first thing, then your desire for the first thing tends to evaporate.

A likely objection to this solution is that we want to be the ultimate creators or originators of our own ends, rather than just to be able freely to keep or drop the ends which we already possess. I suspect though that to create your own ends, in the relevant sense, can indeed be just to choose to do what you already want to do. By choosing to act according to *this* want, when an opposing desire is pushing you towards a different action, you are in a sense making the first want your own, even perhaps helping to create your actual self. This I think fits in with a notion of self-creation which plays a central part in Robert Kane's explication of the traditional idea of free-will.⁹ It is true that the choices by which, according to Kane, we create our ends or purposes are free only if undetermined. And if, as I am assuming, the agent wants them to be basically rational, then problems arise about how it can be true both that they are undetermined and that their rationality plays a part in their explanation. But these are issues which I shall not pursue.¹⁰

So far in this section I have been considering senses of 'rational action' on which the agent acts rationally only if he has performed an infinite number of want-related actions. I have not found a sense of 'rational action', thus desideratively regressive, which makes rational action an object of hopeless but reflectively indestructible desire. So if the sceptic were to define volitionally rational belief in terms of the desideratively regressive descriptive rationality of belief-related action, and on that basis construct a sound argument against volitionally rational belief, then this argument would not pass the negation test. It is also possible to devise regressive senses of 'rational action' on which an agent acts rationally only if he has performed an infinite number of *belief*-related actions. That may allow the sceptic to construct further arguments against volitionally rational belief. But I will postpone what I have to say about this further possibility until the next chapter, and turn now to a non-regressive sense, which may also serve the sceptic's purposes.¹¹

(3) Belief shortage

N's Aing is basically rational if and only if no other action which he is able to perform instead makes more for the satisfaction of his internally unconditional wants. And there is a sense of 'rational action' on which the agent acts rationally only if he has the true belief that it would be basically rational so to act. Initially at least, you may want to act rationally, in this belief-dependent sense. Suppose that just two actions are open to you: A and B. You may think that it will increase your chances of actually doing what is basically rational if either you form a true belief that Aing would be basically rational or a true belief that Bing would be basically rational. But how easy is it for us to form *any* belief, true or false, about the basic rationality of our actions?

Your action is basically rational only if it would have at least as much basically weighted utility as any other action you could take instead. As I explained in section (1), the basically weighted utility of an action depends on the utility for the agent of the U-subsets of its various possible outcomes, on the utility of these subsets of each of its possible outcomes which consist of events which would have utility for the agent. But any action will have an enormous and perhaps infinite number of possible outcomes, and each possible outcome of any action will have an enormous and perhaps infinitely numerous U-subset. The agent cannot imagine more than a tiny part of the U-subset of any of his options. This may well prevent him from believing, of any of these options, that it is basically rational.

It is true that often we act or try to act just on our beliefs about the utility of the few possible outcomes that we can imagine, and ignore all possible outcomes that we cannot imagine. But it doesn't escape us that there are indeed possible outcomes that we can't imagine, and that the basic rationality of our actions depends as much on what we can't imagine as on what we can. Suppose that just two options are open to you: A and B. And suppose you believe that, relative to those possible outcomes of Aing and Bing which you can imagine, your Aing would be basically rational. You may perhaps find yourself assuming that the possible outcomes that you can't imagine 'cancel each other out', so that relative to the outcomes which you can't imagine your Aing will be neither more nor less rational than your Bing. But there is nothing compulsory about this 'cancelling out' assumption.¹²

It may seem that the ease or difficulty of your forming beliefs about the basic rationality of your actions will depend on the actual content of your internally unconditional wants. Suppose that these desiderata include certain states of affairs that actions can produce. So for example you have an internally unconditional desire for people to be happy, for people to have true beliefs, for distributable things to be distributed according to a certain pattern, at least if they are themselves things for which you have internally unconditional desires. With wants like this, you may never be able to believe enough about the future to form beliefs about the basic rationality of your options. There will be too many unimaginable ways in which each action that is open to you could be related to states of affairs of the kind you want. But now suppose by contrast that the only things for whose existence or non-existence you have internally unconditional desires are not states of affairs that actions can produce but rather certain kinds of actions themselves. Won't this make it easier for you to form beliefs about the basic rationality of your options? Your two alternatives are Aing and Bing. Can't you just tell whether an action is basically rational by seeing whether it is an action of one of these kinds? Suppose for example that you have an internally unconditional desire for there to be no acts of deception, and an internally unconditional desire for there to be decisions not to deceive. And suppose you know, as of course you *can* know, that Aing would be an act of deception, and that Bing would be a

decision not to deceive. Do you then need to think about the future at all, in order to form the belief that Bing would be basically rational? The answer seems to be that even if there were nothing for which you had an internally unconditional desire but the non-existence of acts of deception and the existence of decisions not to deceive, still you would have to think about the future. Deciding not to deceive may cause an act of deception. By the deception of one person you may prevent the deception of someone else, or of the same person on another occasion. The difficulty is general: for any type of action X, deciding not to X may cause an Xing, and one Xing may prevent another Xing.

It may now be said that if X is such that necessarily an action is an Xing only if it is an action performed by you, then it is highly unlikely that a decision to X will cause an Xing, or that one Xing will prevent another Xing. By deceiving someone you may prevent an act of deception on somebody else's part. But it isn't likely that by deceiving someone you will prevent some further act of deception on your own part. Nor are you likely, by deciding not to deceive someone, to cause yourself to deceive. But the objection is futile. Even if there is someone who has no internally unconditional desires for the existence or non-existence of anything which is not an action or a decision, there is no one whose internally unconditional desires refer only to decisions or actions of his own.

Belief shortage about the future is then apt to prevent us from forming beliefs about the basic rationality of our options. And initially at least we do want to act rationally in a sense which requires us to form such beliefs. Will reflection destroy this initial desire? A few subversive-seeming considerations do suggest themselves.

Game Theorists tell us that situations are possible in which two agents both act on the rule 'maximize your subjective expected utility' but each would have had more actual utility if both had acted on some other rule. Prisoner's Dilemma is one such situation.

		A	
		Hawk	Dove
B	Hawk	1,1	0,3
	Dove	3,0	2,2

A and B are supposed to know that the utilities of the outcomes of their actions are as given in the figures: so for example each knows that if A hawks and B doves then A's utility will be 3 and B's 0. Suppose that each agent knows that the other will perform one or other of the two actions, and that neither agent has any belief as to the probability that the other agent will perform one of the two actions rather than the other. Then if both agents act on the rule 'maximize your subjective expected utility', both will hawk, with the result that each will have actual utility of 1.¹³ But if both had dove, each would have had actual utility of

2. Are such cases possible, and if they are does their possibility shake your desire to form true beliefs about the basic rationality of your options?

The limitations on your ability to imagine future possibilities which will prevent you from forming such beliefs will also prevent you from forming beliefs merely about the utilities of possible outcomes. So if the numbers in the matrix are taken to represent known utilities of the kind involved in basic rationality, as distinct from say cash payments, then there is no reason to suppose that it represents anything but a merely logically possible situation. Real agents won't know the utilities because they won't even be able to form the necessary beliefs. But even if such situations were actual and common, I do not see that this would shake one's desire to act on true beliefs about the basic rationality of one's options. It might shake a desire just to act on one's *beliefs* about the basic rationality of one's options, to act on the rule 'maximize your *subjective* utility'. But that is another matter.¹⁴

Critics of consequentialist morality often deny what they call the strong doctrine of negative responsibility: 'if I am ever responsible for anything, then I must be just as much responsible for things that I allow or fail to prevent, as I am for things that I myself, in the more everyday restricted sense, bring about'.¹⁵ You are, they insist, especially responsible for your own actions, and less responsible even for those actions of others which you are in a position to allow or prevent. If in consequence of my resignation they give my job to Mary, the basic weighted utility of my own action will be affected by how Mary then does the job. But how she does it is her responsibility, not mine. Is it not possible then that we care less about the basic rationality of our actions than about whether they are rational in a more restricted sense which somehow discounts the utility contribution of what other people do? Perhaps. But if belief shortages about the future prevent you from forming beliefs about the basic rationality of your actions, then they will also prevent you from forming beliefs about their rationality even in this restricted sense. The sceptic about volitionally rational belief could still work with a practical rationality which discounts the utility contribution of alien actions.

A final worry. Aren't we as deliberators often willing to ignore the possible consequences of some of our *own* future decisions, preferring to consider these if and when the time comes to make the decisions themselves. Suppose for example that you are trying to make up your mind about whether to move to Berlin. What happens if you move will partly depend on what, once you have moved, you freely decide to do. Similarly if you stay where you are. In making your decision about whether to move, you may be content to ignore all those possible consequences of moving and of staying which are also possible consequences of those further free decisions which if you made them you would make only after moving or deciding to stay. You may be interested only in what we might call the non-self-mediated rationality of moving or deciding to stay. You will then be willing to ignore some part of what determines the basic rationality of your actions. But once again, the sceptic

would still have something he could use. It is hardly easier to form beliefs about the non-self-mediated rationality of one's actions than it is to form them about their basic rationality.

I have been looking in this chapter for a non-evaluative sense of 'rational action' on which the existence of rational actions is an object of unsatisfiable yet reflectively indestructible desire. If there is such a sense, then maybe the sceptic can use it in order to build sound and negation-proof arguments against volitionally rational belief. It looks as if our shortage of beliefs about the future allows for a sense of 'rational action' which serves the sceptic's purpose. Either it is a sense on which an action is rational if the agent believes truly that it is basically rational. Or it is a sense on which an action is rational if the agent believes truly that it is rational in a sense which is just like the basic sense except that it discounts the utility contributions of alien actions or of the agent's own free decisions.

VIII

ARGUMENTS AGAINST RATIONAL BELIEF

Volitionally interpreted, ‘S is rational in believing that p’ means that an action of S’s is rational which is suitably related to his believing that p. The last chapter suggests that by appealing to our shortage of beliefs about the future the sceptic may be able to build sound and negation-proof arguments against non-evaluatively and volitionally rational belief. How exactly would these arguments go?

Perhaps like this:

- (1) N is rational in believing that P only if he has performed a suitably belief-related and non-evaluatively rational action;
- (2) N cannot perform an action of that kind;

So: (3) N is not rational in believing that P.

If ‘rational belief’ is taken in a volitional sense then (1) will express a conceptual truth. Suppose, additionally, that ‘rational action’ is taken in a sense which requires the agent to believe truly that his action is basically rational. Then belief-shortage considerations can be used to support (2). Here we seem to have sound and sceptical arguments. And, it seems, these arguments will pass the negation test. Reflection would not destroy N’s desire for the rational belief these arguments say he does not have, because we do have a reflectively indestructible desire for actions to be rational, in the sense of ‘rational action’ that the argument employs. Our general conjecture was that no sound sceptical argument passes either the negation or the affirmation test. There seems to be a counterexample.

But actually, there is an ambiguity in our initial formulation. ‘N is rational in believing that P’ could mean ‘N is rational in starting to believe that P’. Or it could mean ‘N is rational in continuing to believe that P’. If N is volitionally rational in starting to believe that P, then the rational action which is suitably related to his believing that P is his getting or allowing himself to start believing that P. If he is volitionally rational in continuing to believe that P, then the rational action which is suitably related to his believing that P is his getting or allowing himself to go on believing that P. Draw this distinction, and the counterexample falls apart.

It is fairly easy to show that so far from enabling us to construct sound, sceptical and negation-proof ‘continuing to believe’ versions of (1)–(3), belief-shortage considerations do not even enable us to construct sound and sceptical arguments of this kind. This we will see in section (1). The ‘starting to believe’ versions are more difficult to dispose of. But as section (2) shows, it can be done.

(1) Continuing to believe

‘Continuing to believe’ versions of (1)–(3) go like this:

(C1) N is rational in continuing to believe that P only if he has rationally got or allowed himself to go on believing that P;

(C2) N cannot rationally get or allow himself to do that;

So: (C3) N is not rational in continuing to believe that P.

I suggest that belief-shortage considerations will not help us to construct sound, sceptical and negation proof (C1)–(C3) arguments because they will not even help us to construct sound and properly sceptical (C1)–(C3) arguments.

We can take (C1) as expressing a conceptual truth: it employs a volitional sense of ‘rational in continuing to believe’. And we may take it to begin with that N rationally gets or allows himself to go on believing that P only if he believes that this action would be basically rational. How, on these assumptions, will N’s general belief shortage about the future affect the truth value of (C2)?

N’s shortage of beliefs about the future will not prevent him from believing that he does at least have the option of getting or allowing himself to go on believing that P. And it will not escape him that if he does have this option then he already believes that P: you cannot go on believing what you do not already believe. Wondering whether to get or allow himself to go on believing that P, N asks himself whether he would then be getting or allowing himself to go on having a true belief. If, as well as believing that he believes that P, he does in fact believe that P, then his answer to this self-addressed question will be that this is indeed what he would be doing. For, as I stressed in chapter I, to believe that p is in part to be disposed to answer ‘Yes, it is’ to the self-addressed question, ‘Is my belief that p true?’¹ Suppose then that N does think that by getting or allowing himself to go on believing that P, he would be getting or allowing himself to go on believing the truth. It seems to me that if in this light he now considers whether it would be basically rational for him to get or allow himself to go on believing that P, then despite his general shortage of beliefs about the future he is unlikely to have much difficulty in believing that this would indeed be basically rational. For it does rather look as if

(HC) For almost any S, if S wonders whether it would be basically rational for him to get or allow himself to go on believing the truth then, other things being equal, he will believe that this would be basically rational.

In chapter I I employed the notion of a consensual proposition. ‘Induction is reliable’ is a consensual proposition, as is the negation of the demonic hypothesis which figured in chapter II. ‘No proposition is both true and false’ is another consensual proposition. A proposition is consensual if and only if (i) almost anyone who entertains it firmly believes it; (ii) almost anyone who entertains (i) firmly believes (i); (iii) almost anyone who entertains (ii) firmly believes (ii); and so on, ad infinitum. Our general belief shortage about the future is no obstacle to the existence of consensual propositions, even when, as in the case of ‘Induction is reliable’, they are partly about the future. That it would be basically rational for me to get or allow myself to go on believing the truth is in some respects like a consensual proposition. Other things being equal, I will believe it when I entertain it. And though it is indeed about the future, my general belief shortage about the future does not prevent its being a proposition that I will believe when I entertain it. (HC) contains a *ceteris paribus* clause. When are other things not equal? You think you have the option of getting yourself to go on believing the truth. You may, however, also have an extremely strong desire to stop thinking about your beliefs and get some sleep. You have been captured by brainwashers, and difficult and dangerous measures will be necessary if you are to preserve your belief in what they think is a heinous falsehood but you take to be a very trivial truth. And so on.

I don’t deny that *some* sound arguments can be constructed of the form (C1)–(C3). That is allowed for by (HC)’s *ceteris paribus* clause. It is also true that when N believes that he has the option of causing himself to continue to believe that P, and accordingly believes that he believes that P, this latter second-order belief could be mistaken. If N mistakenly believes that he believes that P, and asks himself whether by getting or allowing himself to go on believing that P he would be getting or allowing himself to go on having a true belief, then he could answer either way. But even if for these reasons we can construct some sound arguments of the form (C1)–(C3), still we cannot construct sound and properly sceptical arguments of that form. Arguments against rational belief are sceptical only if they do not rely on any premise about N’s purely personal deficiencies or special historical circumstances. But what makes (C1)–(C3) arguments sound prevents them from being sceptical. (HC) contains an ‘other things being equal’ clause. If other things aren’t equal that will be a historical accident. If, improbably, N is mistaken in believing that he believes that P that will be just a personal quirk.

An objection. Suppose N does believe that

- (1) It would be basically rational for him to get or allow himself to go on believing the truth.

And suppose he also believes that by getting or allowing himself to go on believing that P he would be getting or allowing himself to go on believing the truth. It will then be hard for him to avoid believing that it would be basically rational for him to get or allow himself to go on believing that P.

Will he not now be faced with questions about the rationality of believing that (1) itself? How can he rely on his belief that (1) to assure himself that it would be basically rational for him to get or allow himself to go on believing that P, unless he also believes that it would be basically rational for him to get or allow himself to go on believing that (1) itself? But how can he assure himself that it would be basically rational for him to get or allow himself to go on believing that (1) without relying on that selfsame belief? Isn't that viciously circular? I do not think that there is any genuine difficulty here. If N is indeed lucky enough to believe that (1), he will believe that by getting or allowing himself to go on believing that (1) he would be getting or allowing himself to go on believing the truth. And so, since N does believe that (1), he can conclude that it would be basically rational for him to get or allow himself to go on believing it. (1) is general enough to apply to the case in which what is at issue is the rationality of continued belief in itself.

Another objection. I have suggested that, other things being equal, and for any proposition p which you already believe, you are likely when you consider the matter to believe that it would be basically rational to get or allow yourself to go on believing that p. Doesn't this have absurdly conservative consequences? Doesn't it follow that you are unlikely ever to think that anything ought to be subtracted from your existing belief system? That does not follow. It is one thing to judge that you ought to get or allow yourself to go on believing something that you already believe, another to judge that it would be basically rational to do this. To judge that your action is basically rational is not to make a value judgement, but rather to say something about its relation to what you want for its own sake. It is also possible to believe, of any particular proposition which you now believe, that it would be basically rational for you to get or allow yourself to go on believing it, and yet also to believe that your belief system contains many inconsistencies. These you may want to search out, and if you do find them that in itself will make for change. Evidentially unsupported members of patently inconsistent sets will go. (I don't deny that some people believe that the urge to search out such inconsistencies is a neurotic symptom or satanic temptation.)

I have been assuming that in (C1)–(C3) arguments, 'rational action' is so defined that N rationally gets or allows himself to go on believing that P only if he believes that this action would be basically rational. On this assumption belief shortage considerations will not help us to construct sound and sceptical (C1)–(C3) arguments. I wondered at the end of the last chapter if we might not be content enough to act rationally in a sense which requires the agent to believe something less than that his action is basically rational, namely that his action has a rationality just like basic rationality except that no account is taken of the utility contributions of his own free decisions or of alien actions. It is fairly plain that sound and sceptical (C1)–(C3) arguments will be at least as difficult to construct if they employ definitions of 'rational action' in which

basic rationality is replaced by something which is just like it except that account is not taken of the utility contributions of the agent's own free decisions or account is not taken of the utility contributions of alien actions.²

(2) Starting to believe

I now turn to our 'starting to believe' arguments against volitionally rational belief. They go like this:

(S1) N is rational in starting to believe that P only if he has rationally got or allowed himself to start believing that P;

(S2) N cannot rationally do that;

So: (S3) N is not rational in starting to believe that P.

Do belief shortage considerations help us to construct sound and sceptical 'starting to believe' arguments which pass the negation test?

I take (S1) to express a conceptual truth: it employs a volitional sense of 'rational in starting to believe'. And, as before, I take it that N rationally gets or allows himself to start believing that P only if he believes that this action would be basically rational. Can we now say that N's shortage of beliefs about the future will prevent him from believing that his action is thus rational, and so ensure the truth of the (S2) premise? In the corresponding 'continuing to believe' arguments, (S2)'s counterpart was

(C2) N cannot rationally get or allow himself to go on believing that P. And belief shortage considerations made that premise true only at the cost of destroying the philosophical and hence sceptical character of the whole argument. Does the same go for (S2)?

To challenge (C2) I invoked:

(HC) For almost any S, if S wonders whether it would be basically rational for him to get or allow himself to go on believing the truth then, other things being equal, he will believe that this would be basically rational.

The idea was this. Suppose that S does indeed believe that P. Then other things being equal his deliberations about whether to get or allow himself to go on believing it will lead him to believe that were he to do so he would be getting or allowing himself to believe the truth. So given (HC), and despite N's general belief shortage about the future, he is likely, other things being equal, to believe that it would be basically rational for him to get or allow himself to go on believing that P. Given the sense of 'rational action' which is being used in (C2), there is no general belief shortage obstacle to N's rationally causing himself to continue to believe that (C2). To shore up (C2) it was necessary to deny that other things are equal, and that destroyed the sceptical character of the whole argument.

Nothing on these lines works with (S2). It may be true that

(HS) For almost any S, if S wonders whether it would be basically rational for him to get or allow himself to start believing the truth then,

other things being equal, he will believe that this would be basically rational.

But (HS) is less helpful than (HC). We can't use (HS) to reach the conclusion that N is likely, other things being equal, to believe that it would be basically rational for him to get or allow himself to start believing that P. If N wonders whether to get or allow himself to *continue* to believe that P, then, other things being equal, he does in fact already believe that P. To believe that p is in part to be disposed to answer 'Yes, it is' to the self-addressed question, 'Is my belief that p true?'. So, other things being equal, N will if he considers the matter think that by getting or allowing himself to go on believing that P he would be getting or allowing himself to go on believing the truth. At this point, (HC) comes into play. But things are quite different if N is wondering whether to get or allow himself to *start* believing that P. For then, if other things are equal, he does *not* already believe that P. And not believing that P will not in the least incline him to believe that getting or allowing himself to start believing that P is getting or allowing himself to start believing the truth. (HS) can't come into play. We can't use it to obstruct the sceptic who supports (S2) on the grounds of our general belief shortage about the future.

Suppose we grant that, thanks to belief shortage considerations, there are indeed sound and properly sceptical 'starting to believe' arguments against descriptively and volitionally rational belief. Will these arguments pass the negation test? Or is it that if N were confronted with an appropriately interpreted argument of the (S1)–(S3) form, reflection would destroy his desire for the falsity of its conclusion? Only in the latter case can we preserve our general conjecture about sceptical arguments.

Someone may say that however much we may want to form true beliefs about the basic rationality of non-belief-related actions, we are interested in something other than basic rationality when it comes to deciding whether to get or allow ourselves to start believing things. We are interested in epistemic rationality. Is it not easier for N to believe that he would be epistemically rational in causing himself to start believing that P than to believe that this action would be basically rational? And may it not be that by recognizing this N will destroy his frustrated desire to form the belief that the action would be basically rational, and with it his desire for the falsity of the sceptical conclusion (S3)?

We must fix what it is for a belief to be epistemically rational. Most current writers define epistemic rationality not as one might have expected in terms of knowledge, but rather in terms of true belief. Let us go along with this, and take it that S's action is epistemically rational if he has an internally unconditional desire to believe and to continue to believe those propositions which are true, and not to believe or continue to believe those propositions which are false, and if no action which S is able to perform instead makes more for the satisfaction of these wants of S's. I agree that in the face of general belief shortage about the future it may be easier for N to form the belief that it would be epistemically

rational for him to get or allow himself to start believing that P than to form the corresponding belief about the basic rationality of that action. But I do not see why N is supposed to be willing to ignore all his goals apart from those bearing on the epistemic rationality of his action. If he is deliberating about whether to get or allow himself to start believing that P, why should he not want to take into account the truth of other people's beliefs as well as the truth of his own? Why for that matter is he supposed to be willing to ignore his internally unconditional desires for things which have nothing to do with the truth or falsity of anyone's beliefs?³ Why should he have any desire thus to compartmentalize himself? We need a new idea.

Here is one possibility. Think to begin with of how N might actually acquire the desire to believe that it would be basically rational for him to get or allow himself to start believing that P. This seems to be the most likely process. Believing neither that P nor that not-P, N finds that he wants to believe that P. He thinks that one sure way for him to gain this belief is to get or allow himself to start having it. So he wants to take this action. But he also wants not to act in this or any other way without believing that his action would be basically rational. So he wants to believe that it would be basically rational for him to get or allow himself to start believing that P. Given that his want does have this origin, there are various ways in which reflection might undermine it.

One helpfully subversive thought is this. Were he to get or allow himself to start believing that P he would run the risk of ending up with a false belief. There are I admit people who positively want to court such risks, who see an intrinsic value in their acceptance. There will be more about this attitude in the next chapter. But it is not an attitude which we are likely to find in a person who wants to believe that it would be basically rational for him to get or allow himself to start believing that P. I also agree that there are cases in which N thinks that P's truth would be secured by the very fact that he believes it. By starting to believe that he will recover from his illness N will in fact cause his own recovery. Or so he thinks. But in most cases his most natural response will be to wonder whether he cannot somehow get evidence for P's truth. He can investigate, try to discover whether P is true or false. There are in fact three possibilities, (i) He gets evidence strong enough to make him believe that P is true, (ii) He gets evidence strong enough to make him believe that P is false, (iii) He does not get evidence strong enough to have either effect. Let me consider these in turn.

Suppose that (i): he gets evidence strong enough to make him believe that P is true. Then his problem has resolved itself: if he does indeed now believe that P is true, it is no longer possible for him to get or allow himself to start believing it. He wanted to believe that P. Now he does. He has acquired his coveted belief that P without having to get or allow himself to start having it. The belief-generating action in whose basic rationality he wanted to believe is no longer open to him. Maybe his general belief shortage about the future did

indeed make it impossible for him to believe that such action would be basically rational. Now it no longer matters.

Next suppose that (ii): he gets evidence strong enough to make him believe that P is false. Will he still want to get or allow himself to start believing that P? Surely not. Why should he want to get or allow himself to start believing a proposition which he now believes is false? If, however, he no longer wants to get or allow himself to start believing that P, then he will no longer want to believe that such action would be basically rational, and even if his general belief shortage about the future does make it impossible for him to form such a belief, he will no longer care.

The remaining case is (iii): he does not get evidence strong enough either to make him believe that P or to make him believe that not-P. This is the difficult case, and, of course, it is a likelier case than either (i) or (ii). Sitting in his armchair by the fire, N wonders whether there is any coal left in the cellar. He wants to believe that there is. He thinks he can get or allow himself to start believing that there is: it will be enough for him just to dwell on how big a blaze he can make. Since he wants to get or allow himself to believe that some coal is left, he wants to believe that this action would be basically rational. He worries that if, without leaving his armchair, he gets or allows himself to start believing that some coal is left, then he runs the risk of acquiring a false belief. But now he realizes that he has another option. He can go down to the cellar and have a look! Either this will give him evidence strong enough for him to believe that some coal is left, or it will give him evidence strong enough for him to believe that no coal is left. But, often enough, he can't thus easily get evidence either for P or for not-P. Suppose for example that P is a scientific hypothesis which N wants to believe. He thinks he can get or allow himself to believe that P, initially wants to do this, and initially wants to believe that it is basically rational for him to do this. He worries about the risk that this will leave him with a false belief. Can he come up with evidence strong enough to make him believe that P? He believes that there is a proposition Q about the relative complexity of P and its rivals, and about the observation statements which P entails, such that P if and only if Q. And he believes that a research team which investigated the relative complexity of P and its rivals and the truth value of the observation statements which P entails might be able to ensure that they will believe that Q just when Q is true, and believe that not-Q just when Q is false. But N doesn't believe that he can himself do anything to get such evidence. Decades may pass before the hypothesis is tested. It may not be tested before he is dead. Or maybe P is a metaphysical hypothesis which N wants to believe but whose truth value completely eludes investigation.

Suppose then that (iii): N does not get evidence strong enough either to make him believe that P or to make him believe that not-P. He wants to believe that P. So initially at least he has some desire to get or allow himself to start believing that P, and, hence, some desire to believe that this would be basically rational. He does, however, realize that he would then run the risk of

acquiring a false belief. This being so, reflection will I think lead him to dislike the prospect of getting or allowing himself to start believing. And this, however much he wants actually to have the belief. Take the extreme, Pascalian case. There is, N thinks, no good evidence either for or against God's existence, but just a finite probability that God exists. If he exists, belief has infinite benefits. If he does not, belief has only finite disadvantages. Should N get or allow himself to start believing? Pascal says he should.⁴ Many people feel strongly antipathetic to this advice. It is 'absolutely wicked', according to Moore.⁵ Mackie thought that 'from the ascription to God of moral goodness in any sense that we can understand' it follows that he would look with more favour 'on honest doubters or atheists who, in Hume's words, proportioned their belief to the evidence, than on mercenary manipulators of their own understandings'.⁶ Suppose that N agrees. Then even if the case is less extreme, and N thinks that his believing that P would have only finite benefits, reflection will destroy his desire to get or allow himself to start believing. If, however, he does thus lose the desire to get or allow himself to start believing that P, then he will also lose the desire to believe that it would be basically rational for him to do so.

This suggestion about case (iii) would be more plausible if the antipathy to the belief-related action could be seen as something other than a brute fact, could be derived from some more fundamental thought or attitude. In chapter VII mentioned the idea that if you do get or allow yourself to start believing that p, and do this without having evidence for p, you restrict your freedom with respect to what, after the event, you can believe about the origin of your belief that p. Is it fear of this restriction that explains your antipathy? Perhaps. But I came to the conclusion that really there is no such restriction. There is nothing here to make the suggestion more plausible.

Consider this rule:

- (A) Don't get or allow yourself to start believing that p, if no proposition you believe is evidence for p.⁷

Many of us do I think have the vague belief that if one does keep to this rule then whatever the short term consequences somehow all will be well in the end. We may even be prepared to believe that keeping this rule is always basically rational, and that it is never basically rational to get or allow yourself to start believing that p when you do not believe a proposition which is evidence for p. Here we have a large belief about the indefinite future. But, contrary to what the last chapter may have suggested, our belief shortage about the future is not total. We do find ourselves believing certain consensual propositions which are partly about the future, such as that Induction is reliable. And as I said in the last section it is likely that other things being equal you will when you consider the matter believe that it would be basically rational to get or allow yourself to go on believing the truth. I suggest that many of us also believe this about the future: it is never basically rational to break rule (A). If we are for this reason antipathetic to

violations of rule (A), we won't want to perform actions which violate this rule. But if we don't want to perform such actions, we won't after all want to believe that they would be basically rational. To see that this is so, remember the account I gave of how N might acquire a desire to believe that it would be basically rational for him to get or allow himself to start believing that P. He wanted to believe that P, thought that one sure way for him to gain that belief was to get or allow himself to start having it, and wanted not to act in this or any other way without believing that his action would be basically rational. If he loses his desire to get or allow himself to start believing that P, he also loses his desire to believe that such action would be basically rational.

Before I go further, let me recapitulate. Our 'starting to believe' arguments against rational belief go like this:

(S1) N is rational in starting to believe that P only if he has rationally got or allowed himself to start believing that P;

(S2) N cannot rationally do that;

So: (S3) N is not rational in starting to believe that P.

N rationally gets or allows himself to start believing that P only if he believes that this action is basically rational. I am trying to show that even if belief shortage considerations make for sound and properly sceptical 'starting to believe' arguments, no sound and properly sceptical arguments of this kind will pass the negation test. There are three cases to consider: (i) those in which N gets evidence strong enough to make him believe that P, (ii) those in which N gets evidence strong enough to make him believe that p is false, and (iii) those in which N does not get evidence strong enough to have either effect. We have seen that in cases (i) and (ii) reflection will destroy N's desire to believe that he is basically rational in getting himself to start believing that P, and hence his desire for (S3) to be false. So in these cases the (S1)–(S3) argument will not pass the negation test. I suggested that in case (iii) N's desire to believe that he is basically rational in getting or allowing himself to start believing that P, and hence his desire for the falsity of (S3), will be destroyed by his antipathy to any violation of rule

(A) Don't get or allow yourself to start believing that p, if no proposition you believe is evidence for p.

And I suggested that this antipathy can be explained in terms of N's belief that no violation of rule (A) is basically rational.

Someone may now object that in one case at least it is in fact basically rational to violate rule (A). And this is precisely the case which I first used to illustrate the antipathy which (A) enshrines. It is the 'Pascalian' case, where the proposition for which you have no evidence is that God exists. Given that there is a finite probability that God exists, that if he does then belief has infinite benefits, and that if he does not then belief has only finite disadvantages, the expected utility of getting or allowing yourself to believe that God exists is infinitely higher than the expected utility of not doing this. It is therefore basically rational for you to get or allow yourself to start

believing that God exists. If this is right, and if our hostility to Pascal's advice survives our recognition that his argument is sound, then our hostility can't be explained by our belief that no violation of rule (A) is basically rational. To accept the Pascalian argument is to abandon that very belief.

Can the objection be answered by a refutation of the Pascalian argument? It is often said that for all we know God would, if he existed, punish or at least not reward those very people who have been led by such reasoning to get or allow themselves to believe in his existence. This may show that we do not know that the argument is sound. But it does not actually refute the argument. Perhaps then the argument fails because if God did exist he would not leave us in a state of unbelief: each person would sooner or later be caused by God either to believe that God exists or to believe something with a different sense but the same reference. If that were true, the unbeliever would not need to do more than wait. He would not need to take what from his initial position of unbelief he must see as the risk of acquiring a false belief in God's existence.

There is, however, another possible response. We can concede the soundness of the Pascalian reasoning, drop the idea that we have a reflectively indestructible antipathy to all violations of rule (A), and try to show that even on these assumptions belief-shortage considerations do not force us to conclude that in cases of type (iii), where N has no evidence either for or against P, there is a sound and properly sceptical 'starting to believe' argument which will pass the negation test. 'Starting to believe' arguments go like this:

(S1) N is rational in starting to believe that P only if he has rationally got or allowed himself to start believing that P;

(S2) N cannot rationally do that;

So: (S3) N is not rational in starting to believe that P.

I suggest that if the case is of type (iii) then *either* (a) N's general belief shortage about the future will not ensure the truth of (S2), because it will not prevent him from believing that it *would* be basically rational for him to get or allow himself to start believing that P, *or* (b) N will believe enough about the future to believe that it would *not* be basically rational for him to get or allow himself to start believing that P, and that this will prevent him from wanting to take such action, and hence from wanting to believe that such action would be basically rational and hence from wanting (S3) to be false. If the Pascalian reasoning is sound, then (a) will be true when P stands for 'God exists'. For all other values of P, (b) will be true. If (a) is true, the 'starting to believe' argument is unsound. If (b) is true the argument fails the negation test.

(3) Summary

Chapter VII left us with the promise of sound and sceptical arguments against rational belief which would pass the negation test. If the promise had been

fulfilled these arguments would have been counterexamples to our general conjecture that all sound sceptical arguments fail both the negation and the affirmation tests. The promised arguments would employ a volitional sense of rational belief, and would appeal to our belief shortage about the future. In this chapter the relevant arguments against rational belief were divided into 'continuing to believe' and 'starting to believe' kinds. When sound, the former are not properly sceptical. The latter, when sound, will fail the negation test.

I should perhaps emphasize that when these arguments conclude that N is not rational in believing that P it isn't an evaluation which is being negated. The arguments are about beliefs rational in the sense of being suitably related to rational actions; 'rational action', here, is to be so taken that an action is rational only if believed by the agent to be basically rational, where an action is basically rational if, roughly, it is at least as conducive to what the agent himself wants for its own sake as any other action which he is able to perform. I stress this point because it would be easy to gain a false idea of how what I have said in this chapter relates to stock coherentist and foundationalist models of the ideal belief system. The coherentist says that if your belief system is good or as it ought to be then every proposition which you believe is one which you believe on the basis of propositional evidence. How, you might wonder, can that be squared with my claim that whatever you believe reflection would be likely to lead you to the conclusion that it would be rational for you to go on believing it? What about continuing to believe that for which you have no propositional evidence? According to the foundationalist, your belief system will be good or as it ought to be only if *not* every proposition which you believe is one which you believe on the basis of propositional evidence: some propositions are 'properly basic'. Foundationalists differ on which these propositions are. Some think that only those propositions are properly basic which are self-evident or evident to the senses, others think that propositions are properly basic only if they are either incorrigible or self-evident. Plantinga holds that

many kinds of beliefs can be properly basic...for example perceptual beliefs, memory beliefs, beliefs about the mental states of other persons, inductive beliefs...testimonial beliefs...moral beliefs. And...from a theistic perspective it is plausible to follow John Calvin in thinking that belief in God can also be properly basic.⁸

But all foundationalists agree that there are some kinds of propositions which are not properly basic, even if they do not agree on what they are. Does it not follow from my account of the rationality of continuing to believe that any proposition whatsoever is properly basic, just so long as it is a proposition which you do in fact believe? It could seem that what I have said is incompatible with both models.

In fact there is no conflict with either. While coherentism and foundationalism are normative or evaluative doctrines,⁹ there is nothing normative or evaluative about the kind of rationality which I have been discussing in this chapter. I did indeed maintain that whatever *p* stands for and regardless of whether you have propositional evidence for *p* you will probably be volitionally rational in getting or allowing yourself to continue to believe that *p*. But since this is a non-evaluative thesis it is entirely consistent both with the evaluative doctrine that your belief system will be ideal only if you have propositional evidence for *p*, and with the evaluative doctrine that there are limits to what *p* can stand for if your belief system is ideal and you do not have propositional evidence for *p*.

IX

THE AFFIRMATION TEST

'All sound sceptical arguments fail both the negation and the affirmation tests.' If this is right then no sound sceptical argument is more than trivial. For 'P, so Q' is trivial if Q's truth value is a matter of indifference. And Q's truth value is a matter of indifference if 'P, so Q' fails both the negation and the affirmation tests. 'P, so Q' gets through the affirmation test if and only if it meets this condition: were a normally intelligent person to want Q's truth, reflection would leave his desire unscathed. In the negation test, it is a desire for Q's falsity that has to be reflection-proof. In the last eight chapters I have looked for and failed to find a sound sceptical argument which passes the negation test. What then of the affirmation test? Does any sound and sceptical argument get through that?

Often, there are excellent reasons for wanting to be ignorant. You are, let's say, extremely indiscreet. Better perhaps than if you don't get to know too much about Mary's brother. Publicity will only make things worse. But, of course, the sceptic cannot help. The knowledge that you fear to have is not the knowledge which he says you lack. A sound demonic or regressive argument will doubtless show that you know nothing about Mary's brother. Scant consolation if, in some less ambitious reliabilist or evidence-based sense of 'knowledge', you still have ample knowledge about this man. Is there some further and properly sceptical argument to show that you lack even that less ambitious kind of knowledge? It seems not. If you are in fact as ignorant about Mary's brother as you want to be then this will be thanks to special circumstances of the kind that properly sceptical arguments transcend: you didn't listen to what they said, he kept things very quiet, and so forth.

Yet surely there are sound and properly sceptical arguments for ignorance which aren't of either a demonic or regressive kind. There are, for instance, sound evidence-or-acquaintance arguments. Aren't there, among them, some which pass the affirmation test? I looked in chapters IV and V for sound negation-proof evidence-or-acquaintance arguments. Like any argument for ignorance, an evidence-or-acquaintance argument concludes that N does not know that P. There should, I thought, be sound evidence-or-acquaintance arguments in which P was metaphysical. If, as well, this P was non-

consensual, and, further, a proposition whose truth was a reflectively indestructible object of desire, the argument might, I thought, get through the negation test. P's being consensual would rob N of certain stock reasons for wanting to know its truth.¹ But if P was non-consensual, and N wanted it to be true that P then he might well also want to know that P. And if his desire that P was reflectively indestructible the same might go for his desire to know that P. That was the plan. It didn't work. But mightn't a converse strategy nevertheless yield sound evidence-or-acquaintance arguments which pass the affirmation test? We could look for a sound evidence-or-acquaintance argument in which P stands for a metaphysical proposition whose *negation* is non-consensual, and for whose *falsity* N has a reflectively indestructible desire. If N wants that not-P, and if not-P is non-consensual, then it may well be that he wants to know that not-P, and hence wants *not* to know that P. And if his want that not-P is reflectively indestructible, and he does indeed want to know that not-P, then his desire for this knowledge may have the same stability, and, likewise, his desire not to know that P. What might P stand for, in such an argument? Perhaps for some atheistic doctrine, or for a materialism incompatible with immortality, or for some deterministic doctrine incompatible with a freedom which we have a reflectively indestructible desire to possess.

This converse strategy founders on a doubtful inference. The conclusion of an evidence-or-acquaintance argument is that N does not know that P. The argument gets through the affirmation test if N has a reflectively indestructible desire not to know that P, not to have the knowledge which the sceptic argues that he does not have. But why should N have *that* desire? It may well be that if not-P is non-consensual, and if N has a reflectively indestructible desire for P's falsity, then he has a reflectively indestructible desire to know that not-P. But knowing that not-P is not the same as not knowing that P. It is true that the former entails the latter, and that N will hardly be able to avoid believing that this entailment holds. Isn't it true then that if N has a reflectively indestructible desire to know that not-P, and believes that this entails that he does not know that P, then he does also have a reflectively indestructible desire not to know that p? That might be so, given

ENTAILMENT: For any person S and any propositions p and q, if S believes that p entails q, and does not want that q, then he does not want that p.

But, as we saw in chapter III, there are counterexamples to that principle. To exclude them we must add the proviso that S may want that p, even though he believes that p entails q, and does not want that q, provided that the reason why he doesn't want that q is that there is some logically sufficient but not necessary condition for q which he does not want to be fulfilled. So N may want to know that not-P, even though he believes that his knowing that not-P entails his not knowing that P, and does want not to know that P, provided that his reason for wanting not to know that P is that there is a sufficient but not

necessary condition for his not knowing that P which he does not want to be fulfilled. And there may well be such a condition. It is a sufficient but not necessary condition for N not to know that P that he is not justified in believing that P. But if what N wants is precisely that not-P, and that he knows that not-P, then he may well not be particularly interested in merely not being justified in believing that P. His not being justified in believing that P is after all quite compatible with P's being true.

There is however another way in which a sound sceptical argument might fail the affirmation test. The argument might be one in which P stands for a proposition for whose *truth* N has a reflectively indestructible desire, but which he does nevertheless also have a reflectively indestructible desire not to know. Here I come back to the fideist, who wants to be ignorant in order to make room for faith. Maybe the sceptic can show that the fideist is indeed ignorant, or that the evidential situation is as bad as the fideist wants it to be. If reflection leaves the fideist's attitude unscathed, the sceptic's argument gets through the affirmation test. Can the sceptic build the necessary argument? Is the fideist's attitude sufficiently resilient? Let me, in the rest of this chapter, focus on the second of these questions. I doubt that the sceptic can in fact construct the necessary argument. But on that I have nothing to add to what I said in Chapter IV.

Kierkegaard saw probability as the enemy of faith. He went still further: improbability is required for faith. Faith, he thought, demands 'the martyrdom of believing against the understanding'; 'the shrewd and prudent man feels his way with the understanding in the realm of the probable, and finds God where the probabilities are favourable...to believe against the understanding is something different, and to believe with the understanding cannot be done at all'.² 'Faith has in fact two tasks: to take care in every moment to discover the improbable, the paradox; and then to hold it fast with the passion of inwardness.'³ Since God's existence is the object of 'an infinite passionate interest', you cannot want to believe that God exists without being willing to make the greatest possible sacrifice in order to achieve that belief. You must be willing to accept the martyrdom of believing what you take to be improbable, paradoxical, absurd: 'to believe against the understanding is martyrdom; to begin to get the understanding a little on one's side is temptation and regression'.⁴

To believe what you take to be improbable is, perhaps, to believe something of some such form as this: p and there is evidence for p's falsity which is stronger than any evidence for p's truth. I will assume that this is indeed the kind of belief which the fideist wants, and wants the sceptic's help to get. Will reflection destroy a fideistic desire for that kind of belief? If not, then reflection is unlikely to destroy a more moderate fideistic desire for lack of evidence or for ignorance. Someone might say that on this interpretation it is in fact a psychological impossibility to believe what you take to be improbable. That, I assume, is wrong. The feat is possible. We should however press this question.

What is the initial situation of the person who the fideist wants thus to sacrifice himself by believing against the understanding? Is he a person who already believes that God exists, or is he a person who doesn't yet believe?

Suppose the former. His sacrifice would then be, not to start believing what he takes to be improbable, but to go on believing what he takes to be improbable. But why should he see that as a sacrifice? It isn't as if he will think that by getting or allowing himself to go on believing he will courageously run the risk of believing something false. On the contrary. Since, as we are assuming, he does already believe, he will just for that reason be disposed to think that by continuing to believe that this is so he would be continuing to believe the truth. To repeat what by now will be a rather familiar point, S cannot actually believe that p without being disposed to answer 'Yes' to the self-addressed question, 'Is my belief that p true?'.

Suppose then that the person to martyr himself by believing against the understanding is someone who does not yet believe. It is certainly true that if this person does not already believe that God exists, and if he contemplates the possibility of getting or allowing himself to start believing this, and if he thinks that there is stronger evidence for the falsity of theism than any for its truth, then he will think that by taking such action he runs the risk of acquiring a false belief. Will he see this as a sacrifice? Again, the word seems wrong. Sacrifice to whom? Not to God, for, prior to his action of getting or allowing himself to start believing, he does not believe that God exists. But whether or not he sees it as a sacrifice, he may still want to run the risk. From this it may seem to follow that he wants the *opportunity* to run the risk, and will therefore welcome any sceptical argument which shows that his evidential situation gives him this opportunity.

In fact, this does not follow. Two questions must be kept apart: (i) Given that you are an unbeliever and that you think that the sceptic has correctly described your evidential situation, will you think that you should take the risk of getting or allowing yourself to start believing? and (ii) Given that you are an unbeliever and that you think that the sceptic has correctly described your evidential situation, will you want actually to be in that situation? Doubtless, a positive answer to the second question requires a positive answer to the first: unless the unbeliever thinks that, given the nature of his evidential situation, he ought to take the risk, he won't be glad that he is in an evidential situation which gives him this opportunity. But the converse does not hold. It is quite consistent to give a positive answer to the first question, but a negative answer to the second. The believer can think that he should take the risk, without being glad that he finds himself in a situation which makes it possible or necessary for him to take it. An analogy. The boat has capsized; a child is floundering in the sea. You believe that only if you dive in and try to swim over to the upturned boat is there any chance that the child will be saved from drowning. You are an indifferent swimmer; you know that if you dive in you put your own life at risk. Should you dive in? You think you must. You aren't, however, glad to find

THE AFFIRMATION TEST

yourself in a situation in which you have to risk your life, even though it gives you the opportunity courageously to take the risk. To risk your life you see as an unlucky necessity, something required by a situation in which you have no desire to find yourself. Similarly, the unbeliever may think that, given the nature of his evidential situation, it is only by running the risk of acquiring a false belief that he has any chance of gaining the belief that God exists, a belief which he powerfully wants to have. This convinces him that he ought to take the risk. But he may not be in the least glad to find himself in an evidential situation which calls for this response; he doesn't value the poverty of his evidential situation for the opportunity it gives him courageously to take the risk of acquiring a false belief. To take this risk is an unlucky necessity, something required by a situation in which he has no desire to find himself.

What, then, might lead the unbeliever to welcome his actual and, as we are supposing, adverse evidential situation? As I said, he must at least think that, given that his evidential situation is as the sceptic describes it, he ought to run the risk. Perhaps he will indeed think this. He must however also think something else. He must think that there is some *intrinsic* value in running risks of this kind, or want for its own sake to take such risks. It won't be enough for him to suppose that, given his evidential situation, running the risk has more instrumental value than not running it, or that running it is more conducive to his other goals. For that would be compatible with not wanting to be in that evidential situation in the first place. Thinking that, as things are, he ought to run the risk, does nothing to make him glad about the evidential situation, does nothing to make him glad that the sceptic is right. Will the unbeliever have the necessary extra attitude? Will he attach an intrinsic value to running the risk, and hence value the evidential situation for the opportunity it gives him? Some people seek out mortal danger just to brave it. Maybe the unbeliever is inclined to do that. Running the risk of acquiring a false belief might seem to him like risking his own life. He might thus value the adversity of his evidential situation, want the sceptic to be right. I do not see that reflection could undermine that attitude. If the sceptic really could construct the argument, it would pass the affirmation test.

APPENDIX

Semantic objections to demonic arguments

In Chapter II I defended the soundness of demonic arguments for ignorance, and showed that they fail the negation test. I said nothing, however, against certain currently influential semantic objections to demonic arguments. I now append a few critical remarks about these. My general conjecture is that no sound sceptical arguments pass either the negation or the affirmation test. That can stand even if demonic arguments do in the end fall to semantic objections. But even in that case, it would I think be less economical to defend my conjecture by trying to secure semantic objections to demonic arguments than to defend it by failing these arguments in the negation test.

Demonic arguments go like this

- (1) N knows that P entails not-(D);
- (2) For any person S and any propositions p and q, if S knows that p, and knows that p entails q, then he knows that q;
- (3) N does not know that not-(D);

So: (4) N does not know that P.

(D) is the hypothesis that nothing exists but N and a demon who makes him believe whatever it is that in the actual world he does believe, and P is a contingent proposition which entails not-(D). Consider a demonic argument whose conclusion is that N does not know that Mary is looking at a poppy. If the form (1) premise of this argument is true, then N understands the sentence 'Mary is looking at a poppy'. But does he, if the argument is sound?

If N does understand the sentence, then, it seems, he is able to follow a linguistic rule consistently, which is to say, in the same way as on earlier occasions. There must accordingly be some fact in virtue of which a linguistic rule can be consistently followed by N. Wittgenstein has helped to convince many philosophers that this fact can only be a communal practice with which N's individual practice agrees. And this conviction might prompt the following train of thought. There could not be a communal practice governing the use of the sentence 'Mary is looking at a poppy', with which N's practice agreed, unless N had the capacity to know that there are poppies. But that there are poppies is a contingent proposition

which entails not-(D). And if there is a sound demonic argument whose conclusion is that N does not know that Mary is looking at a poppy, then there is also a sound argument of this form:

- (1a) N knows that ‘There are poppies’ entails not-(D);
- (2b) For any person S and any propositions p and q, if S has the capacity to know that p, and knows that p entails q, then he has the capacity to know that q;
- (3a) N does not have the capacity to know that not-(D);

So: (4a) N does not have the capacity to know that there are poppies.

Necessarily (1a) is true only if N understands ‘There are poppies’. But by the rule-following considerations, ‘N understands “There are poppies”’ is incompatible with (4a). So the argument (1a)–(4a) is unsound. So there is no sound demonic argument whose conclusion is that Mary is looking at a poppy. The reasoning can be generalized: there are no sound demonic arguments.

To accept this reasoning, you need to accept that

- (A) The fact in virtue of which an individual can consistently follow a linguistic rule can only be a communal practice.

You need to accept that the fact in virtue of which an individual can consistently follow a linguistic rule is not, for instance, that the individual has previously intended to use certain words when and only when he has thoughts or experiences of certain kinds. Which is easier, to establish (A), or to show that demonic arguments fail the negation test? Has it even been conclusively shown that the fact in virtue of which an individual can consistently follow a linguistic rule is not that the individual has previously intended to use certain words when and only when he has thoughts or experiences of certain kinds? A predictable reply would be that if ‘E’ is the expression on which N has thus bestowed a meaning, still the expressions, other than ‘E’, which figure in his mental meaning-bestowing act would also need to have a meaning. If the meaningfulness of these further expressions is in turn explained in terms of N’s previous mental acts, we have embarked on a regress which we do not know how to terminate. Why, though, should it be assumed that any expressions other than ‘E’ would ‘figure in’ N’s mental act? Certainly it is not obvious that thought is impossible without the employment of a language. One might suppose that the meanings with which the individual would by his own activity endow his words could only be eternally existing and extra-mental concepts which he had somehow ‘grasped’. But there are other ways of looking at the matter. Husserl taught that meanings of words can be seen as species of parts of mental acts. Is it impossible for words to have meanings which are species of parts of the mental acts just of a single individual? It wasn’t easy to show that demonic arguments fail the negation test. But, I suggest, it was easier than it would be properly to establish (A).¹

There are other semantic objections to demonic arguments. To discuss them, let me cast demonic arguments in a slightly different form.² Consider to begin with

- (2) For any person S and any propositions p and q, if S knows that p, and knows that p entails q, then he knows that q.

And consider this definition: For any propositions p and q, q is a *counterpossibility* to p if and only if q is not necessarily false, and p is not necessarily false, and q is incompatible with p. If (2) is true, so clearly is this:

For any person S, and any propositions p and q, if S knows that p, and knows that q is a counterpossibility to p, then S knows that not-q.

Demonic arguments can accordingly be reformulated as follows:

- (1a) N knows that (D) is a counterpossibility to P;
 (2a) For any person S, and any propositions p and q, if S knows that p, and knows that q is a counterpossibility to p, then S knows that not-q;
 (3a) N does not know that not-(D);

So: (4a) N does not know that P.

Replace (D) by

- (V) We are brains in vats,

and let P stand for a contingent proposition incompatible with (V), and you get what can be called envatment arguments for ignorance. These are the arguments which Putnam has criticized, from the standpoint of semantic externalism. He has argued, from this standpoint, that (V) is necessarily false. If this is right then envatment arguments for ignorance are unsound. Since (V) is a counterpossibility to P only if it is *not* necessarily false, it follows from (V)'s necessary falsity that the first premise of an envatment argument is false. One might well think that if, for this reason, envatment arguments are unsound, arguments of form (1a)–(4a) will also be unsound, and so therefore will be all demonic arguments.

(V), for Putnam's purposes, means that nothing exists but brains in vats, together with a computer which generates in its charges streams of sense impressions just like ours. According to Putnam's semantic externalism, if a brain in a vat says 'There is a tree in front of me', that sentence has no meaning unless 'tree' has an extension to which the brain's use of the word 'tree' is causally related. So if the vat-English sentence 'There is a tree in front of me' does have a meaning, then 'tree', in vat-English, has an extension which is not the set of actual trees. Its extension, Putnam supposes, might be 'trees in the image', or in other words the contents of those sense impressions as of trees which can be found in the envatted brain's computer-generated sensory stream. Equally, its extension might be 'the electronic impulses that cause tree experience, or...the features of the program that are responsible for those electronic impulses.' None of these possibilities are according to Putnam ruled out, for 'there is a close causal connection between the use of the word "tree" in vat-English and the presence of trees in the image, the presence of electronic impulses of a certain kind, and the presence of certain features in the machine's program.'³ If we take it that the extension of 'tree' in vat-English is trees in the image then 'There is a tree in front of me' is true in vat-English if and only if I have sense impressions as of trees.

Correspondingly, as Putnam now points out, ‘We are brains in a vat’ is true in vat-English if and only if ‘*we are brains in a vat in the image* or something of that kind (if we mean anything at all).’

How does Putnam now reach his conclusion that ‘We are brains in a vat’ is necessarily false? As follows:

part of the hypothesis that we are brains in a vat is that we aren’t brains in a vat in the image (i.e. what we are ‘hallucinating’ isn’t that we are brains in a vat). So, if we are brains in a vat, then the sentence ‘We are brains in a vat’ says something false (if it says anything). In short, if we are brains in a vat, then ‘We are brains in a vat’ is false. So it is (necessarily) false.⁴

Suppose, in other words, that ‘We are brains in a vat’ is true in vat-English if and only if we are brains in a vat in the image, i.e. if and only if we have sense impressions as of being brains in a vat. Then ‘We are brains in a vat’ is false in vat-English. For it is part of the hypothesis that brains in a vat have streams of sense impressions just like ours. And our sense impressions do *not* include sense impressions as of being a brain in a vat. It is also possible that ‘We are brains in a vat’ is true in vat-English if and only if ‘brains in a vat’ refers to the electronic impulses that cause sense impressions as of being a brain in a vat, or to the features of the program that are responsible for the electronic impulses that cause such experiences. But even on these alternative assumptions ‘We are brains in a vat’ will not be true in vat-English unless brains in a vat have sense impressions as of being brains in vats. And *ex hypothesi* they have no such sense impressions.

Can we show that even if Putnam’s semantic externalist premises were granted his argument would be unsound?⁵ Suppose we can’t. If Putnam’s argument were sound would an analogous argument show that it is necessarily false that

(D) Nothing exists but N and a demon who makes him believe whatever it is that in the actual world he does believe,

and hence that demonic arguments for ignorance are unsound? Suppose it would. Even on these assumptions, there is a long way to go before it can be concluded that demonic arguments for ignorance are unsound. For why should we accept the actual semantic externalist premises? Arguments in their favour have been constructed by Putnam himself and by Burge. But these arguments themselves attract complex objections. Once again, it seems easier just to show that demonic arguments fail the negation test. But let me continue.

Consider the ‘arthritis’ argument of Burge. Alf reports to his doctor that he has arthritis in the thigh. The report is mistaken: though there is indeed something wrong with Alf’s thigh, Alf doesn’t realize that ‘arthritis’ applies only to ailments of the joints. Now suppose that some but not all elements of this situation are different. Suppose that while all N’s physical and non-

intentional states remain the same, and the same thing continues to be wrong with N's thigh, we have a situation in which 'arthritis' is conventionally applied not only to ailments of the joints but also to the sort of thing which is wrong with N's thigh. If, in this altered situation, N reports to his doctor that he has arthritis in the thigh, his report will not be mistaken. From this it is supposed to follow that the contents of our beliefs are not wholly determined by our physical and non-intentional states, but supervene on our linguistic environment. Maybe we can get from here to the semantic premise which Putnam needs, or at any rate to something otherwise incompatible with the soundness of demonic arguments.

There are of course difficulties. In both situations Alf will naturally express what he believes with the sentence 'I have arthritis in my thigh'. But it is plausible to suppose that in both situations what he believes is the same. It is plausible to say that in both situations he believes two things, firstly that there is something wrong with his thigh, and secondly that 'I have arthritis in my thigh' is the right sentence to express that first belief. 'For beliefs to be expressed in words, they have to go via second-order beliefs about which words are the right ones for expressing which beliefs: sentences do not, as it were, just "squirt out" beliefs.'⁶ In both situations, then, the first thing that Alf believes is true. In the initial situation the second thing he believes is false. And in the new situation the second thing he believes is true. The change in his linguistic environment changes the truth value of one of his beliefs, but it doesn't change the content of any of them.⁷

A few remarks now on Putnam's Twin Earth arguments for the externalist premises that his anti-entailment strategy requires. It will be remembered that in Putnam's first Twin Earth story the liquid called 'water' on Twin Earth is not H₂O but a different though superficially indistinguishable liquid whose chemical formula is XYZ. Suppose I have on Twin Earth a Doppelgänger D whose 'narrow' psychological states are and have always been type-identical to mine, where a psychological state is narrow 'if it does not presuppose the existence of any individual other than the subject to whom that state is ascribed'⁸. Suppose I say 'Her shoes let the water in', and D also says 'Her shoes let the water in'. Putnam's argument is then in effect as follows.

- (1) I understand 'water', in the sentence I utter, and D understands 'water', in the sentence he utters;
- (2) 'Water' in the sentence I utter has an extension which is different from the extension of 'water' in the sentence uttered by D;
- (3) Either (a) 'water' in the sentence I utter has a different intension from 'water' in the sentence uttered by D, or (b) 'water' in the sentence I utter has the same intension as 'water' in the sentence uttered by D, or (c) 'water' has no intension either in the sentence I utter or in the sentence uttered by D;
- (4) Not (3a);

- (5) Not (3b) (from (2));
- (6) (3c) (from (3), (4) and (5));
- (7) There are some terms which cannot be understood unless they have extensions (from (1) and (6)).

Why (4)? Why shouldn't 'water' in the sentence I utter have a different intension from 'water' in the sentence uttered by D? Putnam's answer is that this is impossible on the initial assumption that D and I have none but type identical 'narrow' psychological properties. We couldn't each have 'grasped' different intensions, and each have associated 'water' with them, if, at the time, our 'narrow' psychological properties were type-identical. This answer is challenged by Searle, who thinks that Putnam has failed to consider the following possibility.

Let us suppose that Jones on the earth in 1750 indexically identifies and baptizes something as 'water' and twin Jones on twin earth also indexically identifies and baptizes something as 'water'. Let us also suppose that they have type-identical...visual and other sorts of experiences when they make their indexical identification...[But] though they have type-identical visual experiences in the situation where 'water' is for each indexically identified, they do not have type-identical intentional contents...The indexical definitions given by Jones on earth of 'water' can be analysed as follows: 'water' is defined indexically as whatever is identical in structure with the stuff causing *this* visual experience, whatever that structure is. And the analysis for twin Jones on twin earth is: 'water' is defined indexically as whatever is identical in structure with the stuff causing *this* visual experience, whatever that structure is.⁹

Searle is not suggesting that the intension of 'water' is thus indexical, only that Putnam has not shown that it can't be. And Putnam might well be perfectly content to say that his conclusion is true only on the condition that natural kind terms do not have intensions which are indexical in the way that Searle describes. But there is another objection to Putnam's argument, which obliges us to deny that it shows even that much. The objection is that premise (2) is false. Why not say that the extension of 'water' in the sentence I utter has an extension which is the same as the extension of 'water' in the sentence uttered by D, and that in both sentences its extension is H₂O and XYZ. If we discovered that on Twin Earth there was a liquid with the formula XYZ which was superficially indistinguishable from what we call 'water', we would discover that 'not all water has the same microstructure'. The point was made more than twenty years ago by Zemach and Mellor.¹⁰ So far as I can see, it still has not been met.¹¹

In Putnam's second Twin Earth story what we call 'aluminium' is on Twin Earth called 'molybdenum', and what we call 'molybdenum' is on Twin earth called 'aluminium'. As before, I have a Twin Earth Doppelgänger D whose 'narrow' psychological states are and have always been type identical to mine. Suppose I say, 'aluminium is different from molybdenum', and D also says 'aluminium is different from molybdenum'. And suppose that neither of us is able to tell aluminium apart from molybdenum. We now have an argument just like (1)–(7), except that, in the new argument, 'water' is replaced by 'aluminium'. The new argument goes in effect like this:

- (1) I understand 'aluminium' in the sentence I utter, and D understands 'aluminium', in the sentence he utters;
- (2) 'Aluminium' in the sentence I utter has an extension which is different from the extension of 'aluminium' in the sentence uttered by D;
- (3) Either (a) 'aluminium' in the sentence I utter has a different intension from 'aluminium' in the sentence uttered by D, or (b) 'aluminium' in the sentence I utter has the same intension as 'aluminium' in the sentence uttered by D, or (c) 'aluminium' has no intension either in the sentence I utter or in the sentence uttered by D;
- (4) Not (3a);
- (5) Not (3b) (from (2));
- (6) (3c) (from (3), (4) and (5));
- (7) There are some terms which cannot be understood unless they have extensions (from (1) and (6)).

It looks as if this new argument is not open to the objection that 'aluminium' in the sentence that I utter has an extension which is the same as the extension of 'aluminium' in the sentence that D utters. But it still breaks down. Consider the inference of (5) from (2). This relies on the principle that if two terms have different extensions then they cannot have the same intension. But if 'aluminium', in the sentence that I utter, has a different extension from the extension of 'aluminium' in the sentence uttered by D, this is because the two extensions are determined, not by a single intension, which D and I both grasp, but by two different intensions, one grasped by those experts on earth who are better than I am at distinguishing what *we* call 'aluminium' from what *we* call 'molybdenum', the other grasped by those experts on twin earth who are better than D is at distinguishing what *they* call 'aluminium' from what *they* call 'molybdenum'. So if (2) is true, and (5) is so taken that it does indeed follow from (2), then the intensions which (5) says are different are not the ones which (4) says are the same. This equivocation invalidates the inference of (6) from the conjunction of (3), (4) and (5).

NOTES

I ARGUMENTS FOR IGNORANCE AND THE NEGATION TEST

- 1 We need not go into the question of what further restrictions, if any, must be placed on the properties which F and G can stand for if these properties are to be projectible.
- 2 Fumerton 1995:198.
- 3 For a proof, see Lewy 1976:49.
- 4 Hume 1978: xviii.

5 Let UNBELIEF be our name for the principle that you want that p only if you do not believe it. Counterexamples do suggest themselves. We often say things like this: 'I want them to arrive at seven and I believe they will', or 'I want them to approve of Martha, and I believe they do'. But perhaps such remarks should not be taken literally. How do we know in the first case that the speaker isn't heavily emphasizing 'and', with a view to conveying that although he knows you think that he wants them to arrive at seven, he also suspects that you mistakenly think that he doesn't believe that they will arrive then? How do we know that the point of the remark about Martha isn't to get you to realize that although the speaker believes that they approve of her so far, he is anxious for things not to change?

A different objection to UNBELIEF would apply the distinction between dispositional and occurrent attitudes. It does look as if one and the same person can have a dispositional want that p and a simultaneous dispositional belief that p. A purely dispositional version of UNBELIEF seems false. A partly dispositional, partly occurrent version also seems implausible. It is implausible to deny that one and the same person can have a want that p and a simultaneous belief that p when one of the attitudes is dispositional and the other occurrent. This leaves us with the purely occurrent version of UNBELIEF. What is it occurrently to want that p? Perhaps this occurrent wanting is a causal sequence in which either occurrently thinking of p causes you to feel immediate ease, or occurrently thinking of not-p causes you to feel immediate anxiety. If occurrently wanting that p were incompatible with occurrently believing that p, this would have the implausible consequence that occurrently thinking of p would cause you ease, but occurrently believing that p would not.

- 6 Stove 1982:14.
- 7 Miller 1994:54.
- 8 Ayer puts the point as follows:

even if [falsificationism] is the correct account of scientific method it does not eliminate the problem of induction. For what would be the

NOTES

point of testing a hypothesis except to confirm it? Why should a hypothesis which has failed the test be discarded unless this shows it to be unreliable; that is, except on the assumption that having failed once it is likely to fail again? It is true that there would be a contradiction in holding both that a hypothesis had been falsified and that it was universally valid; but there would be no contradiction in holding that a hypothesis which had been falsified was the more likely to hold good in future cases. Falsification might be regarded as a sort of infantile disease which even the healthiest hypotheses could be depended on to catch. Once they had it there would be a smaller chance of their catching it again.

(Ayer 1956:79)

- 9 Wright 1991:88.
- 10 There is a particularly clear exposition of this idea in Papineau 1993: ch. 5.
- 11 Papineau 1993:157.
- 12 van Cleve 1984:558.
- 13 'True opinions are a fine thing and do all sorts of good so long as they stay in their place, but they will not stay long. They run away from a man's mind; so they are not worth much until you tether them by working out the reason' (Plato 1956:97d-98a).
- 14 Craig 1990:18, 25.
- 15 Williams 1978:39.
- 16 Alston 1991:148.
- 17 McIver 1956:29.
- 18 McIver 1956:29-30.
- 19 Wittgenstein 1982: remark no. 427.
- 20 Wittgenstein 1953:190.
- 21 'S knows that p, iff p holds in every possibility uneliminated by S's evidence, except for those possibilities which we are ignoring, and properly ignoring.' For David Lewis, this describes the way we actually use 'knowledge'. So if there is a possibility in which p does not hold, and which is uneliminated by S's evidence, and you are not in fact ignoring this possibility, then you will say the truth if you say that S does not know that p. And 'a possibility *W* is *uneliminated* iff the subject's perceptual experience and memory in *W* exactly match his perceptual experience and memory in actuality' (Lewis 1996:553). This makes sound sceptical arguments fairly easy to construct. But, I assume, they will not be ones which pass the negation test. For why should you want a knowledge that p which you can possess only so long as you do not even think of the fact that p's falsity is consistent with your evidence?

II DEMONIC ARGUMENTS

- 1 The anti-neo-Cartesian phase seems to have begun with Dretske 1970, which criticised the sceptical assumption that knowledge is closed under known entailment. Robert Nozick developed the criticism in his 1981, ch. 3, and an enormous literature has developed out of that. For samples, and some references, see S.Luper-Foy (ed.) 1987 and M.Williams (ed.) 1993. In ch. 1 of his 1981, Hilary Putnam presented a semantic argument for the self-refuting nature of an envatted brain hypothesis. That also provoked intense discussion: see for example Brueckner 1992, Wright 1994. As well as demonic or envatted brain arguments for ignorance, the neo-Cartesian sceptic has arguments for conclusions of the form 'N is not justified in believing that P' which rely on N's

not being justified in believing the negation of a deception hypothesis with which P is taken to be consistent. In demonic arguments, by contrast, P is incompatible with (D). Taking as his point of departure Stroud 1983, Wright constructed an elaborate neo-Cartesian argument of this further kind in his 1991. Here the deception hypothesis is that N is dreaming, and P is any proposition which N can be justified in believing only by perceiving. Wright claims that his Dreaming Argument is not only unsound but also has ‘a good claim to be the distillate of the best sceptical thought in the vicinity’ (Wright 1991:96). It isn’t clear whether the vicinity is supposed to include demonic arguments for ignorance.

- 2 Demonic arguments have been defined as a species of form (1)–(4) arguments. I could have adopted a broader definition, and also counted as demonic those arguments whose form is just like (1)–(4) except that (1) is replaced by

P entails Q,

and (2) is replaced by

For any person S and any propositions p and q, if S knows that p, and if p entails q, then S knows that q,

and in which Q stands for not-(D) and P for any contingent proposition which entails not-(D). For more on these further arguments, see the last note to this chapter.

- 3 How can it be necessarily true that N does not know that not-(D), if that entails that N exists? For the sake of brevity, I am taking propositions of the form ‘N does not know that p’ as equivalent to those of the form ‘It is not the case that N knows that p’.
- 4 Nozick 1981:172–96.
- 5 Condition (iii) has to go, if p is both true and non-contingent, because the antecedent of (iii) would then be necessarily false, and as Nozick points out (1981:186), we do not have a theory of subjunctives which covers such cases.
- 6 Brueckner 1991. Put in my own terms, and with ‘knowledge’ taken in the truth-tracking sense, his argument comes to this. Let P stand for ‘I am sitting in a chair’. And let w be a possible world in which P is true and (a) I believe that I am sitting in a chair, (b) I believe that it is false that

(D1) Nothing exists except me and a demon who makes me believe whatever it is that in w I do believe,

and (c) I believe that P entails not-(D1). Since in w P and (a) are true, I will know in w that P if it is true in w that (d) If I were not sitting in a chair I would not believe that I was sitting in a chair, and that (e) If I were sitting in a chair but under circumstances slightly different from those actually obtaining then I would still believe that I was sitting in a chair. But (d) is true in w because, of those possible worlds in which I am not sitting in a chair, the ones closest to w are ones in which I do not believe that I am sitting in a chair. And (e) is true in w because, of those possible worlds in which I am sitting in a chair, the ones closest to w are ones in which I believe that I am sitting in a chair. So in w I know that P.

Is it true that in w I know that P entails not-(D1)? Since ‘P entails not-(D1)’ is true in all possible worlds it is true in w. It is also true in w that I believe that P entails not-(D1). So I will know in w that P entails not-(D1) if, as we can assume, ‘P entails not-(D1)’ satisfies the fourth condition for truth-tracking knowledge of non-contingent propositions. Will I know in w that not-(D1)? Only if it is true in w that (f) If not-(D1) were false I would not believe not-(D1). If not-(D1) were false (D1) would be true, and if (D1) were true then a demon would make me believe whatever it is that in w I do believe. But one of the things that I do believe in w is that not-(D1). So (f) is false in w. So I do not know in w that not-(D1). And so there is a possible world, namely w, in which I know that P and know that P entails not-(D1) but do not know that not-(D1). So (CP1) is not necessarily true when ‘knowledge’ is taken in the truth-tracking sense.

- 7 Call unmodified truth-tracking knowledge ‘TT knowledge’, and let ‘*TT knowledge’ stand for knowledge of the modified truth-tracking kind. Suppose N *TT knows that P, and that he *TT knows that P entails Q. Does he then *TT know that Q? *TT knowledge is defined as follows: For any S, and any p, S *TT knows that p if and only if (a) he TT knows that p, and (b) there is no proposition q such that (i) S TT knows that p entails q, and (ii) S does not TT know that q. So if N *TT knows that P, and *TT knows that P entails Q, but does not *TT know that Q, this must be either (1) because he does not TT know that Q, or (2) because there is a proposition R such that he TT knows that Q entails R, but does not TT know that R. Since, as we are assuming, N *TT knows that P entails Q, he also TT knows that P entails Q. So if he *TT knows that P then by (b) he TT knows that Q. That rules out (1). (2) is ruled out for the following reason. N TT knows that P entails Q. If (2) is true then there is a proposition R such that he TT knows that Q entails R, but does not TT know that R. But N TT knows that P entails Q. So if (2) is true he TT knows that P entails R. But if (2) is true then he does not TT know that R. And then there is a proposition, namely R, such that he does not TT know that proposition but does TT know that it is entailed by P. But then, by clause (b) of the definition of *TT knowledge, he does not *TT know that P.
- 8 ‘That the sceptic’s concept is “technical” doesn’t by itself render his thesis inconsequential. If, for instance, we use the term “knows” synonymously with “confidently believes”, then, of course, the demon argument couldn’t be invoked to show that I do not know that I am sitting at my desk: the concept doesn’t satisfy (at least) one of the argument’s premises. But if the sceptic can introduce a more stringent concept, “know+”, which does allow the derivation to go through, he won’t automatically be confounded when told the concept isn’t ours. The question would have to be confronted as to the significance of knowledge+ attributions, and the implications of their scarcity. The sceptic could be telling us about a more interesting concept’ (Weintraub 1997:13).
- 9 DeRose 1995; Lewis 1996.
- 10 For some objections, see Schiffer 1996.
- 11 DeRose 1995.
- 12 In note 2 I promised to say something about sceptical arguments for ignorance which go like this:
- (9) P entails not-(D);
- (10) For any person S and any propositions p and q, if S knows that p, and if p entails q, then S knows that q;
- (11) N does not know that not-(D);
- So: (12) N does not know that P.
- Here the sceptic’s difficulty is to find a sense of ‘knowledge’ on which both the form (10) and the form (11) premises are true. Suppose we take ‘knowledge’ so that
- (K1) For any person S and any proposition p, S knows that p only if S believes that p;
- and
- (K2) For any person S and any proposition p, S knows that p only if were p false S would not believe that p.
- Then the form (11) premise will be necessarily true. By (K2), N knows that not-(D) only if were not-(D) false, he would not believe that not-(D). Or in other words, N knows that not-(D) only if were (D) true, he would not believe that not-(D). But suppose that N *does* believe that not-(D). Then were (D) true, the demon would make N believe that not-(D). So N knows that not-(D) only if he does *not* believe that not-(D). But by (K1) he knows that not-(D) only if he does indeed believe that not-(D). So

given (K1) and (K2) he does not know that not-(D). The difficulty is that (10) is incompatible with (K1). It is perfectly possible not to believe a proposition which is entailed by a proposition that you know. So by (10) it is possible to know a proposition without believing it.

The difficulty can be met by replacing (10) with

- (10a) For any person S and any propositions p and q, if S knows that p and believes that q and if p entails q, then S knows that q.

But sound though the resulting argument will be, it fails the negation test. Since arguments of the form (9)–(12) are still valid, when (10) is replaced by (10a), and since they will then have only necessarily true premises, their conclusions will also be necessarily true. Such an argument passes the negation test only if N has a reflectively indestructible desire for the falsity of its conclusion, i.e. a reflectively indestructible desire to know that P. But this desire has an incoherent content if it is necessarily true that N does not know that P. And if N were convinced by the argument he would realise this. We now appeal to

INCOHERENCE: No want survives its owner's belief that it has an incoherent content.

This rapid solution is not available for demonic arguments of form (1)–(4) because they have contingent conclusions. Their conclusions are contingent because they have contingent premises of the form 'N knows that P entails Q'.

III REGRESSIVE ARGUMENTS

- 1 Sosa 1991:150.
- 2 Hume thinks that Reason can't be what propels us in such inferences because we can't derive, either by 'demonstrative' or by 'probable' arguments, the principle that nature is uniform. Demonstrative arguments seem for Hume to be valid arguments from necessarily true premises, and by probable arguments he seems to mean arguments from observed to unobserved instances of empirical properties (see Stove 1973:35–8). This seems to be his thought. If Reason were what propelled us in inferences from past experience to the future then a deductively valid and non-circular argument with true premises could be constructed from past experience to the future, and a further deductively valid and non-question-begging argument with none but true premises could be constructed for any questionable premises in the original argument. Now no sound argument from past experience to the future will be deductively valid unless it includes among its premises the principle that 'the course of nature must continue uniformly the same'. This, however, is a questionable principle: it is not a mere description of the impressions of our senses, and its falsity is not inconceivable. So a further deductively valid and non-circular argument with none but true premises is required for the uniformity principle. And this is impossible to supply. For any deductively valid argument for the uniformity principle which has only true premises will include among its premises the uniformity principle itself, and thus be circular.
- 3 It is less usual for epistemologists to talk of internalist or externalist senses of 'knowledge' or 'justified belief' than for them to talk of internalism and externalism as doctrines about what knowledge or 'epistemically' justified belief essentially is, or about how the concepts of knowledge or of 'epistemically' justified belief are to be correctly analysed. For a good description of the variety of ways in which 'internalism' and 'externalism' have been used to stand for such doctrines, see Fumerton 1995:60–9. As he says, 'internalism' is often used to stand for the idea that the believer's actual or potential 'access' to the conditions that constitute

epistemically justified belief is a necessary condition for his belief's being epistemically justified. On Fumerton's own usage the inferential internalist is someone who believes that in order for me to be justified in believing that P on the basis of E, I must be justified in believing that E makes P probable; the inferential externalist denies this.

4 Fumerton 1995:74–5.

5 Fumerton 1995:75.

6 For a good survey, see Plantinga 1993a: ch. 9.

7 Lewy takes it that p analytically entails q if and only if p strictly implies q and 'q contains no concepts which p does not contain' (Lewy 1976:102). This, he thinks, is equivalent to saying that 'necessarily, anybody who is capable of entertaining p is capable of entertaining q'.

IV EVIDENCE-OR-ACQUAINTANCE ARGUMENTS

1 We don't have to suppose that in our language at least much exists in the way of a special or indispensable evaluative terminology.

Hume says that discretion, caution, enterprise, industry, assiduity, frugality, economy, good-sense, prudence, discernment are endowments 'whose very names force an avowal of their merit'...it is fairly plain that Hume is wrong, if the forcing is supposed to be done by conventions governing English usage. We can easily hear any of these terms, except perhaps 'good sense', negatively...Suppose the lexical items that are derog. and contempt, are excised from the language, as the politically correct would wish. The racist, sexist, fattist or ageist can still convey all she needs by playing the information in neutral words, but to the contempt conveying tune...Suppose, for instance, that the word 'gross' is correctly entered in the dictionary as applying to fat people and derog. The fattist can get by without it, by using the word 'fat' instead, with the right kind of sneery tone. The young sometimes call the old 'wrinklies' or 'crumblied', but before they did that they got by well enough by calling them old in a kind of mock-horror, sepulchral, amazed tone.

(Blackburn 1992:285–90)

In Arabic, the higher the pitch the greater the irony conveyed. Sir Charles Johnston relates that on arrival in Jordan at the time of the Suez crisis to take up the post of British Ambassador he was welcomed in a bat-like squeak.

2 For the alleged impossibility, see for example McDowell 1981, and for a good critique of this position see Blackburn 1998:92–104.

3 Blackburn 1998:78.

4 Blackburn 1998:79.

5 In their 1994, F.Jackson, G.Oppy and M.Smith argue that since there is an analytical tie between asserting something which is true or false and having a belief whose content can be captured by means of the sentence used, it is a mistake to suppose that 'ethical cognitivism' is an immediate consequence of any tenable form of 'minimalism about truth'. 'Ethical cognitivists' hold that ethical sentences, like 'Torture is wrong', have truth conditions. And this implies that there can be beliefs whose contents are given by such sentences. A Humean might deny this on the ground that any such belief would have 'a conceptual connection to motivation that no belief can have'. Some minimalists

about truth would say that the sentence ‘Torture is wrong’ has a truth-condition simply because it can be substituted in the schema “‘p’ is true iff p’. For Jackson et al., that neglects the analytical tie between assertion and belief. The most we can say is that *if* a sentence has a truth-condition then one way to give it is to write down the sentence itself. Before we can conclude that ethical sentences do have truth conditions, we must overcome the Humean objection.

For Jackson et al., ethical cognitivism is a descriptive thesis about what we normally do with ethical sentences, the potential upshot of a piece of conceptual analysis. My sceptic does not tie himself down to the description of our ordinary concepts. He is simply trying to find or devise an intelligible sense of ‘evaluation’ on which we never assert propositions when we evaluate particular objects, a sense which he can use in order to construct an argument for ignorance. In response to the objection that the truth-schema prevents him from finding any such intelligible sense, he need not appeal to the existence of a conventional tie between assertion and belief, but can simply stipulate such a tie.

- 6 According to Horwich, ‘expressivism may and...should be located, not in the thesis that evaluations don’t express belief and can’t be true but rather in the thesis that evaluative utterances express desires’ (Horwich 1994:20). Expressivists ‘should not, strictly speaking, deny that one can believe something to be rational’ for what they really provide is an account of ‘what any such belief consists in (that is, the possession of a certain pro-attitude)’ (Horwich 1993:77).
- 7 Support for the thesis that evaluative argument would still be possible can be extracted from Blackburn’s ‘quasi-realist’ treatment of values. See for example his 1993: chs 9 and 10. For cogent criticisms, see Hale 1993.
- 8 ‘Internalism’ and ‘Externalism’ are names for ‘theories’ about ‘the’ concept of evaluation: the internalist says that there is a conceptual connection between evaluating and willing, wanting or being motivated; the externalist denies this. On the unhelpfulness of these labels see Smith 1994b, 60–3. Would ‘purely conative’ be better than ‘purely dynamic’? No, because ‘conative’ comes from the Latin for ‘to endeavour or make an effort’, and desiring is effortless.
- 9 Lewis 1989. Lewis thinks that when the evaluator talks of what we would desire to desire he does not necessarily take ‘we’ to mean ‘everyone’. When I evaluate, I claim ‘as much as I can get away with’: ‘we’ could be ‘all mankind’, or ‘anyway all nowadays’, or ‘anyway all nowadays except maybe some peculiar people on distant islands; or anyway...’ (Lewis 1989:129).
- 10 Jackson 1998:159.
- 11 The doubly attitudinal possibility seems to be somewhat neglected in current philosophy of value. Blackburn maintains that if we analyse evaluation in terms of ‘*descriptions* of things and their powers to elicit responses from us or from some group’ then we can’t possibly do justice to the fact that

evaluative judgements are *verdicts* rather than hypotheses about the suspected reactions of some group (even our own group) under some putative circumstances. If I am asked whether the picture was beautiful or the play interesting, I dissemble if I say that it was *because* I hypothesize that most people in such-and-such circumstances find it so, although I personally couldn’t stand it. What is expected is that I give my own verdict. ‘Yes, it was fascinating’ expresses how I found it.

(Blackburn 1998:109–10)

Why can’t value judgements have a dual function, serve not only to give the evaluator’s verdict, but also to assert a proposition about the reactions of some group?

Michael Smith invites us to consider these three propositions:

- (1) Moral judgements of the form 'It is right that I O' express a subject's beliefs about what it is right for her to do, a fact about what it is right for her to do.
- (2) If someone judges that it is right that she Os then, *ceteris paribus*, she is motivated to O.
- (3) An agent is motivated to act in a certain way just in case she has an appropriate desire and a means-end belief, where belief and desire are, in Hume's terms, distinct existences.

He thinks that although each of these propositions is plausible by itself, the three of them seem to form an inconsistent set. Smith calls this 'The Moral Problem' to mark what he takes to be its central organizing role in 'contemporary meta-ethics' (Smith 1994b:11–12). If there is a problem here then one solution would be to show that really there is no inconsistency after all. This is what Smith tries to do. He thinks that (1) is true because when I make a moral judgement of the form 'It is right that I O' I express my belief that I have a 'normative reason' to O, and this is a belief about an objective matter of fact. But it is also a belief which 'can be represented' as a belief that I would desire to O if I were fully rational (Smith 1994b:184–5). And if I believe that I would desire to O if I were fully rational then either I want to O or I am not fully rational. But the practicality requirement on moral judgements which is given by (2) is in fact nothing more than that if an agent judges that an act is right then either she is motivated to act accordingly or she is irrational (Smith 1994b:193). So there is no inconsistency between (1) and (2) even on the supposition that, as (3) says, motivation is to be explained in Humean terms.

An objection to this solution is that no one who thinks that there is a problem about the consistency of (1)–(3) is likely to concede that there is such a thing as a purely factual belief that I would desire to O if I were fully rational. Smith admits that he is unable to analyse 'fully rational' without using normative concepts (Smith 1994b:162). But if the judgement that I would desire to O if I were fully rational is itself normative, and if there is a problem about how normative judgements can express factual beliefs, then there is also a problem about how there can be a factual belief to the effect that I would desire to O if I were fully rational. Smith's solution depends on an account of (1) which is itself acceptable only if we already have a solution. A simpler solution would be just to say there is no real inconsistency because one and the same judgement can be doubly attitudinal, at once the expression of a desire and the assertion of a proposition. Smith seems to think that this requires one to say that 'there are beliefs which *are* desires', the line taken by 'anti-Humean theorists of motivation like McDowell' (Smith 1994a:9). But why? I should have thought that the possibility of doubly attitudinal judgements is obvious enough for there not even to be an appearance of inconsistency.

12 Mackie 1977:40.

13 For an exposition of this mystery, first propounded by Casimir Lewy, see Blackburn 1985.

14 This can be seen as a secularised version of a thesis popular among religious apologists about the instability of a purely secular ethics:

Evolutionary theory...suggests that we have the deep-seated intuitions, preferences, inhibitions that we do because they used to serve an evolutionary gain (maybe still do)...Theists, by contrast, held that (despite occasional per-versities) our conscience must be accepted as a message from the Lord, that Supreme Value is also Supreme Cause. There are relevant 'moral facts' because those facts are enough, in

context, to cause our belief in them. It hardly matters whether this causal power works through a special sense (or intuition) or through our inbuilt preferences and inhibitions. Discovering, by contrast, that we are caused to have those preferences and inhibitions by something that is not itself a fit object of admiration and respect, must be to weaken our attachment to them. Nihilism beckons...

(Clark 1994:93)

Cf. O'Hear 1993:513–14:

If prior causal conditioning extraneous to the actual rightness or wrongness of the demands were all there was to it, once we realised that most, if not all, of our moral and aesthetic attitudes were given to us by instinct or upbringing, we would begin to think ourselves as less unconditionally obliged by them than in fact we do.

- 15 For more on this point see chapter VI, section (1).
 16 See for example Kenny 1979: ch. 4.
 17 Blumenfeld 1978.
 18 The comparison is Alston's. He continues: 'Having only the sketchiest grasp of chess, I fail to see any reason for Karpov to have made the move he did at a certain point in the game. Does that entitle me to conclude that he had no good reason for making that move?' (Alston 1996:317).
 19 Plantinga 1996a:73.
 20 In the current terminology, '(i) and (ii), so (iii)' is an evidential argument from evil. I suspect that similar objections apply to Paul Draper's attempt to show that theism is improbable because 'a serious hypothesis that is inconsistent with theism explains some significant set of facts about good and evil much better than theism does.'
 Draper 1996:178; cf. Draper 1989. But Draper's attempt has in any case been effectively criticised on other grounds by Plantinga in his 1996b.
 21 'The Principle of Sufficient Reason' is of course a label which has been attached to various doctrines. Some of them are perhaps demonstrably false. A possible instance is the doctrine that every state of affairs has an explanation. For some plausible arguments against that doctrine, see Ross 1969:303–4; Rowe 1975:99–107.
 22 Van Fraassen 1985:255. On empirical adequacy, see van Fraassen 1980: chs 1–4.
 23 Swinburne 1979:97.
 24 Swinburne 1979:101.
 25 By Bayes's Theorem $P(h/e.k) = \frac{P(e/h.k) \times P(h/k)}{P(e/k)}$

If *h* is the theistic hypothesis, *e* the existence of an orderly physical universe, and *k* background knowledge, then according to Swinburne $P(h/e.k)$ is well away from 0. This because although $P(e/h.k)$ and $P(h/k)$ are both quite low, $P(e/k)$ is very low indeed. $P(e/k)$ is very low indeed because it is equal to $\{P(e/h.k) \times P(h/k) + P(e.\text{not}h/k)\}$, and because $P(e.\text{not}h/k)$ is very low. See Swinburne 1979:281–9.

- 26 Swinburne 1979:291.
 27 Swinburne 1979:254.
 28 Swinburne 1979:272.
 29 Leibniz 1973:199.

- 30 A final possibility. Maybe the sceptic can construct a sound evidence-or-acquaintance argument which passes the negation test and in which P, the proposition that N does not know, stands not for 'God exists' but rather for 'N knows that God exists'. Perhaps he could begin by claiming that N has evidence-or-acquaintance knowledge that God exists only if some explanatory exclusion argument is sound. He could then go on to argue that N does not have evidence-or-acquaintance knowledge that any such argument is sound. Just as there is no evidence for the falsity of the Principle of Sufficient Reason so there is no evidence for its truth. Even if sound, this higher-order argument for ignorance seems doubtfully negation-proof. Would reflection leave N's desire to know that he knows unscathed? It is one thing to want to know something, another to want to know that you know it. What purpose would be served by N's knowing that he knows? Consider the two stock reasons for wanting knowledge which I mentioned in chapter I: the Platonic 'running away' reason, and the reason which refers to the grasp which others have of the truth which you yourself believe. N may hope that he will not lose his true first-order belief that God exists. But why should he mind if this happens to his true second-order belief that he knows that God exists. N may want others to share his true first-order belief. But why should he mind if others do not share his true belief about his own knowledge?

V ANTI-IDEALISM

- 1 There is a curious reluctance among many contemporary analytical philosophers to accept the straightforward distinction between intrinsic and dispositional senses even of colour terms. According to Crispin Wright the truth-conditions for 'x is red' are to be 'construed' in terms of the 'provised biconditional' x is red iff 'for any S : if S were perceptually normal and x were presented to S under perceptually normal conditions, then (S would experience x as red if and only if x was red)' (Wright 1988:14 n.26). Colin McGinn has now abandoned the thesis, defended in *The Subjective View* (1983), that for an object to have a colour property is just for it to have a disposition to cause experiences as of an object having that property in normal perceivers in normal conditions. He now accepts the obvious objection that we don't in fact have experiences as of objects having dispositions. His new idea is that colours are 'categorical, simple, monadic features of external objects' which supervene on dispositions to cause experiences (McGinn 1996:545). So far as I can see, this makes it impossible for the internal objects of experiences to be coloured. Peacocke's distinction between red' and red is in some respects similar to the distinction between intrinsic and dispositional redness, 'x is red' he takes to mean 'x is disposed in normal circumstances to cause the region of the visual field in which it is presented to be red' in normal humans', where red' stands for a 'sensational property' of the visual field. Sensational properties are those which an experience has in virtue of some aspect—other than its representational content—of what it is like to have that experience. But Peacocke complicates matters by building it into the definition of red' that this experience does not represent external objects to be red' (Peacocke 1983:39). On this view it is impossible for external objects both to be red', and to be represented as being red'. One philosopher who does straightforwardly draw the essential distinction between the two senses of colour terms is C.D.Broad.

I should say of a pillar-box that it is 'red'. If I came to distinguish between the pillar-box and the visual sensible which I sense when I look at a pillar-box, I should say that the sensible is 'red' too. As applied to the pillar box the word 'red' is a dispositional adjective; as applied to the

visual sensible it is non-dispositional. By saying that the pillar box is red I mean *at least* that, if any normal observer were to look at it in daylight, it would look red to him. And I *might* mean no more than this. By saying that the visual sensible is red I mean something which could not possibly be expressed by a conditional sentence. A man *may* believe that the pillar-box is red in the non-dispositional sense also.

(Broad 1933:148)

- 2 Consider the internal objects of the experiences of infinitely powerful minds. These internal objects can exist without being thought of or experienced by finitely powerful minds. So they are alpha beings, on my definition. Are they contingent alpha beings? If so, our idealist must deny that they have non-mental properties. But why should he want to do that? Those who think that there might be contingent internal objects of this kind can adjust my formula accordingly.
- 3 Suppose we define the physical world in terms of a modal property: it is a world which could not exist if it were not governed by laws of the type which would be postulated by a physical science employing standard principles of non-deductive inference. It may then be argued that since any world of contingent alpha beings lacks this modal property no world of contingent alpha beings is a physical world. For an attempt to defend Berkeleyan idealism with the help of an argument of this kind see Robinson 1994:222–8. In his *Case for Idealism* Foster maintains that ultimate contingent reality is wholly non-physical on the grounds that it lacks a modal property, to do with nomological organization, which is possessed by the physical world, defined roughly as above (Foster 1982:Part III).
- 4 That ultimate contingent reality is wholly mental and that ultimate contingent reality is wholly non-physical, are both by Foster called kinds of Idealism (Foster 1982:3). For Foster, ultimate contingent reality is the totality of ultimate contingent entities and ultimate contingent facts. An entity *x* is ultimate if there is no fact or set of facts which logically sustains the fact that *x* exists. And a fact or set of facts *F* is logically sustained by a fact or set of facts *F'* iff (a) 'it is logically necessary that if *F'* obtains *F* obtains', (b) 'the obtaining of *F* is achieved through and by means of the obtaining of *F'*', and (c) 'the obtaining of *F* is wholly constituted by and is nothing over and above the obtaining of *F'*'. So for example the fact that John is more than 2 stone heavier than Mary is logically sustained by the fact that John weighs 14 stone and Mary weighs 10 stone. And the set {John, Mary} is not an ultimate entity because the fact that it exists is logically sustained by the facts that John exists and that Mary exists (Foster 1982:5–6). It is neither necessary nor sufficient for an entity to be a contingent alpha being that it is an ultimate contingent entity in Foster's sense. The internal object of an experience appears to be an ultimate contingent entity in Foster's sense, since the obtaining of the fact that it exists does not seem to be through and by means of any fact by the obtaining of which it is wholly constituted.
- 5 McDowell 1983:4.
- 6 See Nathan 1991.
- 7 Parts of the next three paragraphs are taken from my 1992:140–2.
- 8 Jackson 1977:123.
- 9 Mackie 1976:18–20.
- 10 Russell 1912:13–14.
- 11 Coraman 1975:241, 338.
- 12 Eddington 1935:5.
- 13 Sellars 1963:26.
- 14 Foster (1982):110–11.

- 15 Foster (1982):179–80.
 16 For a suggestion as to what these properties might be, see my 1997.
 17 Robinson 1994: ch. 6 (an expanded version of some of his 1985); Valberg 1992:106, 111. Neither author is occupied with quite my problem. Robinson wants a sense-datum theory of perception, from which he can negotiate a further passage to the ultimate haven of Berkeleian idealism. Valberg's concern is to reveal a tension between our commonsensical commitment to a causal picture of the world and what he thinks is our inability not to believe in the immediate presence of external objects, given that we have *any* picture of the world.
 18 qValberg 1992:15.
 19 Valberg 1992:16.
 20 Valberg 1992:16.
 21 Valberg 1992:17.
 22 What is the relation between my reformulation of Valberg's argument, and what Robinson says? According to Robinson, it is

clearly true that

1 It is theoretically possible by activating some brain process which is involved in a particular type of perception to cause a hallucination which exactly resembles that perception in its subjective character.

2 It is necessary to give the same account of both hallucinating and perceptual experience when they have the same neural cause. Thus, it is not, for example, plausible to say that the hallucinatory experience involves a mental image or sense-datum, but that the perception does not, if the two have the same proximate—that is, neural cause. These two propositions together entail that the perceptual processes in the brain produce some object of awareness which cannot be identified with any feature of the external world—that is, they produce a sensedatum.

(Robinson 1994:151–2)

He thinks that 'the general principle lying behind (2)...could be expressed by the slogan "same proximate cause, same immediate effect"' (154).

The reasoning in Robinson's (1) and (2) seems to be as follows:

- (A) For every veridical experience as of an external object *o* there is some neural event which is a sufficient cause of that experience;
- (B) Given any neural event which is a sufficient cause of a veridical experience as of an *o*, it is theoretically possible for a neural event of that type to be a sufficient cause of a hallucinatory experience as of an *o*;
- (C) If particular event *x* is a sufficient cause of particular event *y*, then *x* and *y* belong to maximally specific types *X* and *Y* such that whenever there is an *X* type event there is a *Y* type event;
- (D) If neural event *n* is a sufficient cause of an experience which has external object *o* as its content, then this experience does not belong to any maximally specific type which also covers experiences with sense data as their contents;
- (E) Hallucinations have sense data as their contents;
- (F) Veridical experiences have sense data and not external objects as their contents (from (A)–(E)).

The slogan 'same cause same effect' could I suppose be used to describe (C). But Robinson thinks that the slogan describes a contingent principle: 'one is under no logical compulsion to accept [it], if by compulsion one means logical necessity' (157). And (C), if true at all, seems rather to be a conceptual truth about causation. As I develop and criticize my reformulation of Valberg's argument, points will emerge which threaten other premises in (A)–(E).

- 23 Robinson makes a similar objection in his 1994:155–6.
 24 I derive these questions from Robinson 1994:157–8.
 25 The theist has another alternative. He can accept (E) and rise to the explanatory challenge. Originally, all our veridical experiences had alpha beings as their actual contents. God wanted all our experience to be veridical, and it was simpler thus to ensure its veridicality than to arrange for the contents of our experiences to be internal objects which mirror real alpha beings. Non-veridical experience began with the Fall: b-like brain events which had previously produced only experiences with alpha beings as their contents began sometimes to produce non-veridical experiences whose contents were mere internal objects. But at least veridical experience did not altogether cease, and such veridical experiences as we do continue to enjoy have, as before, alpha beings as their contents, and not internal objects.
 26 Broad 1962:3–4.
 27 Mundle 1971:134.
 28 See chapter III, section (2.2).
 29 Sellars 1963: ch. 1.
 30 Cf. Mundle, writing of his rejection of 'Realism concerning colours':

I find this hard to swallow, and wish it were avoidable. It is hard to believe when one is using one's eyes and not merely thinking about their use. I have a well nigh irresistible urge to ascribe colours to physical things. So I end, like Hume, with a conflict between reason and one of my senses.

To forget the conflict he does not need to play backgammon or remind himself about English usage. 'All I need do is look at things and drink in the colours which I cannot but see as *theirs*' (Mundle 1971:179–80).

- 31 Scientific realism is tenable only given some such principle of inference to the best explanation as

(F*) If q is a set of experiential data then q is good evidence for p if p is entailed by the best explanation of q.

And we may insist that all such principles themselves depend on the arbitrary assumption that all experiential data have some explanation or other. This is one of van Fraassen's objections to these principles, and he thinks that rather than make the assumption, and adopt in consequence a realist interpretation of science, we should accept scientific theories only in the sense of believing that they are empirically adequate, that they tell the truth about observable things and events.

Van Fraassen is in fact acutely conscious of our initial desire for the world to be as we believe it to be when we naively admire shapes and colours, and he suspects that part of the reason why scientific realists are baffled by his own philosophy of science is that they have

a different appreciation of just how unimaginably different is the world we may faintly discern in the models science gives us from the world we experientially live in (the scientific image from the manifest image, the intentional correlate of the scientific orientation from the

phenomenological life world). This difference has been stretched by empiricists from the beginning—once atoms had no colour; now they also have no shape, place or volume. (Except that is on certain hidden variable interpretations, which are in my view at best metaphysical baggage but which in any case engender paradoxes of their own).

(van Fraassen 1985:258)

- 32 Sprigge 1984:455. Sprigge remarks (456) that this is ‘very much the theme of Heidegger’s later philosophy, in which he speaks of *Dasein* as the shepherd of being, as the clearing in the forest into which things must come out of the surrounding darkness in order to be themselves’.

VI ARGUMENTS AGAINST JUSTIFIED BELIEF

- 1 This condition was explained at the beginning of Chapter I.
- 2 Alston 1993:532. For a less tolerant approach, see Plantinga 1993a and 1993b. Warrant is Plantinga’s name for whatever ‘distinguishes knowledge from mere true belief (1993b: v). He thinks that ‘twentieth century received tradition in Anglo-American epistemology—a tradition going back at least to Locke—sees justification as essentially deontological but also as necessary and nearly sufficient (sufficient up to Gettier) for warrant’. But the tradition is mistaken: ‘justification properly so-called—deontological justification—is not so much as necessary for warrant. Justification is a fine thing, a valuable state of affairs—intrinsically as well as extrinsically; but it is neither necessary nor sufficient for warrant’ (1993a: 45). I think it is probably too late to insist on the impropriety of non-deontological uses of ‘justified belief’.
- 3 Moser 1989:42.
- 4 Moser 1989:43.
- 5 Fumerton 1995:19–20.
- 6 Bonjour takes it that knowledge is distinguished from true belief by something normative or evaluative, and takes ‘justification’ to stand for this (1985:5–8). Zagzebski defines knowledge as ‘a state of true belief arising out of acts of intellectual virtue’ (1996:271), and on this basis gives yet another job to ‘justified belief. She sees the difference between knowledge and justified belief as analogous to the difference between act and rule utilitarianism. In rule utilitarianism an act is right because it follows a rule the following of which tends to have good consequences. Similarly, a belief may be justified because it is ‘a belief that an intellectually virtuous person might have in the circumstances’ (268). In act utilitarianism an act is right because *that* act leads to good consequences. And when you know, intellectual virtue does actually yield a particular true belief. There is, however, ‘a component of knowledge that is something like justification’ (268).
- 7 Newman: 157.
- 8 Newman: 189.
- 9 See Cohen 1992: ch. 1.
- 10 Price 1954:20.
- 11 ‘On the whole, we believe willingly because we want some good to be gained by believing...where our wanting goods to be gained is part of a cognitive system which has a targeted finality (the survival-then-fulfilment of the active person, in a basically hospitable environment). So “believing what you want to” is not the pathology of religion and madness but the engine of adapted cognition’ (Ross 1993:237).
- 12 Williams 1973b: 148.

- 13 ‘Augustine and Aquinas argue [that] belief caused by the will, going beyond “the weight of evidence”, may and regularly does, yield *knowledge*. That’s because the will, an ability of the rational appetite, is appropriately *biased* towards the good of the person, where “appropriate bias” requires a successful adaptation to reality. In a word, it is not true that beliefs that arise from the will, an “internal cause”, arise irrespectively of whether they are true or not. They are not, therefore, typically illusory’ (Ross 1986:221–2).
- 14 Isn’t at least this much true: there are no cases in which S believes that p, believes that he came to believe that p without believing any proposition which he believed to be sufficient evidence for p, believes that he does not now believe any proposition which is sufficient evidence for p, and believes that he was not and still isn’t caused to believe that p by what makes p true? Surely there are no cases in which you believe that p, and believe that nothing at all ever did or does now incline you to believe that p other than that the proposition was drummed into you by your Uncle? I don’t see why there shouldn’t be. Barbara Winters thinks that necessarily there is no case in which N believes in full consciousness that he believes that p and that his belief that p is not sustained by any truth considerations (Winters 1979:245). As she explains, a belief can be sustained by considerations other than those which originally brought it into being; S may believe that he holds a belief originally acquired independently of truth considerations, if he thinks that he now believes it for different, truth-related reasons. Suppose it is after all a conceptual truth that if S believes in full consciousness that he believes that p and that his belief that p is not sustained by any truth considerations, then he does not believe that p. That won’t show that it is impossible for him to believe in full consciousness both that he believes that p and that his belief that p is not sustained by any truth considerations. For he might just be wrong about what the concept of belief requires.
- 15 There is more about this in section (2) of chapter VIII.
- 16 According to Alston (1989:142), ‘N is justified in believing that p’ can mean that N is not ‘intellectually to blame’ for believing that p. Here again no universal conditional will be asserted when it is said that a belief is justified, but at most the negation of the conditional that you would have asserted if you had said that N *was* intellectually to blame.

Alston holds that for N to be free of intellectual blame for believing that p it is necessary not just that if he intentionally produced his belief that p then he was not intellectually to blame for doing so, but also that his believing that p was not the unintended product of any other actions for which he was intellectually to blame. Suppose that N did not intentionally produce his belief that p. He would still be intellectually to blame for believing that p if his belief resulted from a general habit of credulity formed by actions for which he was intellectually to blame. Alston offers this analogy: ‘although I did not do anything with the intention of bringing about my cholesterol buildup, still I could have prevented it if I had done certain things I could and should have done, for instance, reduce fat intake’ (1989:137). The formula is

S is justified in believing that p *iff* it is not the case that if S had fulfilled all her intellectual obligations, then S’s belief-forming habits would have changed, or S’s access to relevant adverse considerations would have changed, in such a way that S would not have believed that p (143).

Alston omits from the formula the condition that S was not intellectually to blame for intentionally producing his belief that p because he thinks that there is negligible

- scope for intentional production of specific beliefs. As his reference to intellectual blame and intellectual obligations makes clear, he is specifically concerned with epistemically justified belief. We could, however, drop the 'epistemic' qualification and develop a similar 'no fault' interpretation of all evaluatively justified belief. We could say that for N's belief that p to be evaluated as justified is for the judgement to be made that he is not morally to blame for believing that p, where it is sufficient for him not to be morally blameworthy that he performed no morally impermissible action of intentionally producing his belief that p, and no morally impermissible actions of which his believing that p is the unintended product. If 'justified belief is interpreted in this way, the sceptical argument from human diversity cannot get off the ground.
- 17 If on reflection you would be content with a less than universalizing evaluation of one thing as better than another, you would likewise be content with a less than universalizing evaluation of doing something as no worse than doing it, and likewise with a less than universalizing evaluation of doing something as permissible.
- 18 Barry 1995:3–4.
- 19 As Mackie put it, 'the lack of objective values is not a good reason for abandoning subjective concern or for ceasing to want anything' (Mackie 1977:34).

VII FROM RATIONAL ACTION TO RATIONAL BELIEF

- 1 According to Gibbard, 'to call a thing rational is to endorse it in some way' (Gibbard 1990:6). More particularly, it is to 'express one's acceptance of a norm that permits it'. According to Brandt, '*the proper*' analysis or definition of 'it would be rational for X to A' makes it equivalent, on a first approximation, to 'I hereby recommend that X do A...while taking as my sole objective maximizing X's goal-realization (utility-maximization), and as having his information' (Brandt 1990:399, 401). Williams (1979) seems to allow both for evaluative and for purely descriptive uses.
- 2 Wants can also be externally conditional or externally unconditional. Your want that p is externally unconditional if you want to want that p and it is for the sake of wanting that p that you want to want that p, and your want that p is externally conditional if you want to want that p and it is for the sake of something other than wanting that p that you want to want that p. A want that p can be either internally unconditional, or internally conditional, or both, or neither. And a want that p can be either externally unconditional, or both, or neither.
- 3 'Expressive rationality' seems to have both these meanings in Hargreaves Heap et al. 1992:21–4.
- 4 Nozick 1993:27.
- 5 Nozick 1993:29.
- 6 Why not refine the basic sense of 'rational action' with the aid of causal decision theory, rather than a theory which uses just simple conditional probabilities? I have no firm objection to the alternative route, and nothing in the rest of the chapter seems to hang on the choice. For a defence of simple conditional probabilities, see Eells 1982. For an attempt to combine both types of decision theory see Nozick 1993: ch. 2.
- 7 Strawson 1986: v.
- 8 Strawson 1986:28–9. For another statement of the argument, see Strawson 1994.
- 9 Kane 1996.
- 10 The key to their resolution lies I think in Kane's insistence (1996:174–9) that there is a perfectly good sense in which an action can be non-deterministically explained in terms of the agent's motives and character without its being true that its probability, given these mental factors, is greater than the probability given those same factors of his acting in a different way.

- 11 For a treatment of the doxastically regressive possibility, see ch. VIII, note 2.
- 12 Someone may say that since each action open to you does indeed have an infinite number of outcomes, each with a finite or infinite U-subset, each action open to you will have an infinite amount of basic weighted utility, and hence no action open to you will have more basic weighted utility than any other, and hence each action open to you will be basically rational. I doubt that the conclusion follows. For some ways to avoid it, see Vallentyne 1995.
- 13 If each agent puts the probability of the other's hawking at x , and the probability of the other's diving at $1-x$, then for each agent the subjective expected utility of his hawking is $(x+3)(1-x)$ and the subjective expected utility of his diving is $(x+2)(1-2x)$.
- 14 Would it even shake a desire to act on the rule 'maximise your subjective expected utility' (SEU)? If one knew that everyone else was going to act on a rule X such that if everyone acted on it each person would have more utility than would be the case if one acted on SEU, then one might well oneself prefer to act on rule X rather than on SEU. But why should one prefer to act on rule X in the absence of such knowledge about alien actions? And how is such knowledge to be got? If one did prefer to act on rule X without having such knowledge, one would have to face the possibility that if one acted on it when other people did not, then one would have less utility than if one hadn't acted on it.
- 15 Williams 1973a: 95.

VIII ARGUMENTS AGAINST RATIONAL BELIEF

- 1 This is a thought which first emerged in section (2) of chapter I, when I was discussing the properly 'existential' attitude to one's own attitudes.
- 2 In the last chapter, at the end of the section on infinitely regressive practical rationality, I mentioned the possibility that the sceptic could construct a sound and negation-proof argument against volitionally rational belief by taking 'rational action' in a doxastically rather than desideratively regressive sense. He might in other words take 'rational action' in a sense on which the agent acts rationally only if he has performed an infinite number of belief-related actions. We might want to act rationally, in some such sense, and yet be unable to do so. The desire and the impossibility would transfer themselves to the corresponding volitional interpretation of 'rational belief'.

Suppose that just two actions are open to you: A and B. You may want either (1) to believe truly that your Aing would be basically rational (B:RA), and to A, or (2) to believe truly that your Bing would be basically rational (B:RB), and to B. It will then occur to you that if B:RA you would be free either to cause yourself to continue to believe that RA (KB:RA) or to cause yourself to stop believing that RA (SB:RA). But then you will want either (1.1) to believe truly that your (KB:RA) ing would be basically rational (B:RKB:RA), and to KB:RA, or (1.2) to believe truly that your (SB:RA) ing would be basically rational (B:RSB:RA), and to SB:RA. It will also occur to you that if B:RB you would be free either to cause yourself to continue to believe that RB (KB:RB) or to cause yourself to stop believing that RB (SB:RB). But then you will want either (2.1) to believe truly that your (KB:RB) ing would be basically rational (B:RKB:RB), and to KB:RB, or (2.2) to believe truly that your (SB:RB) ing would be basically rational (B:RSB:RB), and to SB:RB. So, as well as wanting either to A or to B, you will want to perform another action. You will want either to KB:RA, or to SB:RA, or to KB:RB, or to SB:RB. It seems, however, that the series which we have now begun has no non-arbitrary stopping place. It will occur to you that if (B:RKB:RA) you would be free either to cause yourself to

continue to believe that RKB:RA (KB:RKB:RA) or to cause yourself to stop believing that RKB:RA (SB:RKB:RA). It will occur to you that if B:RSB:RA you would be free either to cause yourself to continue to believe that RSB:RA (KB:RSB:RA) or to cause yourself to stop believing that RSB:RA (SB:RSB:RA). And so on. And then you will want to perform not just two actions but three. You will want either to A or to B; and either to KB:RA, or to SB:RA, or to KB:RB, or to SB:RB; and either KB:RKB:RA, or to SB:RKB:RA, or to KB:RSB:RA or to SB:RSB:RA, or to...and so on. As the series continues there are more and more actions that you want to perform. Reflecting about the arbitrariness of any particular stopping place, you may be left with the unsatisfiable desire to perform an infinite number of actions.

A way of scotching this sceptical possibility is suggested by what I have said about believing in the basic rationality of causing yourself to continue to believe. Suppose you believe that

(HC) For almost any S, if S wonders whether it would be basically rational for him to cause himself to continue to believe the truth then, other things being equal, he will believe that it would be.

You are then likely to believe that if you believe truly that your Aing would be basically rational, then on considering whether to cause yourself to continue to have this belief about your Aing (KB:RA) or to stop having this belief about your Aing (SB:RA), then you will believe that it is basically rational to keep the belief. But that means that you will be no less likely to A if you don't even consider whether to continue or stop having your belief that Aing would be basically rational. And in the same way, you are likely to believe that if you believe truly that your Bing would be basically rational, you will be no less likely to B if you don't consider whether to continue or stop having your belief that Bing would be basically rational (whether to KB:RB or SB:RB). Why want either to KB:RA, or to SB:RA, or to KB:RB, or to SB:RB, if performing this extra action makes no difference to the likelihood either of your Aing or of your Bing? So far from wanting to perform an infinite number of actions, you may well be content just either to A or to B.

- 3 Richard Foley's treatment of epistemically rational belief prompts similar questions. For Foley,

a judgment to the effect that an individual S's belief p is rational, like all judgments of rationality, is a judgment that is made from some point of view, where a point of view is defined by a goal (or goals) and a perspective. Specifically, a judgment that S's belief p is rational is a judgment made from some particular perspective that S's believing that p is an effective means to some particular goal (or goals) [of S].

(Foley 1987:139)

The perspective may be radically objective, in which case the judgement will be true if and only if S's believing that p really is an effective means. It may be radically subjective: the judgement will be true if and only if someone believes that S's believing that p is an effective means. It may be reflectively subjective, in which case the judgement will be true if and only if someone would on reflection believe that S's believing that p would be an effective means. If subjective, the perspective may be either egocentrically or sociocentrically subjective: in the former case it is S himself who must believe that his believing that p is an effective means. Finally, we can classify judgements of rationality according to those of S's goals in relation to which the effectiveness of the means is assessed. So, for example, if the only goals in question are epistemic, then we have a judgement of

epistemic rationality. Foley gives an extremely clear and careful account of the conditions under which it will be epistemically rational for S to believe that p, given that the perspective is egocentric and reflectively subjective and the epistemic goal is now to believe the propositions which are true and now not to believe the propositions which are false. But why should the agent not have longer-term epistemic goals? And how does Foley's account help the real deliberator, who must choose between his options in the light of *all* his goals, epistemic and non-epistemic?

- 4 Or *someone* says he should. Pascal himself did not believe that there was no good evidence for theism. See Quinn 1994:69.
- 5 Levy 1979:214.
- 6 Mackie 1982:203.
- 7 By Clifford's famous principle, it is 'wrong everywhere, always, and for everyone to believe anything upon insufficient evidence' (1901:175). (A) endorses something on these lines, but only so far as concerns starting to believe.
- 8 Plantinga 1993b: 183.
- 9 As Plantinga says, foundationalism and coherentism are normative or evaluative doctrines about 'noetic structures', where 'S's noetic structure is the set of propositions he believes, together with certain epistemic relations that hold among him and them' (Plantinga 1993a: 72). Foundationalism and coherentism are theses or groups of theses

about how a system of beliefs *ought* to be structured, about the properties of a correct, or acceptable, or rightly structured system of beliefs. The normativity in question could be *deontological*: one who conforms to his intellectual duties, on this suggestion, will be such that his noetic structure satisfies the theses. The normativity could be *axiological*: the state of affairs consisting in the existence of a noetic structure that satisfies the theses in question is intrinsically valuable. The normativity could be *aretaic*: there are (intrinsically or extrinsically) valuable noetic or intellectual states; there are also the corresponding intellectual virtues, the habits of acting that produce or promote or enhance those valuable states of affairs; and the noetic structure of a person with the appropriate intellectual virtues will satisfy the theses in question. The normativity in question could be understood in terms of what an idealised human being would be like, as in some of the Bayesian literature... Finally the normativity could be 'functional'...: the sort involved when we say of a diseased heart or knee or immune system that it isn't functioning properly, isn't working the way it ought to work; and the claim would be that a properly functioning epistemic structure... would satisfy the theses in question.

(Plantinga 1993a: 72–3)

IX THE AFFIRMATION TEST

- 1 See chapter I, section (2).
- 2 Kierkegaard 1941:208.
- 3 Kierkegaard 1941:209.
- 4 Kierkegaard 1941:209.

APPENDIX: SEMANTIC OBJECTIONS TO
DEMONIC ARGUMENTS

- 1 two of the most influential movements in the philosophy of mind, behaviourism and functionalism, have encouraged us to view the mental through the concepts of input and output. A natural concomitant of this, part cause part effect, has been a climate highly unfavourable to subjective inner states in which the fact—as I strongly suspect—that the arguments about them are all holding each other up by their bootlaces can easily remain undetected amidst the general low visibility. In that case what is at stake here is far more than some relatively technical question about what elements are admissible in a theory of meaning: it is our estimate of where the highlights should fall in our picture of the human being, whether on what it feels like from within, or on how it works when viewed from without. That is the kind of contrast of attitude which can affect an entire culture.
- (Craig 1991:280)
- 2 In this reformulation, I follow Brueckner 1994.
- 3 Putnam 1981:14.
- 4 Putnam 1981:15.
- 5 Brueckner has made the point that its conclusion is merely metalinguistic in character. Its conclusion is not that we are not brains in a vat, but only that no truth is expressed by ‘We are brains in a vat’ whether this is taken as an English or as a vat-English sentence. And there are difficulties in supposing that vat-English is composed of sentences with disquotational truth conditions which allow us to move from the falsity of ‘I am a brain in a vat’ to the claim that I am not a brain in a vat. Brueckner tries to meet this difficulty in his 1992.
- 6 Crane 1991:18.
- 7 Burge does consider, as an objection to his argument, that Alf’s mistake, in the initial situation, should be ‘construed as purely a metalinguistic mistake’ (1979:96). But he makes things easy for himself by assuming that the objector denies that Alf has any ‘object-level thought-content’ at all.
- 8 Putnam 1975:220.
- 9 Searle 1983:207–8.
- 10 Zemach 1976; Mellor 1991.
- 11 Sterelny attempted to dismiss it by modifying Putnam’s thought experiment.

The story is the same as before, with this difference. H₂O, if any were to exist on Twin Earth, would be fatal and foul tasting to Twin Earthians. Conversely, if there were XYZ on Earth, it would be fatal and foul tasting to us. In these circumstances, surely not even the most stubborn could feel any temptation to say extension of ‘water_E’=extension of ‘water_{TE}’=H₂O v XYZ.

(Sterelny 1996:101)

Surely?

WORKS CITED

- Alston, W. (1989) 'Concepts of Epistemic Justification', in *Epistemic Justification*, Ithaca, NY: Cornell University Press.
- Alston, W. (1991) *Perceiving God*, Ithaca, NY: Cornell University Press.
- Alston, W. (1993) 'Epistemic Desiderata', *Philosophy and Phenomenological Research*, vol. 53, pp. 527–51.
- Alston, W. (1996) 'Some (Temporarily) Final Thoughts on Evidential Arguments from Evil', in D.Howard-Snyder (ed.) *The Evidential Argument from Evil*, Bloomington, IN: Indiana University Press.
- Ayer, A.J. (1956) *The Problem of Knowledge*, London: Macmillan.
- Barry, B. (1995) *Justice as Impartiality*, Oxford: Clarendon Press.
- Blackburn, S. (1985) 'Supervenience Revisited', in I.Hacking (ed.) *Exercises in Analysis*, Cambridge: Cambridge University Press.
- Blackburn, S. (1992) 'Morality and Thick Concepts', *Proceedings of the Aristotelian Society*, supp. vol. 66, pp. 85–299.
- Blackburn, S. (1993) *Essays on Quasi-Realism*, New York: Oxford University Press.
- Blackburn, S. (1998) *Ruling Passions*, Oxford: Clarendon Press.
- Blumenfeld, D. (1978) 'On the Compossibility of the Divine Attributes', *Philosophical Studies*, vol. 34, pp. 91–103.
- Bonjour, L. (1985) *The Structure of Empirical Knowledge*, Cambridge, MA: Harvard University Press.
- Brandt, R. (1990) 'The Concept of Rational Action', in P.Moser (ed.) *Rationality in Action*, Cambridge: Cambridge University Press.
- Broad, C.D. (1933) *Examination of McTaggart's Philosophy*, vol. I, Cambridge: Cambridge University Press.
- Broad, C.D. (1962) *Lectures on Psychological Research*, London: Routledge.
- Brueckner, A. (1991) 'Unfair to Nozick', *Analysis*, vol. 51, pp. 61–4.
- Brueckner, A. (1992) 'Semantic Answers to Scepticism', *Pacific Philosophical Quarterly*, vol. 73, pp. 200–19.
- Brueckner, A. (1994) 'The Structure of the Sceptical Argument', *Philosophy and Phenomenological Research*, vol. 54, pp. 827–35.
- Burge, T. (1979) 'Individualism and the Mental', *Midwest Studies in Philosophy*, vol. 4, pp. 73–121.

WORKS CITED

- Clark, S. (1994) 'Companions on the Way', *Philosophical Quarterly*, vol. 44, pp. 90–100.
- Clifford, W.K. (1901) 'The Ethics of Belief', in *Lectures and Essays*, vol. II, London: Macmillan.
- Cohen, J. (1992) *An Essay on Belief and Acceptance*, Oxford: Clarendon Press.
- Cornman, J. (1975) *Perception, Common Sense and Science*, New Haven, CT: Yale University Press.
- Craig, E. (1990) *Knowledge and the State of Nature*, Oxford: Clarendon Press.
- Craig, E. (1991) 'Advice to Philosophers: Three New Leaves to Turn Over', *Proceedings of the British Academy*, vol. 76, pp. 265–81.
- Crane, T. (1991) 'All the Difference in the World', *The Philosophical Quarterly*, vol. 41, pp. 1–25.
- DeRose, K. (1995) 'Solving the Sceptical Problem', *The Philosophical Review*, vol. 104, pp. 1–52.
- Draper, P. (1989) 'Pain and Pleasure: An Evidential Problem for Theists', *Nous*, vol. 23, pp. 331–50.
- Draper, P. (1996) 'The Sceptical Theist', in D.Howard-Snyder (ed.) *The Evidential Argument from Evil*, Bloomington, IN: Indiana University Press.
- Dretske, F. (1970) 'Epistemic Operators', *Journal of Philosophy*, vol. 67, pp. 1007–23.
- Eddington, Sir A. (1935) *The Nature of the Physical World*, London: Dent.
- Eells, E. (1982) *Rational Decision and Causality*, Cambridge: Cambridge University Press.
- Foley, R. (1987) *The Theory of Epistemic Rationality*, Cambridge, MA: Harvard University Press.
- Foster, J. (1982) *The Case for Idealism*, London: Routledge.
- Fumerton, R. (1995) *Metaepistemology and Scepticism*, Lanham, MD: Rowman and Littlefield.
- Gibbard, A. (1990) *Wise Choices, Apt Feelings*, New York: Oxford University Press.
- Hale, B. (1993) 'Can there be a Logic of Attitudes?', in J.Haldane and C.Wright (eds) *Reality, Representation, and Projection*, New York: Oxford University Press.
- Hargreaves Heap, S., Hollis, M., Lyons, B., Sugden, R. and Weale, A. (1992) *The Theory of Rational Choice*, Oxford: Blackwell.
- Horwich, P. (1993) 'Gibbard's Theory of Norms', *Philosophy and Public Affairs*, vol. 22, pp. 67–78.
- Horwich, P. (1994) 'The Essence of Expressivism', *Analysis*, vol. 54, pp. 19–20.
- Howard-Snyder, D. (ed.) (1996) *The Evidential Argument from Evil*, Bloomington, IN: Indiana University Press.
- Hume, A. (1978) *Treatise of Human Nature*, Selby-Bigge and Nidditch edition, Oxford: Clarendon Press.
- Jackson, F. (1977) *Perception*, Cambridge: Cambridge University Press.
- Jackson, F., Oppy, G. and Smith, M. (1994) 'Minimalism and Truth Aptness', *Mind*, vol. 103, pp. 287–302.
- Jackson, F. (1998) *From Metaphysics to Ethics*, Oxford: Clarendon Press.
- Kane, R. (1996) *The Significance of Free Will*, New York: Oxford University Press.
- Kenny, A. (1979) *The God of the Philosophers*, Oxford: Clarendon Press.
- Kierkegaard, S. (1941) *Concluding Unscientific Postscript* (tr. David F.Swenson and Walter Lowrie), London: Oxford University Press.

WORKS CITED

- Leibniz, G.W. (1973) 'Principles of Nature and Grace' in *Philosophical Writings*, Parkinson edition, London: Dent.
- Lewis D. (1989) 'Dispositional Theories of Value', *Proceedings of the Aristotelian Society*, supp. vol. 63, pp. 113–37.
- Lewis, D. (1996) 'Elusive Knowledge', *Australasian Journal of Philosophy*, vol. 74, pp. 549–67.
- Levy, P. (1979) *Moore*, London: Weidenfeld and Nicolson.
- Lewy, C. (1976) *Meaning and Modality*, Cambridge: Cambridge University Press.
- Luper-Foy, S. (ed.) (1987) *The Possibility of Knowledge*, Totowa, NJ: Rowman and Littlefield.
- McDowell, J. (1981) 'Non-Cognitivism and Rule-Following', in S.Holtzman and S.Leich (eds) *Wittgenstein: To Follow a Rule*, London: Routledge.
- McDowell, J. (1983) 'Aesthetic Value, Objectivity and the Fabric of the World', in E. Schaper (ed.) *Pleasure, Preference and Value*, Cambridge: Cambridge University Press.
- McGinn, C. (1983) *The Subjective View*, Oxford: Clarendon Press.
- McGinn, C. (1996) 'Another Look at Color', *Journal of Philosophy*, vol. 93, pp. 537–53.
- McIver, A. (1956) 'Report on the Problem "How can I think it possible that I may be mistaken?"', *Analysis*, vol. 17, pp. 25–30.
- Mackie, J. (1976) *Problems from Locke*, Oxford: Clarendon Press.
- Mackie, J. (1977) *Ethics*, Harmondsworth: Penguin.
- Mackie, J. (1982) *The Miracle of Theism*, Oxford: Clarendon Press.
- Mellor, D.H. (1991) 'Natural Kinds', in *Matters of Metaphysics*, Cambridge: Cambridge University Press.
- Miller, D. (1994) *Critical Rationalism*, La Salle, IL: Open Court.
- Moser, P.K. (1989) *Knowledge and Evidence*, Cambridge: Cambridge University Press.
- Mundle, C. (1971) *Perception: Facts and Theories*, Oxford: Oxford University Press.
- Nathan, N. (1991) 'McTaggart's Immaterialism', *The Philosophical Quarterly*, vol. 41, pp. 442–56.
- Nathan, N. (1992) *Will and World*, Oxford: Clarendon Press.
- Nathan, N. (1997) 'Self and Will', *International Journal of Philosophical Studies*, vol. 5, pp. 81–94.
- Newman, J.H. (1955) *The Grammar of Assent*, New York: Image Books.
- Nozick, R. (1981) *Philosophical Explanations*, Cambridge, MA: Bellknap Press.
- Nozick, R. (1993) *The Nature of Rationality*, Princeton, NJ: Princeton University Press.
- O'Hear, A. (1993) 'Science and Religion', *British Journal for the Philosophy of Science*, vol. 44, pp. 505–16.
- Papineau, D. (1993) *Philosophical Naturalism*, Oxford: Blackwell.
- Peacocke, C. (1983) *Sense and Content*, Oxford: Clarendon Press.
- Plantinga, A. (1993a) *Warrant: The Current Debate*, New York: Oxford University Press.
- Plantinga, A. (1993b) *Warrant and Proper Function*, New York: Oxford University Press.
- Plantinga, A. (1996a) 'Epistemic Probability and Evil', in D.Howard-Snyder (ed.) *The Evidential Argument from Evil*, Bloomington, IN: Indiana University Press.
- Plantinga, A. (1996b) 'On Being Evidentially Challenged', in D.Howard-Snyder (ed.) *The Evidential Argument from Evil*, Bloomington, IN: Indiana University Press.
- Plato (1956) *Meno* (tr. W.Guthrie), Harmondsworth: Penguin.

WORKS CITED

- Price, H. (1954) 'Belief and the Will', *Proceedings of the Aristotelian Society*, supp. vol. 28, pp. 1–26.
- Putnam, H. (1975) 'The Meaning of "Meaning"', in *Philosophical Papers*, vol. II, Cambridge: Cambridge University Press.
- Putnam, H. (1981) *Reason, Truth and History*, Cambridge: Cambridge University Press.
- Quinn, P. (1994) 'Moral Objections to Pascalian Wagering', in J.Jordan (ed.) *Gambling on God: Essays on Pascal's Wager*, Lanham, MD: Rowman and Littlefield.
- Robinson, H. (1985) 'The General Form of the Argument for Berkeleyian Idealism', in J.Foster and H.Robinson (eds) *Essays on Berkeley*, Oxford: Clarendon Press.
- Robinson, H. (1994) *Perception*, London: Routledge.
- Ross, J. (1969) *Philosophical Theology*, Indianapolis, IN: Bobbs-Merrill.
- Ross, J. (1986) 'Believing for Profit', in G.McCarthy (ed.) *The Ethics of Belief Debate*, Atlanta, GA: Scholar's Press.
- Ross, J. (1993) 'Cognitive Finality', in L.Zagzebski (ed.) *Rational Faith*, Notre Dame, IN: Notre Dame University Press.
- Rowe, W. (1975) *The Cosmological Argument*, Princeton, NJ: Princeton University Press.
- Russell, B. (1912) *The Problems of Philosophy*, London: Williams and Norgate.
- Schiffer, S. (1996) 'Contextualist Solutions to Scepticism', *Proceedings of the Aristotelian Society*, vol. 96, pp. 317–33.
- Searle, J. (1983) *Intentionality*, Cambridge: Cambridge University Press.
- Sellars, W. (1963) *Science, Perception and Reality*, London: Routledge.
- Smith, M. (1994a) 'Why Expressivists about Value Should Love Minimalism about Truth', *Analysis*, vol. 54, pp. 1–11.
- Smith, M. (1994b) *The Moral Problem*, Oxford: Blackwell.
- Sosa, E. (1991) *Knowledge in Perspective*, Cambridge: Cambridge University Press.
- Sprigge, T. (1984) 'Non-Human Rights: An Idealist Perspective', *Inquiry*, vol. 27, pp. 439–61.
- Sterelny, K. (1996) 'Natural Kind Terms', in A.Pessin and S.Goldberg (eds) *The Twin Earth Chronicles*, New York: M.E.Sharpe.
- Stove, D. (1973) *Probability and Hume's Inductive Scepticism*, Oxford: Clarendon Press.
- Stove, D. (1982) *Popper and After*, Oxford: Pergamon Press.
- Strawson, G. (1986) *Freedom and Belief*, Oxford: Clarendon Press.
- Strawson, G. (1994) 'The Impossibility of Moral Responsibility', *Philosophical Studies*, vol. 75, pp. 5–24.
- Stroud, B. (1983) *The Significance of Philosophical Scepticism*, Oxford: Clarendon Press.
- Swinburne, R. (1979) *The Existence of God*, Oxford: Clarendon Press.
- Valberg, J. (1992) *The Puzzle of Experience*, Oxford: Clarendon Press.
- Vallentyne, P. (1995) 'Infinite Utility: Anonymity and Person-Centredness', *Australasian Journal of Philosophy*, vol. 73, pp. 413–20.
- van Cleve, J. (1984) 'Reliability, Justification, and the Problem of Induction', *Midwest Studies in Philosophy*, vol. 9, pp. 555–67.
- van Fraassen, B. (1980) *The Scientific Image*, Oxford: Clarendon Press.
- van Fraassen, B. (1985) 'Empiricism in the Philosophy of Science', in P.Churchland and C.Hooker (eds) *Images of Science*, Chicago: Chicago University Press.
- Weintraub, R. (1997) *The Sceptical Challenge*, London: Routledge.
- Williams, B. (1973a) 'A Critique of Utilitarianism', in J.J.C.Smart and B.Williams, *Utilitarianism: For and Against*, Cambridge: Cambridge University Press.

WORKS CITED

- Williams, B. (1973b) 'Deciding to Believe', in *Problems of the Self*, Cambridge: Cambridge University Press.
- Williams, B. (1978) *Descartes*, Harmondsworth: Penguin.
- Williams, B. (1979) 'Internal and External Reasons', in R.Harrison (ed.) *Rational Action*, Cambridge: Cambridge University Press.
- Williams, M. (ed.) (1993) *Scepticism*, Aldershot: Dartmouth.
- Winters, B. (1979) 'Believing at Will', *Journal of Philosophy*, vol. 76, pp. 243–56.
- Wittgenstein, L. (1953) *Philosophical Investigations*, Oxford: Blackwell.
- Wittgenstein, L. (1982) *Last Writings on the Philosophy of Psychology*, vol. I, Oxford: Blackwell.
- Wright, C. (1988) 'Moral Values, Projection and Secondary Qualities', *Proceedings of the Aristotelian Society*, supp. vol. 62, pp. 1–26.
- Wright, C. (1991) 'Scepticism and Dreaming: Imploding the Demon', *Mind*, vol. 100, pp. 87–116.
- Wright, C. (1994) 'On Putnam's Proof that We Are Not Brains in Vats', in P.Clark and B.Hale (eds) *Reading Putnam*, Oxford: Blackwell.
- Zagzebski, L. (1996) *Virtues of the Mind*, Cambridge: Cambridge University Press.
- Zemach, E. (1976) 'Putnam's Theory on the Reference of Substance Terms', *Journal of Philosophy*, vol. 73, pp. 116–27.

INDEX

- acceptance 97
acquaintance 4–5, 40, 64, 68, 81–7
action at a distance 84
admiration 90–1
affirmation test 2, 137
alpha beings 72–92
Alston, W. 13, 93, 150, 157, 162, 163
analytic entailment 47
Aquinas 163
arguments from evil 64–7
arguments for ignorance 3; demonic 19–30; doubly-regressive 36–9, 43–9; evidence-or-acquaintance arguments 17, 50–93, 137–8; evidentialist 32–4, 40–3; neo-Cartesian 12, 19, 142–8, 150–1; regressive 31–49; semi-evidentialist 34–9, 43–9
arguments against justified belief 93–4, 100–7
arguments against rational belief 93, 108, 124–36
assent 96–7
Augustine 163
Ayer, A.J. 149

Barry, B. 164
basic limiting principles 86
basic weighted utility 113, 120
Bayes's Theorem 157
belief: existential attitudes towards one's own 14–16, 125, 140; in the improbable 139; shortage of about the future 119–23, 125, 128, 132; stability of 41–2, 62; and will 96–100

Berkeley, G. 74, 75, 160
Blackburn, S. 52, 154, 155, 156
Blumenfeld, D. 157
Bonjour, L. 162
Brandt, R. 164
Broad, C.D. 86, 158, 161
Brueckner, A. 21, 150, 151, 168
Burge, T. 145, 168

cake: having and eating one's 29
Cartesian dualism 74
chromatic qualativism 74, 75–8, 87, 91
Clark, S. 157
Clifford, W.K. 167
closure, *see* knowledge
Cohen, L.J. 162
coherentism 135–6
colours 73–4, 75–8, 87–8, 90–1, 158–9
consensual propositions 11, 16, 26, 45, 126
contextualism 23–4
cooperation 103–4
Cornman, J. 159
cosmological arguments 67–70
Craig, E. 150, 168
Crane, T. 168
Critical Rationalism 8

deductivism 31–6
deontological and consequentialist principles 66, 122
DeRose, K. 151
DISJUNCTION 61
dispositional properties 73–4

INDEX

- Draper, P. 157
 Dretske, F. 150
 dynamic objectivity 57–63
 dynamic unity 55
- Eddington, Sir A. 78, 159
 Eells, E. 164
 ‘eliminative materialism’ 74
 ends 109, 119, *and see* wanting,
 internally and externally
 unconditional
 ‘Enlightenment project’ 107
 ENTAILMENT 7, 26, 44; modification
 of 47–8, 62, 89, 138; objections to
 27–8, 46–7
 evaluation 51; and assertion 51; doubly
 attitudinal 54; expressivism about 53;
 purely assertive 54; purely dynamic
 54; and truth 51; universal
 conditionals in 54, 103–7
 Eventual Evidence Principle 98
 evidence 4–5, 32, 41, 65; and
 explanation 39, 76–7, 79
 Exclusion Principle 69
 expected utility 112–13, 120, 121,
 132, 165
 explanatory exclusion 67–70
- fideism 1, 139
 Foley, R. 166, 167
 Foster, J. 80–1, 159, 160
 foundationalism 40, 135–6
 freedom 19, 68, 89, 114, 119, 138
 Fumerton, R. 40, 94, 149,
 153–4, 162
- Gettier, E. 25, 94, 162
 Gibbard, A. 164
- Hale, B. 155
 Hargreaves Heap, S. 164
 Heidegger, M. 162
 Horwich, P. 155
 human diversity 54, 55, 57, 102
 Hume, D. 6, 8, 35, 112, 132, 149, 153,
 154, 156
 Husserl, E. 143
- Idealism and Anti-Idealism: defined
 71–5; knowledge of Anti-Idealism
 75–88; wanting Idealism’s falsity
 71, 89–91
 impartiality 104–5
 INCOHERENCE 26, 29, 53, 64, 75
 Induction 4–16, 35, 38–9, 45
 infinite regression, *see* knowledge;
 rationality of belief; rationality of
 action
 internalism and externalism: in
 epistemology 35, 37, 153; in ethics
 54, 155
 intrinsic properties 71–2; mental 74,
 80–1
- Jackson, F. 76, 154, 155, 159
 justified belief 93–4; and evaluation,
 95–6, 101–3; and knowledge 94–5;
 and rational belief 108; and volition
 95–100
- Kane, R. 164–5
 Kant, I. 111, 112
 Kenny, A. 157
 Kierkegaard, S. 139, 167
 knowledge: activist conditions for
 34; closure under entailment
 151, 152–3; closure under known
 entailment 21, 27, 37; ‘conjectural’
 8; doubly regressive conditions for
 37; evidentialist conditions for 31;
 evidence-or-acquaintance conditions
 for 4, 50; infallibilist conditions for
 17; intrinsic value of 14; of modality
 59–60; optional conditions for 17;
 regressive conditions for 31;
 reliabilism 9–11, 40–1; semi-
 evidentialist conditions for 34;
 stock reasons for wanting
 10–12; truth-tracking conditions
 for 21
- laws of nature 83
 Leibniz, G.W. 70, 74, 158
 Levy, P. 167
 Lewis, D. 150, 152, 155
 Lewy, C. 149, 154, 156

INDEX

- logical probability 4–5
 Luper-Foy, S. 150
- McDowell, J. 75, 154, 159
 McGinn, C. 158
 McIver, A. 15, 150
 Mackie, J. 60, 76, 132, 159, 164
 McTaggart, J. 74, 76
 manifest image 78, 89, 161
 Mellor, D.H. 147, 168
 mental properties 74, 80
 Miller, D. 8, 149
 mind-independence 72
 Missing Object Argument 81–5
 Moore, G.E. 132
 Moser, P.K. 94, 162
 multiplication of utility 105–6
 Mundle, C. 161
- Nathan, N. 159
 natural theology 67
 negation test 2, 6
 negation-proof arguments 51
 negative responsibility 122
 neo-Cartesian scepticism 19, 150–1
 Newman, J.H. 96, 162
 Nozick, R. 21, 111, 150, 164
- obligation 101
 Occasionalism 3
 Ockham's Razor 67, 76
 O'Hear, A. 157
 Oppy, G. 154
- Papineau, D. 150
 paradox 24
 Pascal's Wager 132–4, 167
 Peacocke, C. 158
 permissibility 101
 personal identity 90
 Plantinga, A. 135, 154, 157, 162, 167
 Plato 10, 150
 Popper, K. 8, 17
 practical agreement, *see* dynamic unity
 Price, H. 162
 Principle of Sufficient Reason 67, 157, 158
 Prisoner's Dilemma 121–2
 psychologism about modality 59
- Putnam, H. 144, 146, 150, 168
- question-beggingness 4, 33, 39, 59, 66,
 79; and vicious circularity 10
 Quinn, P. 167
- rationality of action 109; basic 110,
 112–13, 124; and belief shortage
 119–23; desideratively regressive
 113–19; doxastically regressive 112,
 165–6; expressive 110; instrumental
 109–10; non-self-mediated 122;
 symbolic 111
 rationality of belief 93, 108; epistemic
 129–30, 166–7; and evaluation 108;
 starting and continuing to believe
 124–5; and volition 108
- Reason 35, 109, 153
 relational properties 72–3
 risking false belief 130, 139–41
 Robinson, H. 81, 83, 160, 161
 Ross, J. 157, 162, 163
 Rowe, W. 157
 rule following considerations 23, 142–3
 Russell, B. 74, 77, 159
- Sartre, J.-P. 14
 sceptical and non-sceptical arguments:
 for ignorance 3, 137; against justified
 belief 102; against rational belief 126
 Schiffer, S. 152
 scientific image 89
 scientific realism 78, 79
 Searle, J. 147, 168
 Sellars, W. 78, 91, 160, 161
 semantic externalism 23, 144–8
 Smith, M. 154, 155, 156
 solidarity 104
 Sosa, E. 33
 space 80–1
 Sprigge, T. 162
 Sterelny, K. 168
 stoicism 1, 23
 Stove, D. 8, 149, 153
 Strawson, G. 114, 164
 Stroud, B. 151
 Swinburne, R. 68, 157

INDEX

- terminological eccentricity 5–6, 21–2,
24, 32
to-be-pursuedness 60
'track-record' arguments 13
truth-tracking 21–3
twin earth 146–8
- UNBELIEF 7–8, 149
- Valberg, J. 81, 160, 161
Vallentyne, P. 165
van Cleve, J. 150
van Fraassen, B. 67, 79, 157, 161–2
- wanting 149; internally and externally
unconditional 109, 164; rationality of
112, *and see* rationality of action,
desideratively regressive;
conjunctions and their conjuncts as
objects of 29–30, 45; wanting to want
30; destruction of 62, 106, 156–7, *and
see* DISJUNCTION, ENTAILMENT,
INCOHERENCE
- Weintraub, R. 152
Williams, B. 13, 98, 150, 162, 164, 165
Williams, M. 150
Winters, B. 163
Wittgenstein, L. 15, 142, 150
Wright, C. 9, 150, 151, 158
- Zagzebski, L. 162
Zemach, E. 147, 168